

AoI-optimal Joint Sampling and Updating for Wireless Powered Communication Systems

Mohamed A. Abd-Elmagid, Harpreet S. Dhillon, and Nikolaos Pappas

Abstract—This paper characterizes the structure of the Age of Information (AoI)-optimal policy in wireless powered communication systems while accounting for the time and energy costs of generating status updates at the source nodes. In particular, for a single source-destination pair in which a radio frequency (RF)-powered source sends status updates about some physical process to a destination node, we minimize the long-term average AoI at the destination node. The problem is modeled as an average cost Markov Decision Process (MDP) in which, the generation times of status updates at the source, the transmissions of status updates from the source to the destination, and the wireless energy transfer (WET) are jointly optimized. After proving the monotonicity property of the value function associated with the MDP, we analytically demonstrate that the AoI-optimal policy has a *threshold-based* structure w.r.t. the state variables. Our numerical results verify the analytical findings and reveal the impact of state variables on the structure of the AoI-optimal policy. Our results also demonstrate the impact of system design parameters on the optimal achievable average AoI as well as the superiority of our proposed *joint sampling and updating policy* w.r.t. the *generate-at-will* policy.

I. INTRODUCTION

AoI provides a rigorous way of quantifying the freshness of information about a physical process at a destination node based on the status updates it receives from a source node [1]. In [2], AoI was first defined as the time elapsed between the generation of a status update at the source and its reception at the destination. Since then, AoI has been extensively used to quantify the performance of various communication networks that deal with time-sensitive information, including multi-hop networks [3], multicast networks [4], broadcast networks [5], [6], and ultra-reliable low-latency vehicular networks [7]. Interested readers are advised to refer to [8], [9] for comprehensive surveys.

Recently, the concept of AoI has been argued to have an important role in designing *freshness-aware Internet of Things* (IoT) networks (which can enable a broad range of real-time applications) [10]–[14]. A common assumption in most of the literature on AoI is to neglect the costs of generating status updates, however, the IoT devices (source nodes in the context of AoI setting) are currently expected to perform *sophisticated tasks* while generating status updates [12], [15]. In that sense, it is crucial to incorporate the energy and time costs of generating status updates in the design of future freshness-aware IoT networks. To further enable a sustainable operation

Copyright (c) 2020 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

M. A. Abd-Elmagid and H. S. Dhillon are with Wireless@VT, Department of ECE, Virginia Tech, Blacksburg, VA (Email: {maelaziz, hdhillon}@vt.edu). N. Pappas is with the Department of Science and Technology, Linköping University, SE-60174 Norrköping, Sweden (Email: nikolaos.pappas@liu.se). The support of the U.S. NSF (Grant CPS-1739642) is gratefully acknowledged.

of such networks, RF energy harvesting has emerged as a promising solution for charging low-power IoT devices [16]. In particular, the ubiquity of RF signals even at hard-to-reach places makes them more suitable to power IoT devices than other potential sources of energy harvesting, such as solar or wind. In addition, the implementation of RF energy harvesting modules is usually cost efficient, which is another important aspect of the deployment of IoT devices. The main focus of this paper is on investigating the structural properties of the AoI-optimal joint sampling and updating policy for freshness-aware RF-powered IoT networks.

The AoI-optimal policy for an energy harvesting source has already been investigated under various system settings [17]–[24]. The energy harvesting process is commonly modeled as an independent external stochastic process. However, when the source is assumed to be RF-powered, the harvested energy depends on the channel state information (CSI) and its variation over time, which makes the characterization of the AoI-optimal policies very challenging. It is worth noting that [25]–[27] have very recently explored the AoI-optimal policy in wireless powered communication systems. However, none of the proposed policies took into account the time and energy costs of generating status updates at the source. In addition, [25] and [26] did not incorporate the evolution of the battery level at the source and the variation of CSI over time in the process of decision-making. This paper makes the first attempt to analytically characterize the structural properties of the AoI-optimal joint sampling and updating policy while: *i*) *considering the dynamics of battery level, AoI, and CSI, and* *ii*) *accounting for the costs of generating status updates in the process of decision-making*.

Contributions. Our main contribution is the analytical characterization of the structure of the AoI-optimal policy for an RF-powered single source-destination pair system setup while incorporating the time and energy costs for generating status updates at the source. In particular, we model the problem as an average cost MDP¹ with finite state and action spaces for which its corresponding value function is shown to be a monotonic function w.r.t. the state variables. Using this property, the AoI-optimal policy is proven to have a threshold-based structure w.r.t. different state variables. Our numerical results verify our analytical findings and reveal the impact of state variables as well as the energy required for generating a status update at the source on the structure of the AoI-optimal policy. Our results also demonstrate the superiority of our proposed joint sampling and updating policy over the generate-at-will policy in terms of the achievable average AoI.

¹The theory of MDPs is useful for problems in which the objective is to obtain an optimal mapping between the system state and action spaces.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. Network Model

We consider a single source-destination pair model in which the source contains: i) a sensor that keeps sampling the real-time status of a physical process and ii) a transmitter that sends status update packets about the observed process to the destination, as shown in Fig. 1. Since the single source-destination pair model may actually be sufficient to study a diverse set of applications [2] (e.g., safety of an intelligent transportation system, predicting and controlling forest fires, and efficient energy utilization in future smart homes), our analysis in this paper will be of interest in many applications. The scenario of having multiple source nodes is left as a promising direction of future work.

We assume that the source node may perform sophisticated sampling tasks, e.g., initial feature extraction and pre-classification using machine learning tools [12]. Hence, unlike most of the existing literature, the time and energy costs of generating an update packet at the source node cannot be neglected. While the destination node is assumed to be always connected to the power grid, the source node is powered through WET by the destination node. Particularly, the destination node transmits RF signals in the downlink to charge the source node. The energy harvested by the source node is then stored in its battery, which has a finite capacity of B_{\max} joules. The source and destination nodes share the same channel and they have a single antenna each. Hence, the source can either harvest energy or transmit data at a given time instant.

We assume discrete time and the time slots are of equal size. Let $B(n)$, $A(n)$ and $\tau(n)$ denote the amount of available energy in the battery at the source node, the AoI at the destination node, and the time passed since the generation instant of the current update packet available at the source node (i.e., the AoI of the status updates at the source node), respectively, at the beginning of time slot n . Denote by $h(n)$ and $g(n)$ the uplink and downlink channel power gains between the source and destination nodes over slot n , respectively. We assume that the channels are influenced by quasi-static flat fading. This, in turn, means that the channels are fixed over a time slot, and independently vary from one slot to another.

B. State and Action Spaces

The state of the system at slot n can be expressed as $s(n) \triangleq (B(n), A(n), \tau(n), h(n), g(n)) \in \mathbb{S}$; where \mathbb{S} is the state space which contains all the combinations of the system state variables. We also assume that the state variables can take discrete values² and obtain a lower bound to the performance of the continuous system (as it will be clear in the sequel). In particular, we have $B(n) \in \{0, 1, \dots, b_{\max}\}$ where b_{\max} denotes the battery capacity, such that each energy

²Note that constructing a finite state space of an MDP by discretizing the state variables and/or defining upper bounds on their maximum values is very common in the literature to obtain the optimal policy numerically as well as characterize its structure properties analytically using standard optimization techniques such as the Value Iteration Algorithm (VIA) or Policy Iteration Algorithm (PIA). See [12], [22], [23], [28] for representative examples.

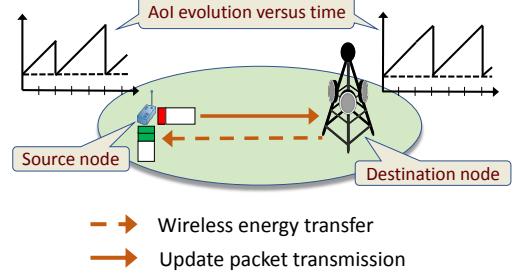


Fig. 1. An illustration of the system setup.

quantum in the battery is equivalent to $e_q = \frac{B_{\max}}{b_{\max}}$ joules. Note that both the energy consumed from the battery for an update packet transmission and the harvested energy need to be expressed in terms of the energy quanta. In addition, if the channel power gains are originally modeled using continuous random variables, we discretize them into a finite number of intervals whose probabilities are determined from the probability density function (PDF) of the fading gain. In particular, each interval is then represented by a discrete level of channel power gain which has the same probability as that of this interval. Without loss of generality, we also assume that $A(n)$ ($\tau(n)$) is upper bounded by a finite value A_{\max} (τ_{\max}) which can be chosen to be arbitrarily large [12], [22]. This assumption is quite reasonable since, in practice, the protocols will be designed such that a new packet is transmitted by the device if the monitor has not received an update over a certain observation window because otherwise the information will become too stale. This naturally leads to a bound on AoI where A_{\max} can be set to the length of the observation window. Based on $s(n)$, two actions are decided at slot n : i) the first action $a_1(n) \in \mathbb{A}_1 \triangleq \{S, I\}$ determines whether the source generates a new update packet in slot n or not, and ii) the second action $a_2(n) \in \mathbb{A}_2 \triangleq \{T, H\}$ determines whether slot n is allocated for an update packet transmission from the source to the destination or WET by the destination. Specifically, when $a_1(n) = S$, a new update packet is generated by the source, which replaces the currently available one, if any, since there is no benefit of sending out-of-date packets to the destination. We also consider that generating an update packet takes one time slot (as a *time cost*) and requires an amount of energy E^S (as *energy cost* expressed in energy quanta). When $a_2(n) = T$, the source sends its currently available packet (that was generated from $\tau(n)$ time slots) to the destination. The required energy for a packet transmission of size M bits in slot n , according to Shannon's formula, is $E^T(n) = \frac{\sigma^2}{h(n)} (2^{M/W} - 1)$, where σ^2 is the noise power at the destination and W is the channel bandwidth. When $a_2(n) = H$, slot n is allocated for WET by the destination to charge the battery at the source. We consider a practical non-linear energy harvesting model [29] such that the energy harvested by the source is given by

$$E^H(n) = \frac{P_{\max} (1 - \exp[-a P_{\text{rec}}(n)])}{1 + \exp[-a (P_{\text{rec}}(n) - b)]}, \quad (1)$$

where a and b are constants representing the steepness and the inflexion point of the curve that describes the input-output power conversion, P_{\max} is the maximum power that can be harvested through a particular circuit configuration, and $P_{\text{rec}}(n) = P_t g(n)$ such that P_t is the average transmit power by the destination. Hence the system action at slot n can be expressed as $a(n) = (a_1(n), a_2(n)) \in \mathbb{A} = \mathbb{A}_1 \times \mathbb{A}_2$, where \mathbb{A} is the action space of the system.

Note that the system state is assumed to be available at the destination node at the beginning of each time slot to take decisions. In particular, we assume that the location of the source node is known *a priori*, and hence the average channel power gains are pre-estimated and known at the destination node. In particular, at the beginning of an arbitrary time slot, the destination node has perfect knowledge about the channel power gains in that slot, and only a statistical knowledge for future slots [28]. Further, given some initial values for the remaining system state parameters (i.e., $B(0)$, $\tau(0)$ and $A(0)$), the destination node updates their values based on the action taken at each time slot. More specifically, $B(n+1)$ can be expressed as a function of the system action at slot n ($a(n)$) as

$$\begin{cases} B(n) - \lceil E^T(n)/e_q \rceil, & \text{if } a(n) = (I, T), \\ B(n) - E^S - \lceil E^T(n)/e_q \rceil, & \text{if } a(n) = (S, T), \\ \min \{b_{\max}, B(n) + \lfloor E^H(n)/e_q \rfloor\}, & \text{if } a(n) = (I, H), \\ \min \{b_{\max}, B(n) - E^S + \lfloor E^H(n)/e_q \rfloor\}, & \text{if } a(n) = (S, H), \end{cases} \quad (2)$$

where we used the ceiling and floor with $E^T(n)$ and $E^H(n)$, respectively. Thus, we obtain a lower bound to the performance of the original continuous system. Let $\mathbb{A}(s(n))$ denote the action space associated with state $s(n)$, i.e., $\mathbb{A}(s(n))$ contains the possible actions that can be taken at $s(n)$. We assume that $a(n) \in \mathbb{A}(s(n))$ only if $B(n)$ is greater than the energy required for taking action $a(n)$, hence we always have $B(n+1) \geq 0$. Furthermore, $A(n+1)$ and $\tau(n+1)$ can be expressed, respectively, as

$$A(n+1) = \begin{cases} \min \{A_{\max}, \tau(n) + 1\}, & \text{if } a(n) = (a_1(n), T) \\ \min \{A_{\max}, A(n) + 1\}, & \text{otherwise.} \end{cases} \quad (3)$$

$$\tau(n+1) = \begin{cases} 1, & \text{if } a(n) = (S, a_2(n)), \\ \min \{\tau_{\max}, \tau(n) + 1\}, & \text{otherwise,} \end{cases} \quad (4)$$

where $a(n) = (a_1(n), T)$ means that $a(n) \in \{(I, T), (S, T)\}$. This also applies to $(S, a_2(n))$ w.r.t. $a_2(n)$.

C. Problem Formulation

A policy is a mapping from the system state space to the system action space. Under a policy π , the long-term average AoI at the destination with initial state $s(0)$ is given by

$$\bar{A}^{\pi} \triangleq \limsup_{N \rightarrow \infty} \frac{1}{N+1} \sum_{n=0}^N \mathbb{E}[A(n) | s(0)]. \quad (5)$$

We take the expectation w.r.t. the channel conditions and policy in (5). We then aim at finding the policy π^* that

achieves the minimum average AoI, i.e., $\pi^* = \arg \min_{\pi} \bar{A}^{\pi}$. Owing to the independence of channel power gains over time and the nature of the dynamics of remaining state variables, as described by (2)-(4), the problem can be modeled as an MDP. Recall that the system state space is finite (the state variables are discretized) and the system action space is clearly finite as well. In this case, the MDP at hand is a *finite-state finite-action MDP*, for which there exists an optimal stationary deterministic policy (i.e., we take a deterministic action at each state that is fixed over time) that can be obtained using the VIA or PIA [30]. Therefore, in the sequel, we omit the time index and explore this stationary deterministic policy. In the next section, we characterize the AoI-optimal policy π^* and derive its structural properties.

III. ANALYSIS OF THE AOI-OPTIMAL POLICY

A. Optimal Policy Characterization

Given a stationary deterministic policy π , the probability of moving from state $s = (B, A, \tau, h, g)$ to state $s' = (B', A', \tau', h', g')$ can be expressed as

$$\begin{aligned} \mathbb{P}(s' | s, \pi(s)) &\triangleq \mathbb{P}(B', A', \tau', h', g' | B, A, \tau, h, g, \pi(s)) \\ &\stackrel{(a)}{=} \mathbb{P}(B', A', \tau' | B, A, \tau, h, g, \pi(s)) \mathbb{P}(h') \mathbb{P}(g') \\ &\stackrel{(b)}{=} C \mathbb{P}(B' | B, h, g, \pi(s)) \mathbb{P}(A' | A, \tau, \pi(s)) \mathbb{P}(\tau' | \tau, \pi(s)), \end{aligned} \quad (6)$$

where $\pi(s)$ denotes the action taken at state s according to π , $\mathbb{P}(h')$ and $\mathbb{P}(g')$ denote the probability mass functions for the uplink and downlink channel power gains, and $C = \mathbb{P}(h') \mathbb{P}(g')$. Step (a) follows since the channel power gains are independent over time from each other and from other random variables. Note that for the case of a Markovian fading channel model, the conditional probabilities $\mathbb{P}(h' | h)$ and $\mathbb{P}(g' | g)$ will replace $\mathbb{P}(h')$ and $\mathbb{P}(g')$, respectively. These conditional probabilities are determined according to the Markovian fading channel model considered in the problem. However, all our analytical results regarding the structure of the AoI-optimal policy (derived in the next subsection) will remain the same. Step (b) follows due to the fact that given s and $\pi(s)$, we can obtain B' , A' and τ' in a deterministic way separately from each other using (2)-(4). The optimal policy π^* can be characterized using following Lemma.

Lemma 1. *The policy π^* can be obtained by solving the following Bellman's equation for average cost MDPs [30]*

$$\bar{A}^* + V(s) = \min_{a \in \mathbb{A}(s)} Q(s, a), s \in \mathbb{S}, \quad (7)$$

where $V(s)$ is the value function, \bar{A}^* is the achievable average AoI by π^* which is independent of the initial state $s(0)$, and $Q(s, a)$ is the expected cost due to taking action a in state s , which is given by

$$Q(s, a) = A + \sum_{s' \in \mathbb{S}} \mathbb{P}(s' | s, a) V(s'), \quad (8)$$

where $\mathbb{P}(s' | s, a)$ can be computed using (6). In addition, the optimal action taken at state s can be evaluated as

$$\pi^*(s) = \arg \min_{a \in \mathbb{A}(s)} Q(s, a). \quad (9)$$

The value function $V(s)$ can be obtained iteratively using the VIA [30]. Particularly, according to the VIA, the value function at iteration k , $k = 1, 2, \dots$, is evaluated as

$$\begin{aligned} V(s)^{(k)} &= \min_{a \in \mathbb{A}(s)} Q(s, a)^{(k-1)} \\ &= \min_{a \in \mathbb{A}(s)} \left\{ A + \sum_{s' \in \mathbb{S}} \mathbb{P}(s' | s, a) V(s')^{(k-1)} \right\}, \end{aligned} \quad (10)$$

where $s \in \mathbb{S}$. Hence, $\pi^*(s)$ at iteration k is given by

$$\pi^*(s) = \arg \min_{a \in \mathbb{A}(s)} Q(s, a)^{(k-1)}. \quad (11)$$

Note that in each iteration of the VIA, the optimal action at each system state needs to be computed using (11) (this is referred to as the policy improvement step). Under any initialization of value function $V(s)^{(0)}$, according to the VIA, the sequence $\{V(s)^{(k)}\}$ converges to $V(s)$ which satisfies the Bellman's equation in (7), i.e.,

$$\lim_{k \rightarrow \infty} V(s)^{(k)} = V(s). \quad (12)$$

In the next subsection, we will use the VIA to explore the structural properties of π^* . Note that the obtained analytical results can be derived using Relative VIA (RVIA) as well [30].

B. Structural Properties of the Optimal Policy

Lemma 2. *The value function $V(s)$ corresponding to π^* is: (i) non-decreasing w.r.t. A and τ , and (ii) non-increasing w.r.t. B , g and h .*

Proof: We first prove that $V(B, A, \tau, h, g)$ is non-decreasing w.r.t. A . Let $s \setminus x$ denote the combination of state s variables excluding the variable x . Define $s_1 = (B_1, A_1, \tau_1, h_1, g_1)$ and $s_2 = (B_2, A_2, \tau_2, h_2, g_2)$ such that $A_1 \leq A_2$ and $s_1 \setminus A_1 = s_2 \setminus A_2$. Therefore, the goal is to show that $V(s_1) \leq V(s_2)$. Clearly, it is sufficient to show that the relation holds over all iterations of the VIA, i.e., $V(s_1)^{(k)} \leq V(s_2)^{(k)}, \forall k$. We prove that using mathematical induction as follows. For $k = 0$, the relation holds since we can choose the initial values $\{V(s)^{(0)}\}_{s \in \mathbb{S}}$ arbitrary. Now, for an arbitrary value of k , we show that having $V(s_1)^{(k)} \leq V(s_2)^{(k)}$ leads to $V(s_1)^{(k+1)} \leq V(s_2)^{(k+1)}$. From (10) and (11), $V(s_1)^{(k+1)}$ and $V(s_2)^{(k+1)}$ are given, respectively, by

$$\begin{aligned} V(s_1)^{(k+1)} &= A + \sum_{s'_1 \in \mathbb{S}} \mathbb{P}(s'_1 | s_1, \pi^*(s_1)) V(s'_1)^{(k)} \\ &\stackrel{(a)}{\leq} A + \sum_{s'_1 \in \mathbb{S}} \mathbb{P}(s'_1 | s_1, \pi^*(s_2)) V(s'_1)^{(k)} \\ &\stackrel{(b)}{=} A + C \sum_{g'_1} \sum_{h'_1} V(\bar{B}_1, \bar{A}_1, \bar{\tau}_1, h'_1, g'_1)^{(k)}, \end{aligned} \quad (13)$$

$$\begin{aligned} V(s_2)^{(k+1)} &= A + \sum_{s'_2 \in \mathbb{S}} \mathbb{P}(s'_2 | s_2, \pi^*(s_2)) V(s'_2)^{(k)} \\ &\stackrel{(b)}{=} A + C \sum_{g'_2} \sum_{h'_2} V(\bar{B}_2, \bar{A}_2, \bar{\tau}_2, h'_2, g'_2)^{(k)}, \end{aligned} \quad (14)$$

where step (a) follows since it is not optimal to take action $\pi^*(s_2)$ in state s_1 ; step (b) follows from (2)-(4)

and (6) where for a given $\pi^*(s_2)$: 1) \bar{B}_i and $\bar{\tau}_i$ are determined using (2) and (4), respectively, and 2) \bar{A}_i is evaluated from (3), $i \in \{1, 2\}$. Note that $\bar{B}_1 = \bar{B}_2$ and $\bar{\tau}_1 = \bar{\tau}_2$ for $\pi^*(s_2) \in \mathbb{A}$ since we have $B_1 = B_2$ and $\tau_1 = \tau_2$. On the other hand, since $A_1 \leq A_2$, we can observe from (3) that $\bar{A}_1 \leq \bar{A}_2$ for $\pi^*(s_2) \in \mathbb{A}$, and hence $V(\bar{B}_1, \bar{A}_1, \bar{\tau}_1, h'_1, g'_1)^{(k)} \leq V(\bar{B}_2, \bar{A}_2, \bar{\tau}_2, h'_2, g'_2)^{(k)}$. Therefore, $V(s_2)^{(k+1)}$ is greater than or equal to the expression in (13) which makes $V(s_1)^{(k+1)} \leq V(s_2)^{(k+1)}$ and indicates that the value function is non-decreasing w.r.t. A . Using the same approach, we can show that $V(B, A, \tau, h, g)$ is non-decreasing (non-increasing) w.r.t. τ (B). Finally, note that increasing h (g) reduces E^T (increases E^H), which increases the battery level at the next time slot and hence the value function is reduced. Therefore, $V(B, A, \tau, h, g)$ is non-increasing w.r.t. h and g . ■

Using the monotonicity property of the value function, as demonstrated by Lemma 2, the following Theorem characterizes some structural properties of the AoI-optimal policy π^* .

Theorem 1. *For any $s_1 = (B_1, A_1, \tau_1, h_1, g_1)$ and $s_2 = (B_2, A_2, \tau_2, h_2, g_2)$, the AoI-optimal policy π^* has the following structural properties:*

- (i) *When $B_1 \geq B_2$, $s_1 \setminus B_1 = s_2 \setminus B_2$ and $B_2 \geq b_{\max} - \lfloor E_1^H/e_q \rfloor$, if $\pi^*(s_1) = (I, H)$, then $\pi^*(s_2) = (I, H)$.*
- (ii) *When $B_1 \geq B_2$, $s_1 \setminus B_1 = s_2 \setminus B_2$ and $B_2 \geq b_{\max} - \lfloor E_1^H/e_q \rfloor + E^S$, if $\pi^*(s_1) = (a_1, H)$, then $\pi^*(s_2) = (a_1, H)$.*
- (iii) *When $A_2 \geq A_1$ and $s_1 \setminus A_1 = s_2 \setminus A_2$, if $\pi^*(s_1) = (a_1, T)$, then $\pi^*(s_2) = (a_1, T)$.*
- (iv) *When $\tau_2 \geq \tau_1$ and $s_1 \setminus \tau_1 = s_2 \setminus \tau_2$, if $\pi^*(s_1) = (S, a_2)$, then $\pi^*(s_2) = (S, a_2)$.*

Proof: We first notice from (9) that when $\pi^*(s_1) = a$, we have $Q(s_1, a) - Q(s_1, a') \leq 0, \forall a' \in \mathbb{A}(s_1)$. Hence, proving that $\pi^*(s_1) = a$ leads to $\pi^*(s_2) = a$ is equivalent to showing

$$Q(s_2, a) - Q(s_2, a') \leq Q(s_1, a) - Q(s_1, a'), \forall a' \neq a. \quad (15)$$

For instance, to prove (i), we need to show that (15) holds when $a = (I, H)$ and $a' \in \{(I, T), (S, H), (S, T)\}$. In the following, we prove part (i) while parts (ii), (iii) and (iv) can be proven similarly. According to (2), the next battery level for both states s_1 and s_2 when taking action $a = (I, H)$ is b_{\max} since we have $B_1 \geq B_2$ and $B_2 \geq b_{\max} - \lfloor E_1^H/e_q \rfloor$. Therefore, we have $Q(s_2, a) = Q(s_1, a)$ since $s_1 \setminus B_1 = s_2 \setminus B_2$, and showing that (15) holds for (i) reduces to showing that $Q(s_1, a') \leq Q(s_2, a'), \forall a' \neq a$. Now, since $B_1 \geq B_2$, we note from (2) that the next battery level of s_1 is greater than or equal to the associated next battery level with s_2 for all possible values of $a' \neq a$. Therefore, based on Lemma 2 ($V(s)$ is non-increasing w.r.t. B), we have $Q(s_1, a') \leq Q(s_2, a'), \forall a' \neq a$ from (6) and (8). This completes the proof of (i). ■

Remark 1. *Theorem 1 demonstrates the threshold-based structure of the AoI-optimal policy π^* w.r.t. each of the system state variables. Specifically from (i) and (ii), we can see that π^* has a threshold-based structure w.r.t. B when taking action (I, H) for $B \geq b_{\max} - \lfloor E_1^H/e_q \rfloor$ (when taking action (a_1, H) for $B \geq b_{\max} - \lfloor E_1^H/e_q \rfloor + E^S$). For instance, for a fixed $s \setminus B$, if B_{th} is the maximum value of $B \geq b_{\max} - \lfloor E_1^H/e_q \rfloor$*

for which it is optimal to take an action $a = (I, H)$, then for all states s such that $B \leq B_{\text{th}}$, the optimal decision is (I, H) as well. Similarly, from (iii) and (iv), we observe that π^* has a threshold-based structure w.r.t. A and τ when taking actions (a_1, T) and (S, a_2) , respectively. This essentially means that π^* aims to restrict the occurrence of the scenario of having a large AoI value at the destination node. In fact, in such a scenario, π^* would allocate a time slot for update packet transmission as soon as the source node has enough energy required for performing that action so that the average AoI at the destination node (expressed in (5)) is minimized.

One can also show that (15) does not hold when $B_1 < B_2$ in parts (i) and (ii), $A_2 < A_1$ in part (iii) or $\tau_2 < \tau_1$ in part (iv). Because of this, it is not possible to discuss structural properties in this case.

Remark 2. *Based on Remark 1, the threshold-based structure of π^* w.r.t. the system state variables can be exploited to reduce the computational complexity of the VIA in terms of the number of required evaluations. More specifically, due to the threshold-based structure of π^* , the optimal actions at some states can be directly determined based on the optimal actions taken at some other states without performing any evaluations. This, in turn, reduces the number of evaluations needed for the policy improvement step, and hence the computational complexity of the VIA is reduced. We refer the readers to [12], [31] for a detailed treatment of this point.*

IV. NUMERICAL RESULTS

We model the uplink and downlink channel power gains between the source and destination as $g = h = \delta\theta^2 d^{-\beta}$; δ is the gain of the signal power at a distance of 1 meter, $d^{-\beta}$ models power law path-loss with exponent β , and $\theta^2 \sim \exp(1)$ denotes the small-scale fading gain. Each state variable is discretized into 10 levels. Considering a similar simulation setup to that of [29], we use $W = 1$ MHz, $d = 25$ meters, $P_t = 37$ dBm, $P_{\max} = 12$ dBm, $\sigma^2 = -95$ dBm, $M = 12$ Mbits, $B_{\max} = 0.3$ mJoules, $a = 1500$, $b = 0.0022$, $\delta = 4 \times 10^{-2}$, and $\beta = 2$. We also consider that the sensitivity of the power received at the RF energy harvesting circuit is -13 dBm. Note that we use the red (blue) color to represent $a_2 = T$ ($a_2 = H$) whereas the circle (square) marker to represent $a_1 = S$ ($a_1 = I$).

First, from Figs. 2a, 2b, 2c and 2d, we can verify the analytical structural properties of π^* derived in Theorem 1. For instance, we can observe from Figs. 2a and 2b that π^* has a threshold-based structure w.r.t. A (τ) when action (a_1, T) (action (S, a_2)) is taken, as derived in parts (iii) and (iv) of Theorem 1. In addition, parts (i) and (ii) of Theorem 1 can be verified from Figure 2c. For instance, since $\lceil E^H/e_q \rceil = 9$ and $E^S = 4$, we can see that: 1) since the optimal action at the point $(3, 4)$ is (I, H) , it is optimal to take action (I, H) at the points $(B, 4), 0 \leq B \leq 3$ (part (i) in Theorem 1), and 2) since the optimal action at the point $(9, 4)$ is (S, H) , it is optimal to take action (S, H) at the points $(B, 4), 4 \leq B \leq 9$ (part (ii) in Theorem 1). Second, the impact of E^S on π^* is revealed in Figs. 2a and 2b, where $\lceil E^T/e_q \rceil = 2$. In particular, we

discuss this impact in two different regimes: 1) the value of E^S is comparable with B ($E^S/B = 3/5$ in Fig. 2a), and 2) E^S is small w.r.t. B ($E^S/B = 3/9$ in Fig. 2b). We observe that when E^S is comparable with B and τ is relatively large, it is optimal to take action (S, H) and save energy that could be used for an update packet transmission for future packet transmissions when τ is small. Note that this insight can also be obtained for small values of A (e.g., $A = 1$ in Fig. 2d).

Third, we show the impact of M on the optimal achievable average AoI (\bar{A}^*) in Fig. 2e. As expected, \bar{A}^* monotonically increases w.r.t. M since the larger M , the larger is E^T required for its transmission. Finally, in Fig. 2f, we demonstrate the importance of our proposed joint sampling and updating policy by comparing its achievable average AoI with that of the generate-at-will policy proposed in [27]. The generate-at-will policy just decides whether to allocate each time slot for an update packet transmission or WET such that the update packets are only generated at the beginning of the time slots allocated for update packet transmissions. This means that the generate-at-will policy does not optimize the timing of update packet generations, and hence Fig. 2f captures the impact of optimally generating update packets on \bar{A}^* . We observe from Fig. 2f that the achievable average AoI by our proposed policy significantly outperforms that of the generate-at-will policy [27] especially when M is large and/or when E^S is large. This happens since it becomes crucial in such cases to wisely decide the timing of update packet generations so that the energy available at the battery can be efficiently utilized to achieve a small value of average AoI.

V. CONCLUSION

This paper has studied the long-term average AoI minimization problem for wireless powered communication systems while taking into account the costs of generating status updates at the source nodes. The problem was modeled as an average cost MDP for which its corresponding value function was shown to be monotonic w.r.t. the state variables. We analytically demonstrated the threshold-based structure of the AoI-optimal policy w.r.t. the state variables. Our results also demonstrated the importance of optimally generating status updates by showing that the performance of our proposed joint sampling and updating policy significantly outperforms that of the generate-at-will policy in terms of the achievable average AoI. A promising avenue of future work is to extend our analysis and results to the scenario with multiple source nodes. Given the prohibitive complexity of the problem resulting from the extreme curse of dimensionality in the state space of its associated MDP, it is difficult to tackle it with conventional approaches. A feasible option is to use deep reinforcement learning-based algorithms to reduce the complexity of the state space while learning the optimal policy at the same time.

REFERENCES

- [1] M. A. Abd-Elmagid, N. Pappas, and H. S. Dhillon, "On the role of age of information in the Internet of things," *IEEE Commun. Magazine*, vol. 57, no. 12, pp. 72–77, 2019.
- [2] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc., IEEE INFOCOM*, 2012.

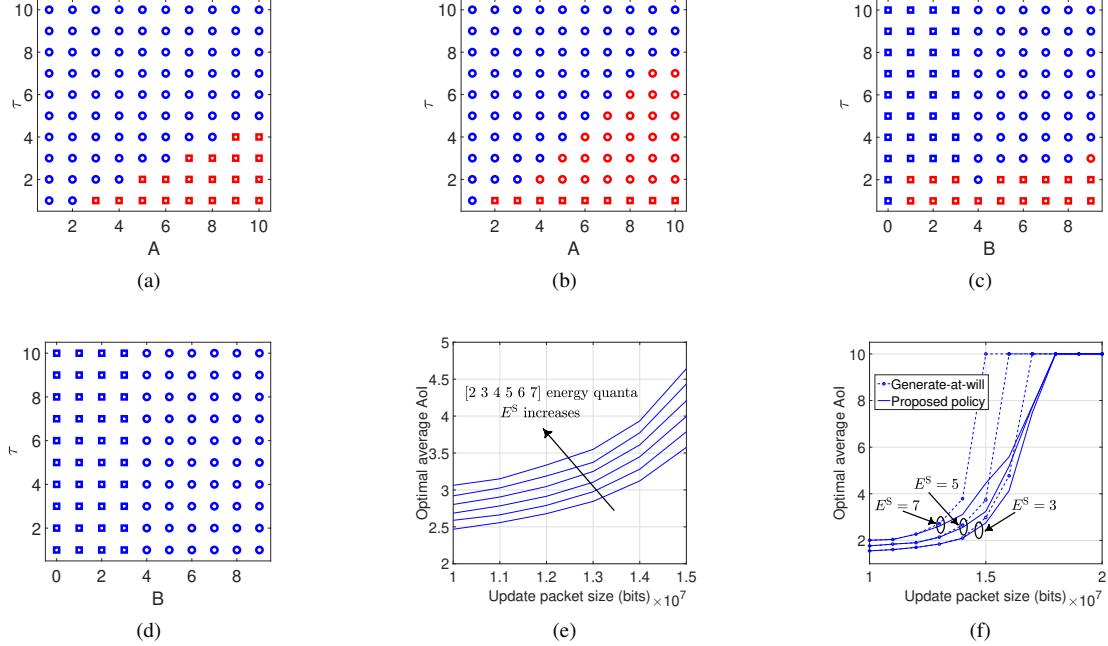


Fig. 2. Structure of the AoI-optimal policy: (a) $E^S = 3$ and $B = g = h = 5$, (b) $E^S = 3$, $g = h = 5$, and $B = 9$, (c) $E^S = 4$, $g = h = 6$, and $A = 5$, and (d) $E^S = 4$, $g = h = 6$, and $A = 1$. System design insights: (e) Impact of E^S on the optimal achievable average AoI, and (f) Comparison between the performance of the proposed joint sampling and updating policy and that of the generate-at-will policy proposed in [27].

- [3] R. Talak, S. Karaman, and E. Modiano, "Minimizing age-of-information in multi-hop wireless networks," in *Proc., Allerton Conf. on Commun., Control, and Computing*, 2017.
- [4] B. Buyukates, A. Soysal, and S. Ulukus, "Age of information in Two-hop multicast networks," in *Proc., IEEE Asilomar*, 2018.
- [5] I. Kadota, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Minimizing the age of information in broadcast wireless networks," in *Proc., Allerton Conf. on Commun., Control, and Computing*, 2016.
- [6] M. Bastopcu and S. Ulukus, "Who should google scholar update more often?" in *Proc., IEEE INFOCOM Workshops*, 2020.
- [7] M. K. Abdel-Aziz, C.-F. Liu, S. Samarakoon, M. Bennis, and W. Saad, "Ultra-reliable low-latency vehicular networks: Taming the age of information tail," in *Proc., IEEE Globecom*, 2018.
- [8] A. Kosta, N. Pappas, and V. Angelakis, "Age of information: A new concept, metric, and tool," *Foundations and Trends in Networking*, 2017.
- [9] Y. Sun, I. Kadota, R. Talak, and E. Modiano, "Age of information: A new metric for information freshness," *Morgan and Claypool Publishers*, 2019.
- [10] Y. Gu, H. Chen, Y. Zhou, Y. Li, and B. Vucetic, "Timely status update in internet of things monitoring systems: An age-energy tradeoff," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 5324–5335, 2019.
- [11] M. A. Abd-Elmagid and H. S. Dhillon, "Average peak age-of-information minimization in UAV-assisted IoT networks," *IEEE Trans. on Veh. Technology*, vol. 68, no. 2, pp. 2003–2008, Feb 2019.
- [12] B. Zhou and W. Saad, "Joint status sampling and updating for minimizing age of information in the Internet of Things," *IEEE Trans. on Commun.*, vol. 67, no. 11, pp. 7468–7482, 2019.
- [13] M. A. Abd-Elmagid, A. Ferdowsi, H. S. Dhillon, and W. Saad, "Deep reinforcement learning for minimizing age-of-information in UAV-assisted networks," in *Proc., IEEE Globecom*, 2019.
- [14] P. D. Mankar, Z. Chen, M. A. Abd-Elmagid, N. Pappas, and H. S. Dhillon, "Throughput and age of information in a cellular-based IoT network," 2020, available online: arxiv.org/abs/2005.09547.
- [15] E. Fountoulakis, N. Pappas, M. Codreanu, and A. Ephremides, "Optimal sampling cost in wireless networks with age of information constraints," in *Proc., IEEE INFOCOM Workshops*, 2020.
- [16] M. A. Abd-Elmagid, M. A. Kishk, and H. S. Dhillon, "Joint energy and SINR coverage in spatially clustered RF-powered IoT network," *IEEE Trans. on Green Commun. and Networking*, vol. 3, no. 1, pp. 132–146, March 2019.
- [17] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *Proc., IEEE ISIT*, 2015.
- [18] B. T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu, "Age of information under energy replenishment constraints," in *Proc., IEEE ITA*, 2015.
- [19] X. Wu, J. Yang, and J. Wu, "Optimal status update for age of information minimization with an energy harvesting source," *IEEE Trans. on Green Commun. and Networking*, vol. 2, no. 1, pp. 193–204, 2018.
- [20] S. Feng and J. Yang, "Age of information minimization for an energy harvesting source with updating erasures: With and without feedback," 2018, available online: arxiv.org/abs/1808.05141.
- [21] A. Arafa and S. Ulukus, "Timely updates in energy harvesting two-hop networks: Offline and online policies," *IEEE Trans. on Wireless Commun.*, vol. 18, no. 8, pp. 4017–4030, Aug. 2019.
- [22] E. T. Ceran, D. Gündüz, and A. György, "Reinforcement learning to minimize age of information with an energy harvesting sensor with harq and sensing cost," in *Proc., IEEE INFOCOM Workshops*, 2019.
- [23] G. Stamatakis, N. Pappas, and A. Tragantitis, "Control of status updates for energy harvesting devices that monitor processes with alarms," in *Proc., IEEE GLOBECOM Workshops*, 2019.
- [24] O. Ozel, "Timely status updating through intermittent sensing and transmission," 2020, available online: arxiv.org/abs/2001.01122.
- [25] Y. Lu, K. Xiong, P. Fan, Z. Zhong, and K. B. Letaief, "Optimal online transmission policy in wireless powered networks with urgency-aware age of information," in *Proc., IEEE IWCMC*, 2019.
- [26] I. Krikidis, "Average age of information in wireless powered sensor networks," *IEEE Wireless Commun. Letters*, 2019.
- [27] M. A. Abd-Elmagid, H. S. Dhillon, and N. Pappas, "A reinforcement learning framework for optimizing age of information in RF-powered communication systems," *IEEE Trans. on Commun.*, vol. 68, no. 8, pp. 4747 – 4760, Aug. 2020.
- [28] A. Biason and M. Zorzi, "Battery-powered devices in wpcns," *IEEE Transactions on Communications*, vol. 65, no. 1, pp. 216–229, 2017.
- [29] E. Boshkovska, D. W. K. Ng, N. Zlatanov, and R. Schober, "Practical non-linear energy harvesting model and resource allocation for SWIPT systems," *IEEE Commun. Letters*, 2015.
- [30] D. P. Bertsekas, "Dynamic programming and optimal control 3rd edition, volume ii," *Belmont, MA: Athena Scientific*, 2011.
- [31] Y.-P. Hsu, E. Modiano, and L. Duan, "Scheduling algorithms for minimizing age of information in wireless broadcast networks with random arrivals," *IEEE Trans. on Mobile Computing*, to appear.