Optimal Status Updating with a Finite-Battery Energy Harvesting Source

Baran Tan Bacinoglu, Yin Sun, Elif Uysal, and Volkan Mutlu

Abstract: **We consider an energy harvesting source equipped with a finite battery, which needs to send timely status updates to a remote destination. The timeliness of status updates is measured by a non-decreasing penalty function of the age of information (AoI). The problem is to find a policy for generating updates that achieves the lowest possible time-average expected age penalty among all online policies. We prove that one optimal solution of this problem is a monotone threshold policy, which satisfies (i) each new update is sent out only when the age is higher than a threshold and (ii) the threshold is a non-increasing function of the instantaneous battery level. Let τ_B denote the optimal threshold corresponding to the full battery level B, and $p(\cdot)$ denote the age-penalty function, then we can show that $p(\tau_B)$ is equal to the optimum objective value, i.e., the minimum achievable time-average expected age penalty. These structural properties are used to develop an algorithm to compute the optimal thresholds. Our numerical analysis indicates that the improvement in average age with added battery capacity is largest at small battery sizes; specifically, more than half the total possible reduction in age is attained when battery storage increases from one transmission's worth of energy to two. This encourages further study of status update policies for sensors with small battery storage.**

Index Terms: **Age of information, age-energy tradeoff, battery capacity, energy harvesting, non-linear age penalty, optimal threshold, threshold policy.**

I. INTRODUCTION

THE *age of information* (AoI), or simply the age, was proposed in [2], [3] as a performance metric that measures the freshness of information in status-update systems. For a flow of information updates sent from a source to a destination, the age is defined as the time elapsed since the newest update available was generated at the source. That is, if $U(t)$ is the largest among the time-stamps of all packets received by time t, the age is defined as:

$$\Delta(t) = t - U(t). \tag{1}$$

AoI is a particularly relevant performance metric for status-update applications that have growing importance in remote monitoring [4], [5], machine-type communication, industrial manufacturing, telerobotics, Internet of Things and social networks.

In many applications, the timeliness of status updates not only determines the quality of service, but also affects other design goals such as the controllability of a dynamical system that relies on the updates of sensing and control signals. AoI quantifies the timeliness of status-updates from the perspective of the receiver rather than throughput or delay based measures that are actually channel-centric. Moreover, AoI is also related to measures such as the time-average mean-square error (MSE) for remote estimation. An example of this is the result in [6] which showed remote estimation of a Wiener process minimizing MSE reduces to an AoI optimization problem when the sampling times at the transmitting side are independent of the process. While AoI optimization based on linear functions of the age $\Delta(t)$ is a relevant performance goal for most scenarios, the performance of some applications may be related to *non-linear* functions of the age. For example, the change in the value of stale data can be less/more significant as its age grows. In such cases, the penalty of data staleness can be modelled as a non-linear function $p(\Delta(t))$ of the age $\Delta(t)$, i.e., the *age-penalty*. This function is chosen to be non-decreasing so that a decrease in *age-penalty* can be only possible when the age is less. Accordingly, the optimization of the age-penalty parallels to average AoI optimization while it might have distinct optimality conditions.

Ideally, AoI is minimized when status updates are frequent and fresh. That is, good AoI performance requires packets with low delay received regularly. A limitation in the minimization of AoI is a constraint on the long-term average update rate which may be due to an average power budget for the channel over which status updates are sent. A stricter constraint is to keep a detailed budget on the number of status updates by allowing update transmission when a replenishable resource becomes available. This is the case of energy harvesting communication systems where each update consumes a certain amount of the harvested energy, if available. In the related literature of AoI optimization for energy harvesting communication systems, energy harvesting process is considered as an arrival process where each energy arrival carries the energy required for an update [7]–[14]. The goal of AoI optimization in such formulations is to find an optimal timing of update instants in order to minimize average AoI while transmission opportunities are subject to the availability of energy. Energy arrivals occur irregularly or randomly, which models an energy harvesting scenario. The main challenge in optimizing time average expected age under random energy arrivals is that in the case of an energy outages (empty battery), the transmitter must idle for an unknown du-

ration of time. If it is the case that such random durations are inevitable, they introduce a tension for the regulation of inter-update durations. Another challenge is due to the finiteness of battery sizes. Theoretically, it is possible to achieve asymptotically optimal average AoI by employing simple schemes assuming infinite [8] or sufficiently large battery [9] sizes. However, when the battery size is comparable to the energy required per update, such simple schemes do not allow performance guarantees. Consequently, it is important to explore optimal policies under such regimes where performance depends heavily on the statistics of energy arrivals and the battery size.

This study is motivated by the aforementioned challenges of optimizing AoI in energy harvesting systems, capturing both the randomness of energy arrivals and finite energy storage capability. In addition capturing both challenges we go further, by optimizing not only average age itself, but a more general age penalty function $p(\Delta(t))$ that is not necessarily linear (see [15]–[20]). Hence, the problem considered in this study is an age-penalty optimization problem where status updates consume discrete units of energy that are randomly generated, i.e., harvested, such that the number of energy units that can be stored at a time is limited by a finite value which is called battery capacity.

Under the assumption of Poisson energy arrivals, we show the structure of solutions for the age-penalty optimization problem. The structure of the optimal solution reflects a basic intuition about the optimal strategy: Updates should be sent when the update is valuable (when the age is high) and the energy is cheap (the battery level is high). We show that the optimal solution is given by a stopping rule according to which an update is sent when its immediate cost is surpassed by the expected future cost. For Poisson energy arrivals, this stopping rule can be found in the set of policies that we refer as *monotone threshold policies*. Monotone threshold policies have the property that each update is sent only when the age is higher than a certain threshold which is a non-increasing function of the instantaneous battery level. One of our key results is that the value of the age-penalty function at the optimal threshold corresponding to the full battery level is exactly equal to the optimal value of the average age-penalty.

A. Contributions

The contributions of this paper can be summarized as follows:
- We formulate the general average age-penalty optimization problem for sending status updates from an energy harvesting source. This generalizes the AoI optimization goal in the prior studies [1], [8], [9], [11], [12] to a non-linear function of age. In addition to the generalization on the objective, the optimization is carried out over a more general policy space defined only using the causality assumption. We prove that solutions to this general optimization problem can be found among threshold-type policies.
- We show that, for optimal threshold-type policies with non-decreasing thresholds, the value of the penalty function at the threshold corresponding to the highest battery level is equal to the minimum value of the average age-penalty. As this optimal threshold is also the minimum of optimal thresholds at different battery levels, this implies that inter-update durations

under such a policy is always above the minimum value of the average age-penalty.
- For the case when the age-penalty function is linear, i.e., average AoI minimization problem, we provide the optimal thresholds for integer battery size up to 4. These results show that the most significant decrease in the minimum average AoI happens when incrementing the battery capacity of unit size (capable of holding one packet transmission's worth of energy) to two units. The minimum achievable average AoI with a battery size of 4 units is only about 10% larger than the ultimate minimum average AoI with infinite battery capacity. That is a promising result for small sensor systems.
- For average AoI minimization problem, we provide an algorithm that can find near optimal policies achieving average AoI values arbitrarily close to the optimal values for any given battery capacity. This algorithm provides a methodical way to derive near optimal policies utilizing analytical results.

B. Paper Organization

The rest of the paper is organized as follows. In Section II, the related work is discussed and summarized. In Section III, the system model and the formulation of the AoI optimization problem are described. In Section IV, the main results on the structural properties of the solution to the AoI optimization problem are shown and an algorithm to derive solutions for arbitrary integer battery sizes is provided. In Section V, the numerical results validating analytical results and also showing optimal solutions for integer battery size up to 4 are presented. In Section VI, the paper is concluded summaring the results and insights obtained over the course of this study.

II. RELATED WORK

Several studies on AoI considered this performance metric under various queueing system models comparing service disciplines and queue management policies (e.g., [21]–[29]). A common observation in these studies was that many queueing/service policies that are throughput and delay optimal but are often suboptimal with respect to AoI, while AoI-optimal policies can be throughput and delay optimal, at the same time. This showed that AoI optimization is quite different than optimization with respect to classical performance metrics. This required many queueing models to be re-addressed under respect to age related objectives. Moreover, queueing system formulations typically assume no precise control on the transmission or generation times of status updates. However, such control is important for age optimization [16], [17].

A direct control on the generation times of status updates is possible through a control algorithm that runs at the source. This is the *"generate-at-will"* assumption formulated in [7], [10] and studied in [6], [16], [17]. In [7], the problem of AoI optimization for a source, which is constrained by an arbitrary sequence of energy arrivals was studied. In [10], AoI optimization was considered for a source that harvests energy at a constant rate under stochastic delays experienced by the status update packets. The results in these studies showed suboptimality of work-conserving transmission schemes. Often, introducing a waiting time before sending the next update is optimal. That is, for maxi-
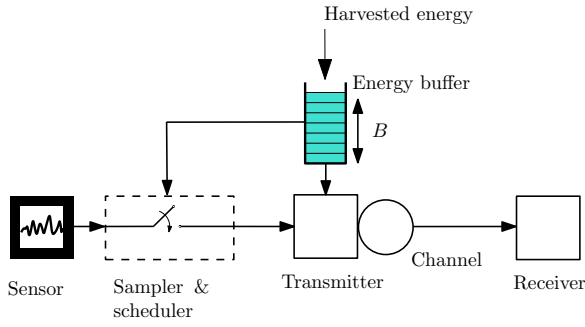
Fig. 1. System model.

mum freshness, one may sometimes send updates at a rate lower than one is allowed to which may be counter-intuitive at first sight.

The problem in [7] was extended to a continuous-time formulation with Poisson energy arrivals, finite energy storage (battery) capacity, and random packet errors in the channel in [8]. An age-optimal threshold policy was proposed for the unit battery case, and the achievable AoI for arbitrary battery size was bounded for a channel with a constant packet erasure probability. The concurrent study in [9], limited to the special cases of unit battery capacity and infinite battery capacity computed the same threshold-type policies under these assumptions. These special cases were investigated also for noisy channels with a constant packet erasure probability in [13], [14]. The case for a battery with 2-units capacity was studied in [11] and the optimal policies for this case characterized as threshold-type policies similar to the optimal policy for unit battery capacity introduced in [8] and [9]. Optimal policies for arbitrary battery sizes were characterized via Lagrangian approach in [12] and using optimal stopping theory in [1].

III. SYSTEM MODEL

Consider an energy harvesting transmitter that sends update packets to a receiver, as illustrated in Fig. 1. Suppose that the transmitter has a finite battery which is capable of storing up to B units of energy. Similar to [8], we assume that the transmission of an update packet consumes one unit of energy. The energy that can be harvested arrive in units according to a Poisson process with rate μ_H. Let $E(t)$ denote the amount of energy stored in the battery at time t such that $0 \leq E(t) \leq B$. The timing of status updates are controlled by a sampler which can monitor the battery level $E(t)$ for all t. We assume that the initial age and the initial battery level are zero, i.e., $\Delta(0) = 0$ and $E(0) = 0$.

Let $H(t)$ and $A(t)$ denote the number of energy units that have arrived during $[0, t]$ and the number of updates sent out during $[0, t]$, respectively. Hence, $\{H(t), t \geq 0\}$ and $\{A(t), t \geq 0\}$ are two counting processes. If an energy unit arrives when the battery is full, it is lost because there is no capacity to store it.

The system starts to operate at time $t = 0$. Let Z_k denote the generation time of the k-th update packet such that $0 = Z_0 \leq Z_1 \leq Z_2 \leq \cdots$. An update policy is represented by a sequence of update instants $\pi = (Z_0, Z_1, Z_2, \cdots)$. Let X_k

represent the inter-update duration between updates $k - 1$ and k, i.e., $X_k = Z_k - Z_{k-1}$. In many status-update systems (e.g., a sensor reporting temperature [30]), update packets are small in size and are only sent out sporadically. Typically, the duration for transmitting a packet is much smaller than the difference between two subsequent update times, i.e., X_ks are typically large compared to the duration of a packet transmission. With such systems in mind, in our model, we will approximate the packet transmission durations as zero. In other words, once the k-th update is generated and sent out at time $t = Z_k$, it is immediately delivered to the receiver. Hence, the age of information $\Delta(t)$ at any time $t \geq 0$ is

$$\Delta(t) = t - \max\{Z_k : Z_k \leq t\}, \tag{2}$$

which satisfies $\Delta(t) = 0$ at each update time $t = Z_k$. Because an update costs one unit of energy, the battery level reduces by one upon each update, i.e.,

$$E(Z_k) = E(Z_k^-) - 1, \tag{3}$$

where Z_k^- is the time immediately before the kth update. Further, because the battery size is B, the battery level evolves according to

$$E(t) = \min\{E(Z_k) + H(t) - H(Z_k), B\}, \tag{4}$$

when $t \in [Z_k, Z_{k+1})$ is between two subsequent updates.

In terms of energy available to the scheduler, we can define update policies, that do not violate causality, as in the following:

Definition 1: A policy π is said to be *energy-causal* if updates only occur when the battery is non-empty, that is, $E(Z_k^-) \geq 1$ for each packet k.

Another restriction on update instants is due to the information available to the scheduler which we define as follows,

Definition 2: Information on the energy arrivals and updates by time t is represented by the filtration [1] $\mathcal{F}_t = \sigma(\{(H(t'), A(t')), 0 \leq t' < t\})$ which is the σ-field generated by the sequence of energy arrivals and updates, i.e., $\{(H(t'), A(t')), 0 \leq t' < t\}$.

Similar to the definition of energy-causal policies, in the policy space that we will consider we merely assume the causality of available information besides energy causality. To formulate this assumption, we use the definition of $\mathcal{F}_t$. In terms of information available to the scheduler, any random time instant θ does not violate causality if and only if $\{\theta \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$. We will refer such random instants as *Markov times* [31] and consider update times as Markov times based on the filtration $\mathcal{F}_t$ in general. Notice that such update times do not have to be finite, however, we will refer Markov times that are also finite with probability 1 (w.p.1.) as *stopping times* [31]. For a policy trying to regulate age, it is legitimate to assume that update instants are always finite w.p.1. as otherwise the age may grow unbounded with a positive probability. With this in mind, we will consider only the update instants that are stopping times.

Accordingly, we can define the *online* update policies combining the causality assumptions on available energy and infor-

---

[1]Note that the filtration is right continuous as both $H(t)$ and $A(t)$ are right continuous.

mation as follows:

Definition 3: A policy is said to be *online* if (i) it is energy causal, (ii) no update instant is determined based on future information, i.e., all update times are stopping (finite Markov) times based on $\mathcal{F}_t$, i.e., Z_k is finite w.p.1. while $\{Z_k \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$ and $k \geq 1$.

Let Π^{online} denote the set of online update policies. To evaluate the performance of online policies, we consider an *age-penalty function* that relates the age at a particular time to a cost which increases by the age. This function is defined as in below:

We consider an *age-penalty function* $p(\cdot)$ that maps the age $\Delta(t)$ at time t to a penalty $p(\Delta(t))$:

Definition 4: A function $p : [0, \infty) \to [0, \infty)$ of the age is said to be an *age-penalty function* if

- $\lim_{\Delta \to \infty} p(\Delta) = \infty$.
- $p(\cdot)$ is a non-decreasing function.
- $\int_0^\infty p(t)e^{-\alpha t}dt < \infty$ for all $\alpha > 0$.

Observe that the definition of age-penalty functions covers any non-decreasing function of age that is of sub-exponential order[2] and grows to infinity.

The time-average expected value of the age-penalty or simply the *average age-penalty* can be expressed as

$$\bar{p} = \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}\left[\int_0^T p(\Delta(t))dt\right]. \tag{5}$$

Let $\bar{p}_\pi$ denote the average age-penalty achieved by a particular policy π. The goal of this paper is to find the optimal update policy for minimizing the average age-penalty, which is formulated as

$$\min_{\pi \in \Pi^{\text{online}}} \bar{p}_\pi. \tag{6}$$

IV. MAIN RESULTS

We begin with a result guaranteeing that the space of threshold-type policies (see Definition 5) contains optimal update policies hence we can focus our attention to these policies for finding solutions to (6).

Note that at time $t = Z_k$, the age $\Delta(t)$ is equal to 0. In the meanwhile, the battery level $E(t)$ will grow as more energy is harvested. In threshold policies, the threshold $\tau_{E(t)}$ changes according to the battery level $E(t)$ and a new sample is taken at the earliest time that the age $\Delta(t)$ exceeds the threshold $\tau_{E(t)}$. We define such policies as follows:

Definition 5: When $E(t) \in \{\ell = 1, \cdots, B\}$ represents the battery level at time t, an online policy is said to be a *threshold policy* if there exists τ_ℓ for $\ell = 1, \cdots, B$ s.t.

$$Z_{k+1} = \inf\left\{t \geq Z_k : \Delta(t) \geq \tau_{E(t)}\right\}. \tag{7}$$

Note that a policy is said to be *stationary* if its actions depend only on a current state while being independent of time. An immediate observation is that given $\Delta(t)$ and $E(t)$ threshold policies do not depend on time, hence:

Proposition 1: All threshold policies are stationary.

---

[2]This is due to the third property in the definition, which is a technical requirement for the proofs.

Proof: By definition, the update instants of a threshold policy only depend on the time elapsed since the last update, i.e., $\Delta(t)$, and the current battery level. □

We expect that such stationary policies can minimize $\bar{\Delta}$ among all online policies as energy arrivals follow a Poisson process which is memoryless. Due to the memorylessness of energy arrivals, the evolution of the system can be understood through a renewal type behaviour which suggests that an optimal policy should be stationary.

Indeed, we note the following as the first key result of this paper.

Theorem 1: There exists a threshold policy that is optimal for solving (6).

Proof: See Appendix A. □

One significant challenge in the proof of Theorem 1 is that (6) is an infinite time-horizon time-averaged MDP which has an uncountable state space. When the state space is countable, one can analyze infinite time-horizon time-averaged MDP by making a unichain assumption. However, this method cannot be directly applied when state space is uncountable. To resolve this, we use a modified version of the "vanishing discount factor" approach [32] to prove Theorem 1 in two steps:

1. Show that for every $\alpha > 0$, there exists a threshold policy that is optimal for solving

$$\min_{\pi \in \Pi^{\text{online}}} \mathbb{E}\left[\int_0^\infty e^{-\alpha(t-a)}p(\Delta(t))dt\right].$$

2. Prove that this property also holds when the discount factor α vanishes to zero.

In our search for an optimal policy, we can further reduce the space of policies:

Definition 6: A threshold policy is said to be a *monotone threshold* policy if $\tau_1 \geq \tau_2 \geq \cdots \geq \tau_B$.

Note that the definition of monotone threshold policies refers only to the case of thresholds that non-increasing in battery levels as opposed to the non-decreasing case.

Let Π^{MT} be the set of monotone threshold policies, then, the following is true:

Theorem 2: There exists a monotone threshold policy $\pi \in \Pi^{\text{MT}}$ that is optimal for solving (6).

Proof: See Appendix B. □

Theorem 2 implies that in the optimal update policy, update packets are sent out more frequently when the battery level is high and less frequently when the battery level is low. This result is quite intuitive: If the battery is full, arrival energy cannot be harvested; if the battery is empty, update packets cannot be transmitted when needed and the age increases. Hence, both battery overflow and outage are harmful. Monotone threshold policies can address this issue. When the battery level l is high, the threshold τ_l is small to reduce the chance of battery overflow; when the battery level l is low, the threshold τ_l is high to avoid battery outage.

For a policy in Π^{MT}, the state $(\Delta(t), E(t))$ does not spend a measurable amount of time anywhere $\Delta(t) \geq \tau_{E(t)}$ in which an update is sent out instantly reducing the battery level. Otherwise, the battery level is incremented upon energy harvests while the age is increasing linearly in time. The illustration in Fig. 2 shows the time evolution of the state $(\Delta(t), E(t))$ for
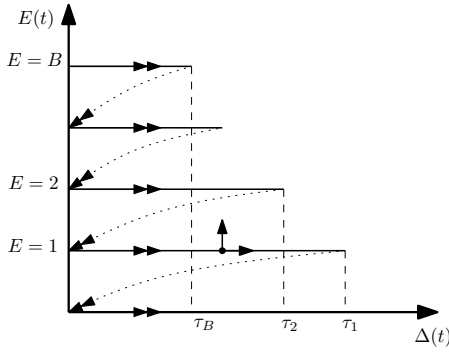
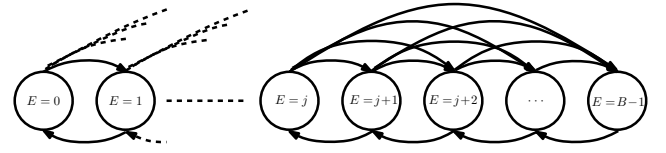Fig. 2. An illustration of a monotone threshold policy.

Fig. 3. The DTMC for energy states sampled at update times.

threshold policy that satisfies the following

$$p(\tau_B^*) = \bar{p}_{\pi^*} = \min_{\pi \in \Pi^{\text{online}}} \bar{p}_\pi, \qquad (11)$$

where π^* is a monotone threshold policy solving (6) and τ_B^* is its age threshold for the full battery case.

Proof: See Appendix E.   □

The result in Theorem 3 exhibits a structural property of optimal policies which also appears in the sampling problem that was studied in [19]. The sampling problem in [19] considered sources without energy harvesting, where the packet transmission times were *i.i.d.* and non-zero. On the one hand, the optimal sampling policy in Theorem 1 of [19] is a threshold policy on an expected age penalty term, and the threshold is exactly equal to the optimal objective value. On the other hand, a sampling problem for an energy harvesting source with zero packet transmission time is considered in the current paper. The optimal sampling policy in Theorem 3 can be rewritten as

$$Z_{k+1} = \inf\left\{ t \geq Z_k : p(\Delta(t)) \geq p(\tau_{E(t)}^*) \right\},$$

which is a multi-threshold policy on the age penalty function, each threshold $p(\tau_\ell^*)$ corresponding to a battery level ℓ. Further, the threshold $p(\tau_B^*)$ associated with a full battery size $E(t) = B$ is equal to the optimal objective value. The results in these two studies are similar to each other. Together, they provide a unified view on optimal sampler design for sources both with and without energy harvesting capability. The proof techniques in these two studies are of fundamental difference.

A. Average Age Case

If we take the age-penalty function as an identity function, i.e., $p(\Delta) = \Delta$, then (6) becomes the problem of minimizing the time-average expected age. In this case, the result in Theorem 3 implies that in optimal monotone threshold policies, inter-update durations can be small as much as the minimum average AoI only when the battery is full. From results in [8] and [9], we know that the minimum average AoI for the infinite battery case is $1/2\mu_H$ and this can be achieved asymptotically using the best-effort scheme in [9] or with a threshold policy [8] where all thresholds are nearly equal to $1/\mu_H$. On the other hand, according to Theorem 3, the optimal threshold for the full battery level tends to $1/2\mu_H$ as the battery capacity increases. This shows that the optimal monotone threshold policies remain structurally dissimilar to asymptotically optimal policies when the battery capacity is approaching to infinity. The result is more useful when the battery capacity is finite as it may lead to the optimal threshold values of the other battery levels. We will use this in an algorithm for finding near optimal policies for any given integer sized battery capacity. In addition, the special case of Theo-

policies in Π^{MT}. If the energy level is $E(Z_k) = j$ upon the previous update, then the inter-update time $X_{k+1} \in [\tau_m, \tau_{m-1}]$ holds if and only if $m - j$ packets arrive during the inter-update time. In other words, reaching the battery state m or higher is necessary and sufficient for the next inter-update duration being shorter than some x when $x \in [\tau_m, \tau_{m-1})$. Let Y_i denote the duration required for $i \geq 1$ successive energy arrivals, which obeys the Erlang distribution at rate μ_H with parameter i,

$$P(Y_i \leq x) = 1 - \sum_{v=0}^{i-1} \frac{1}{v!} e^{-\mu_H x} (\mu_H x)^v, \qquad (8)$$

and let $Y_i = 0$ for $i \leq 0$.

Accordingly, for policies in Π^{MT}, the cumulative distribution function (CDF) of inter-update durations, can be expressed as

$$\Pr(X_{k+1} \leq x \mid E(Z_k) = j) =$$
$$\begin{cases} 0, \text{if } x < \tau_B, \\ \Pr(Y_{m-j} \leq x), \text{if } \tau_m \leq x < \tau_{m-1}, \forall m \in \{2, \cdots, B\}, \\ \Pr(Y_{1-j} \leq x), \text{if } \tau_1 \leq x. \end{cases}$$
$$(9)$$

From (9), an expression for the transition probability $\Pr(E(Z_{k+1}) = i \mid E(Z_k) = j)$ for $i = 0, 1, \cdots, B - 1$ can be derived[3]

$$\Pr(E(Z_{k+1}) = i \mid E(Z_k) = j) =$$
$$\begin{cases} \Pr(Y_{B-j} \leq \tau_{B-1}), \text{if } i = B - 1, \\ \Pr(Y_{1+i-j} \leq \tau_i) - \Pr(Y_{2+i-j} \leq \tau_{i+1}), \text{if } i < B - 1. \end{cases}$$
$$(10)$$

Hence, energy states sampled at update instants can be described as a discrete time Markov chain (DTMC) with the transition probabilities in (10) (See Fig. 3). When thresholds are finite, this DTMC is ergodic as any energy state is reachable from any other energy state in $B - 1$ steps with positive probability.

Any optimal policy in Π^{MT} has the following property:

Theorem 3: An optimal policy for solving (6) is a monotone

---

[3]Note that the event $E(Z_{k+1}) = i$ happens if and only if $X_{k+1} \in [\tau_{i+1}, \tau_i)$, accordingly $\Pr(E(Z_{k+1}) = i \mid E(Z_k) = j) = \Pr(X_{k+1} \leq \tau_i \mid E(Z_k) = j) - \Pr(X_{k+1} \leq \tau_{i+1} \mid E(Z_k) = j)$.

rem 3 for average age [1] can be derived from a more general result which we provide in Lemma 1. This result shows a relation between the partial derivatives of a non-negative random variable with respect to the thresholds determining the random variable in a similar way to the inter-update duration case.

Lemma 1: Suppose X is a r.v. that satisfies the following:

$$\Pr(X \leq x) = \begin{cases} 0 & \text{if } x < \tau_B, \\ F_i(x) & \text{if } \tau_i \leq x < \tau_{i-1}, \forall i \in \{2, \cdots, B\}, \\ F_1(x) & \text{if } \tau_1 \leq x, \end{cases}$$

where $0 < \tau_B \leq \cdots \leq \tau_2 \leq \tau_1$ and for each $i \in \{1, \cdots, B\}$ $F_i(x)$ is the CDF of a non-negative random variable. Then:

$$\frac{\partial}{\partial \tau_i} \mathbb{E}\left[X^2\right] = 2\tau_i \frac{\partial}{\partial \tau_i} \mathbb{E}\left[X\right].$$

Proof: See Appendix C. □

Corollary 1: The inter-update intervals, X, for any $\pi \in \Pi^{\mathrm{MT}}$ satisfy the following:

$$\frac{\partial}{\partial \tau_i} \mathbb{E}\left[X^2 \mid E = j\right] = 2\tau_i \frac{\partial}{\partial \tau_i} \mathbb{E}\left[X \mid E = j\right], \forall (i, j) \in \{1, 2, \cdots, B\}^2, \tag{12}$$

where $\mathbb{E}\left[X \mid E = j\right] \triangleq \mathbb{E}\left[X_k \mid E(Z_k) = j\right]$ and $\mathbb{E}\left[X^2 \mid E = j\right] \triangleq \mathbb{E}\left[X_k^2 \mid E(Z_k) = j\right]$.

Note that the transition probabilities (10) do not depend on τ_B hence the steady-state probabilities obtained from (10) also do not depend on τ_B. This leads to a property of τ_B the average age case of Theorem 3 as shown in [1]. The unit-battery case, i.e., $B = 1$ case was solved in [8] and [9]. For completeness, this result is summarized in Theorem 4.

Theorem 4: When $B = 1$, the average age $\bar{\Delta}$ can be expressed as

$$\bar{\Delta} = \frac{\frac{1}{2}(\mu_H \tau_1)^2 + e^{-\mu_H \tau_1}(\mu_H \tau_1 + 1)}{\mu_H(\mu_H \tau_1 + e^{-\mu_H \tau_1})}, \tag{13}$$

and $\tau_1^* = \bar{\Delta}_{\pi^*} = \frac{1}{\mu_H} 2W\left(\frac{1}{\sqrt{2}}\right)$ where $W(\cdot)$ is the Lambert-W function.

Proof: See Appendix F. □

Theorem 5: When $B = 2$, the average age $\bar{\Delta}$ can be expressed as:

$$\bar{\Delta} = \frac{\frac{\alpha_2^2}{2} + e^{-\alpha_2}[\alpha_2 + 1 + \rho_1(\alpha_2^2 + 2\alpha_2 + 2)] - e^{-\alpha_1}[\alpha_1 + 1 + \rho_1(\alpha_1^2 + \alpha_1 + 1)]}{\mu_H(\alpha_2 + e^{-\alpha_2}[1 + \rho_1(\alpha_2 + 1)] - e^{-\alpha_1}[1 + \rho_1 \alpha_1])}, \tag{14}$$

where

$$\rho_1 = \frac{e^{-\alpha_1}}{1 - e^{-\alpha_1} \alpha_1},$$

and

$$\alpha_1 = \mu_H \tau_1, \alpha_2 = \mu_H \tau_2.$$

Proof: See Appendix G. □

B. An Algorithm for Finding Near Optimal Policies

We propose an algorithm to find a near optimal policy $\pi \in \Pi^{\mathrm{MT}}$ such that $\bar{\Delta}_\pi - \bar{\Delta}_{\pi^*} \leq \frac{1}{2^{q+1}\mu_H}$ for any given B and $q \in$

Algorithm 1 Find $\pi \in \Pi^{\mathrm{MT}}$ such that $\bar{\Delta}_\pi - \bar{\Delta}_{\pi^*} \leq \frac{1}{2^{q+1}\mu_H}$.

Require: $B \geq 1 \wedge q \geq 1$
Ensure: $\bar{\Delta}_\pi - \bar{\Delta}_{\pi^*} \leq \frac{1}{2^{q+1}\mu_H}$
  $\tau_B^- \leftarrow \frac{1}{2\mu_H}, \tau_B^+ \leftarrow \frac{1}{\mu_H}$
  for $i = 1, 2, \cdots, q$ **do**
    $\hat{\tau}_B \leftarrow \frac{\tau_B^- + \tau_B^+}{2}$
    if $\exists \tau_{B-1} \leq \cdots \leq \tau_2 \leq \tau_1$ s.t. $\tau_{B-1} \geq \hat{\tau}_B$ and
    $2\hat{\tau}_B m_1(\tau_1, \tau_2, \cdots, \hat{\tau}_B) - m_2(\tau_1, \tau_2, \cdots, \hat{\tau}_B) = 0$ **then**
      $\tau_B^+ \leftarrow \hat{\tau}_B$
    else
      $\tau_B^- \leftarrow \hat{\tau}_B$
    end if
  end for
  Solve $2\hat{\tau}_B m_1(\tau_1, \tau_2, \cdots, \hat{\tau}_B) - m_2(\tau_1, \tau_2, \cdots, \hat{\tau}_B) = 0$
  return $\pi = (\tau_1, \tau_2, \cdots, \hat{\tau}_B)$

$\mathbb{Z}^+$. Let $m_1(\tau_1, \tau_2, \cdots, \tau_B)$ and $m_2(\tau_1, \tau_2, \cdots, \tau_B)$ denote the functions such that:

$$m_1(\tau_1, \tau_2, \cdots, \tau_B) = \sum_{j=0}^{B-1} \mathbb{E}\left[X \mid E = j\right] \Pr(E = j), \tag{15}$$

$$m_2(\tau_1, \tau_2, \cdots, \tau_B) = \sum_{j=0}^{B-1} \mathbb{E}\left[X^2 \mid E = j\right] \Pr(E = j), \tag{16}$$

where $\Pr(E = j)$ is the steady-state probability for energy state j, $\mathbb{E}\left[X \mid E = j\right] \triangleq \mathbb{E}\left[X_k \mid E(Z_k) = j\right]$ and $\mathbb{E}\left[X^2 \mid E = j\right] \triangleq \mathbb{E}\left[X_k^2 \mid E(Z_k) = j\right]$.

Note that it is straight forward to derive $m_1(\tau_1, \tau_2, \cdots, \tau_B)$ and $m_2(\tau_1, \tau_2, \cdots, \tau_B)$ using (9) and (10), hence we assume these functions are available for any B.

In the below theorem, we state the main result that we will use in an algorithm for finding near optimal policies:

Theorem 6: For $B > 1$, the equation

$$2\tau_B m_1(\tau_1, \tau_2, \cdots, \tau_B) - m_2(\tau_1, \tau_2, \cdots, \tau_B) = 0, \tag{17}$$

has a solution with monotone non-increasing thresholds, i.e., $\tau_B \leq \cdots \leq \tau_2 \leq \tau_1$ if and only if $\tau_B \geq \bar{\Delta}_{\pi^*}$.

Proof: See Appendix H. □

Algorithm 1 uses this result to find a near optimal policy $\pi \in \Pi^{\mathrm{MT}}$ such that $\bar{\Delta}_\pi - \bar{\Delta}_{\pi^*} \leq 1/(2^{q+1}\mu_H)$. Each iteration in Algorithm 1 halves the interval where the minimum average AoI can be found based on the existence of solution to (6) with the current estimate of the smallest threshold $\hat{\tau}_B$. Accordingly, it is guaranteed that Algorithm 1 finds a solution within a gap to the optimal value that is $1/(2^{q+1}\mu_H)$.

Algorithm 1 assumes a numerical solver that can solve the transcendental equation in (17), however, the exact solution is required only once at the final step while iterations only require verifying the existence of a solution to (6).

V. NUMERICAL RESULTS

For battery sizes $B = 1, 2, 3, 4$, the policies in Π^{MT} are numerically optimized giving AoI versus energy arrival rate (Pois-
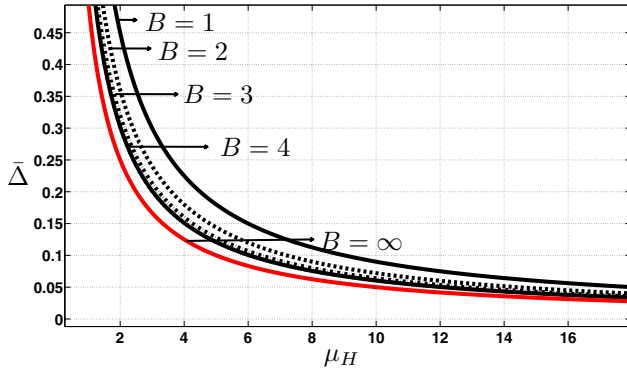
Fig. 4. AoI versus energy arrival rate (Poisson) for different battery sizes $B = 1, 2, 3, 4$.
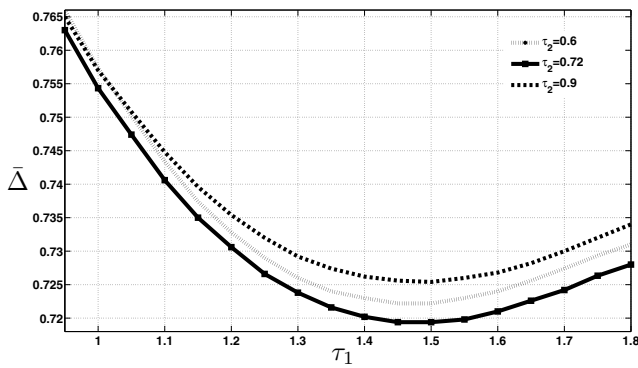

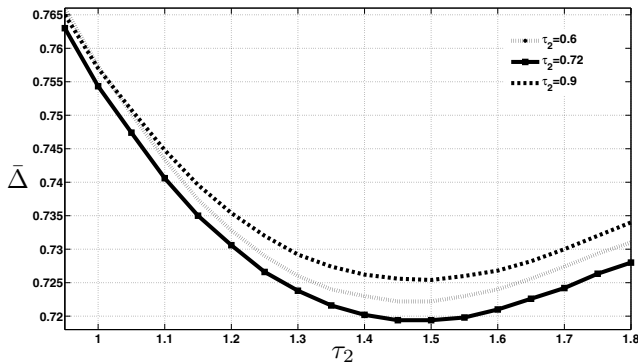Fig. 5. AoI versus τ_1 against various τ_2 values for $B = 2$ and $\mu_H = 1$.

Fig. 6. AoI versus τ_2 against various τ_1 values for $B = 2$ and $\mu_H = 1$.

son) curves in Fig. 4. We give the corresponding threshold values in Table 1. These results were obtained through exhaustive search for possible threshold values, and Monte Carlo analysis for approximating AoI values in the simulation of the considered system and policies without relying on analytical results. It can be seen that these optimal thresholds and corresponding AoI values (in Table 1) validate Theorem 3. Figs. 5 and 6 show the dependency of AoI on threshold values τ_1 and τ_2 which is consistent with the result in Theorem 5 for the special case of $B = 2$.

Table 1. Optimal thresholds for different battery sizes for $\mu_H = 1$.

	τ_1	τ_2	τ_3	τ_4	$\bar{\Delta}_{\pi^*}$
$B = 1$	0.90	-	-	-	0.90
$B = 2$	1.5	0.72	-	-	0.72
$B = 3$	1.5	1.2	0.64	-	0.64
$B = 4$	1.5	1.2	0.86	0.604	0.604

VI. CONCLUSION

We have studied optimizing a non-linear age penalty in the generation and transmission of status updates by an energy harvesting source with a finite battery. An optimal status updating policy for minimizing the time-average expectation of a general non-decreasing age function $p(\cdot)$ has been obtained. The policy has a monotonic threshold structure: (i) Each new update is sent out only when the age is higher than a threshold and (ii) the threshold is a non-increasing function of the instantaneous battery level such that the updates are sent out more frequently when the battery level is high. Furthermore, we have identified an interesting relationship between the smallest optimal threshold τ_B^* (i.e., the threshold corresponding to a full battery level) and the optimal objective value $\bar{p}_{\pi^*}$ (i.e., the minimum achievable time-average expected age penalty), which is given by

$$p(\tau_B^*) = \bar{p}_{\pi^*}. \tag{18}$$

APPENDIX A
THE PROOF OF THEOREM 1

In order to prove Theorem 1, we use a modified version of the "vanishing discount factor" approach [32] which consists of 2 steps:

Step 1. Show that for every $\alpha > 0$, there exists a threshold policy that is optimal for solving

$$\min_{\pi \in \Pi^{\text{online}}} \mathbb{E}\left[\int_0^\infty e^{-\alpha t} p(\Delta(t)) dt\right].$$

Step 2. Prove that this property still holds when the discount factor α vanishes to zero.

We first discuss *Step 1*. Recall that $\mathcal{F}_t$ represents the information about the energy arrivals and the update policy during $[0, t]$. Given $\mathcal{F}_a$, we are interested in finding the optimal online policy during $[a, \infty)$, which is formulated as

$$\min_{\pi \in \Pi^{\text{online}}} \mathbb{E}\left[\int_a^\infty e^{-\alpha(t-a)} p(\Delta(t)) dt \,\bigg|\, \mathcal{F}_a\right]. \tag{19}$$

Observe that, in (19), the term $e^{-\alpha(t-a)}$ ensures that the exponential decay always starts from unity so that the problem is independent of a given $\mathcal{F}_a$. In addition, this problem has the following nice property:

Lemma 2: There exists an optimal solution to (19) that depends on $\mathcal{F}_a$ only through $(\Delta(a), E(a))$. That is, $(\Delta(a), E(a))$ is a sufficient statistic for solving (19).

Proof: In Problem (19), the age evolution $\{\Delta(t), t \geq a\}$ is determined by the initial age $\Delta(a)$ at time a and the update policy during $[a, \infty)$. Further, the update policy during $[a, \infty)$ is determined by the initial age $\Delta(a)$, the initial battery level

$E(a)$, and the energy counting process $\{H(t) - H(a), t \geq a\}$. Hence, $\{\Delta(t), t \geq a\}$ is determined by $\Delta(a)$, $E(a)$, and $\{H(t) - H(a), t \geq a\}$.

Recall that $\Delta(0)$ and $E(0)$ are fixed. Hence, for any online update policy, the online update decisions during $[0, a]$ depends only on $\{H(t), t \leq a\}$. Hence, $\mathcal{F}_a$ is determined by $\{H(t), t \leq a\}$. Because $\{H(t), t \geq 0\}$ is a compound Poisson process, $\{H(t) - H(a), t \geq a\}$ is independent of $\{H(t), t \leq a\}$. Hence, $\{\Delta(t), t \geq a\}$ depends on $\mathcal{F}_a$ only through $\Delta(a)$ and $E(a)$. By this, $(\Delta(a), E(a))$ is a sufficient statistic for solving (19). □

By using Lemma 2, we can simplify (19) as (20) and define a cost function $J_\alpha(\Delta(a), E(a))$ which is the optimal objective value of (20):

$$J_\alpha(\Delta(a), E(a)) := \min_{\pi \in \Pi^{\text{online}}} \mathbb{E}\left[\int_a^\infty e^{-\alpha(t-a)} p(\Delta(t)) dt \middle| \mathcal{F}_a\right]$$

$$= \min_{\pi \in \Pi^{\text{online}}} \mathbb{E}\left[\int_a^\infty e^{-\alpha(t-a)} p(\Delta(t)) dt \middle| \Delta(a), E(a)\right]. \quad (20)$$

Furthermore, one important question is: Given that the previous update occurs at $Z_k = a$, how to choose the next update time Z_{k+1}. This can be formulated as

$$\min_{(Z_1, \cdots, Z_k = a, Z_{k+1}, \cdots) \in \Pi^{\text{online}}}$$
$$\mathbb{E}\left[\int_a^\infty e^{-\alpha(t-a)} p(\Delta(t)) dt \middle| Z_k = a, \Delta(a) = 0, E(a)\right], \quad (21)$$

where we have used the fact that if $Z_k = a$, then $\Delta(a) = \Delta(Z_k) = 0$.

According to the definition of Π^{online}, Z_{k+1} is a finite Markov time, i.e., stopping time, hence the problem of finding Z_{k+1} for a solution to (21) can be formulated as an infinite horizon optimal stopping problem in the interval $[a, \infty)$. We will consider a *gain* [31] process $G = (G_t)_{t \geq a}$ adapted to the filtration $\mathcal{F}_t$ where a stopping time Z_{k+1} for a solution to (21) maximizes $\mathbb{E}\left[G_{Z_{k+1}} \mid \mathcal{F}_a\right]$ when we choose Z_{k+1} from a family of stopping times based on $\mathcal{F}_t$. Let $\mathfrak{M}_a$ denote this family of Z_{k+1}s which can be expressed as:

$$\mathfrak{M}_a =$$
$$\{Z_{k+1} \geq a : \Pr(Z_{k+1} < \infty) = 1, \{Z_{k+1} \leq t\} \in \mathcal{F}_t, \forall t \geq a\}.$$

Note that a stopping time in $\mathfrak{M}_a$ may violate energy causality however our definition of the gain process will guarantee that those stopping times cannot be optimal.

We will define the gain process $(G_t)_{t \geq a}$ based on the value of the discounted cost when an update is sent at a particular time t. The gain process $(G_t)_{t \geq a}$ for $E(t) > 0$ corresponds to the additive inverse of this cost and can be written as follows:

$$G_t = - \min_{\pi \in \Pi^{\text{online}}}$$
$$\mathbb{E}\left[\int_a^\infty e^{-\alpha(w-a)} p(\Delta(w)) dw \middle| Z_k = a, Z_{k+1} = t, E(t)\right],$$
$$E(t) > 0. \quad (22)$$

Note that the stopping time cannot be at time t when $E(t) = 0$ as there is no energy to send another update in that case. To cover this case, we set G_t to $-\infty$ so that a stopping time Z_{k+1} maximizing $\mathbb{E}\left[G_{Z_{k+1}} \mid \mathcal{F}_a\right]$ should satisfy energy causality hence belongs to an online policy. In other words, the stopping time Z_{k+1} in a solution to (21) maximizes $\mathbb{E}\left[G_{Z_{k+1}} \mid \mathcal{F}_a\right]$ among all the stopping times in $\mathfrak{M}_a$.

Alternatively, the gain process $(G_t)_{t \geq a}$ can be expressed in terms of the cost defined in (20) as follows

$$G_t = - \int_a^t e^{-\alpha(w-a)} p(w - a) dw$$
$$- \mathbb{E}\left[\int_t^\infty e^{-\alpha(w-a)} p(\Delta(w)) dw \middle| Z_k = a, Z_{k+1} = t, E(t)\right]$$
$$= - \int_a^t e^{-\alpha(w-a)} p(w - a) dw - e^{-\alpha(t-a)} J_\alpha(0, E(t) - 1), \quad (23)$$

for $t \geq a$ and $E(t) > 0$.

Let's define $J(0, -1) := \infty$ so that (23) holds for the $E(t) = 0$ as well. Notice that the process G_t is driven by the random process $E(t)$ which is not conditioned on any particular value of $E(a)$ while being adapted to the filtration $\mathcal{F}_t$. However, for a policy solving (21), the stopping time Z_{k+1} depends on $E(a)$ as it maximizes $\mathbb{E}\left[G_{Z_{k+1}} \mid \mathcal{F}_a\right]$ which depends on $E(a)$ through the filtration $\mathcal{F}_\dashv$.

Accordingly, we define the stopping problem of maximizing the expected gain in the given interval $[a, \infty)$ as in the following:

$$\max_{t \in \mathfrak{M}_a} \mathbb{E}\left[G_t \mid \mathcal{F}_a\right]. \quad (24)$$

Based on this formulation, we will show that the optimal stopping time exists and is given by the following stopping rule for Z_{k+1}:

$$Z_{k+1} = \inf\{t \geq Z_k = a : G_t = S_t\}, \quad (25)$$

where S is the Snell envelope [31] for G:

$$S_t = \operatorname*{ess\,sup}_{t' \in \mathfrak{M}_t} \mathbb{E}\left[G_{t'} \mid \mathcal{F}_t\right]. \quad (26)$$

Showing that Z_{k+1} in (25) is finite w.p.1 is sufficient to prove the existence of the optimal stopping time and the optimality of the stopping rule in (25) (see [31, Theorem 2.2.]). Consider the lemma below and its proof in order to see the finiteness of Z_{k+1} in (25):

Lemma 3: For the stopping rule in (25) Z_{k+1} is finite w.p.1, i.e., $\Pr(Z_{k+1} < \infty) = 1$.

Proof: Consider the Markov time Q_{k+1} which is defined as follows:

$$Q_{k+1} := \inf\{t \geq Z_k = a : E(t) = B, G_t = S_t\}. \quad (27)$$

Clearly, the stopping time Z_{k+1} chosen in (25) is earlier than Q_{k+1} as Q_{k+1} has an additional stopping condition $E(t) = B$. This means that if $\Pr(Q_{k+1} < \infty) = 1$, then $\Pr(Z_{k+1} < \infty) = 1$.

Accordingly, for the proof of this lemma, it is sufficient to show that Q_{k+1} is finite w.p.1. We will show this by showing the

finiteness of (i) the first time $t \geq Z_k = a$ such that $E(t) = B$, and (ii) the duration between this time and the Markov time Q_{k+1}. Note that $E(t) = B$ condition is always satisfied after it reached for the first time. Let R_{k+1} be the Markov time representing the first time when $E(t) = B$ is satisfied:

$$R_{k+1} := \inf\{t \geq Z_k = a : E(t) = B\}. \tag{28}$$

(i) Observe that the Markov time R_{k+1} is finite w.p.1 as it is stochastically dominated by $a + Y_B$ where Y_B is an Erlang distributed random variable with parameter B which obeys (8) and $\Pr(Y_B < \infty) = 1$.

(ii) In order to see that $Q_{k+1} - R_{k+1}$ is also finite, consider the time period after R_{k+1}, i.e., $[R_{k+1}, \infty)$. As $E(t) = B$ for any $t \geq R_{k+1}$, the evolution of G_t becomes deterministic after $t \geq R_{k+1}$:

$$G_t = \\ -\int_a^t e^{-\alpha(w-a)}p(w-a)dw - e^{-\alpha(t-a)}J_\alpha(0, B-1), \tag{29}$$

for $t \geq R_{k+1}$.

On the other hand, for $t \geq R_{k+1}$, the Snell envelope is $S_t = \operatorname{ess\,sup}_{t' \in \mathfrak{M}_t} G_{t'} = \sup_{t' \geq t} G_{t'}$. We will show that G_t is always non-increasing after some finite time so that $S_t = G_t$ is always satisfied after that time.

In order to see this, consider the change in G_t for $t \geq R_{k+1}$. As

$$-\frac{\partial}{\partial t}\left[\int_a^t e^{-\alpha(w-a)}p(w-a)dw + e^{-\alpha(t-a)}J_\alpha(0, B-1)\right] \\ = e^{-\alpha(t-a)}\left(\alpha J_\alpha(0, B-1) - p(t-a)\right), \tag{30}$$

and $p(t-a)$ is non-decreasing, for $t \geq R_{k+1}$, G_t is non-increasing if $t \geq t_c$ for some t_c such that

$$t_c := \inf\{t \geq a : p(t-a) = \alpha J_\alpha(0, B-1)\}. \tag{31}$$

This implies that, for $t \geq \max\{R_{k+1}, t_c\}$, $G_t = \sup_{t' \geq t} G_{t'}$ and hence $S_t = G_t$. Accordingly, the stopping conditions of Q_{k+1} are satisfied for the first time when $t = \max\{R_{k+1}, t_c\}$ which means $Q_{k+1} = \max\{R_{k+1}, t_c\}$.

As $\alpha J_\alpha(0, B-1)$ is finite, t_c is finite which implies Q_{k+1} is finite w.p.1 as R_{k+1} is finite w.p.1. This completes the proof. □

We just showed that the Markov time in (25) is finite w.p.1 and this means that it is the optimal stopping time by [31, Theorem 2.2.]. Next, we show that the optimal stopping rule in (25) is a threshold policy by using the properties of the cost function in (20). To relate the optimal stopping time and the cost function in (20), we will express the Snell envelope in an alternative way.

Notice that the Snell envelope can be written by substituting (22) in (26) as follows:

$$S_t = \operatorname{ess\,sup}_{t' \in \mathfrak{M}_t} - \min_{\pi \in \Pi^{\text{online}}} \\ \mathbb{E}\left[\int_a^\infty e^{-\alpha(w-a)}p(\Delta(w))dw \,\middle|\, Z_k = a, Z_{k+1} = t', \mathcal{F}_t\right]. \tag{32}$$

Hence,

$$S_t = - \min_{\pi \in \Pi^{\text{online}}} \\ \mathbb{E}\left[\int_a^\infty e^{-\alpha(w-a)}p(\Delta(w))dw \,\middle|\, Z_k = a, Z_{k+1} \geq t, \mathcal{F}_t\right]. \tag{33}$$

Accordingly, using the definition of $J_\alpha(\Delta(a), E(a))$, we can write

$$S_t = - \int_a^t e^{-\alpha(w-a)}p(w-a)dw + e^{-\alpha(t-a)}J_\alpha(\Delta(t), E(t)). \tag{34}$$

Therefore, as the first terms in (29) and (34) are identical, the optimal stopping rule in (25) is equivalent to

$$Z_{k+1} = \inf\{t \geq Z_k = a : J_\alpha(0, E(t)-1) = J_\alpha(\Delta(t), E(t))\}. \tag{35}$$

Next, we show that the stopping rule in (35) is a threshold rule in age. In order to show this, let us define the function $\rho_\alpha(\cdot) : \{0, 1, \cdots, B\} \to [0, \infty)$ such that:

$$\rho_\alpha(\ell) := \inf\{w \geq 0 : J_\alpha(0, \ell-1) = J_\alpha(w, \ell)\}. \tag{36}$$

We can show that for any $\Delta \geq \rho_\alpha(\ell)$, it is guaranteed that $J_\alpha(0, \ell-1) = J_\alpha(\Delta, \ell)$ due to the following reasons:

• For any Δ and $\ell \in \{0, 1, 2, \cdots, B\}$, $J_\alpha(\Delta, \ell)$ is smaller than or equal to $J_\alpha(0, \ell-1)$ as :

$$J_\alpha(\Delta, \ell) = \min_{\pi \in \Pi^{\text{online}}} e^a \\ \mathbb{E}\left[\int_a^\infty e^{-\alpha w}p(\Delta(w))dw \,\middle|\, Z_k = t_a - \Delta, Z_{k+1} \geq t_a, E(t_a) = \ell\right] \\ \leq \min_{\pi \in \Pi^{\text{online}}} e^a \\ \mathbb{E}\left[\int_a^\infty e^{-\alpha w}p(\Delta(w))dw \,\middle|\, Z_k = t_a - \Delta, Z_{k+1} = a, E(t) = \ell\right] \\ = J_\alpha(0, \ell-1),$$

where the inequality is true as the expectation is conditioned on policies with $Z_{k+1} = t_a$.

• For any $\ell \in \{0, 1, 2, \cdots, B\}$, $J_\alpha(\Delta, \ell)$ is non-decreasing in Δ as :

$$J_\alpha(\Delta, \ell) \\ = \min_{\pi \in \Pi^{\text{online}}} \mathbb{E}\left[\int_a^{Z_{k+1}} e^{-\alpha(w-a)}p(w + \Delta - t_a)dw \,\middle|\, \theta(\Delta)\right] \\ + \mathbb{E}\left[\int_{Z_{k+1}}^\infty e^{-\alpha(w-a)}p(\Delta(w))dw \,\middle|\, \theta(\Delta)\right] \\ \leq \min_{\pi \in \Pi^{\text{online}}} \mathbb{E}\left[\int_a^{Z_{k+1}} e^{-\alpha(w-a)}p(w + \Delta' - t_a)dw \,\middle|\, \theta(\Delta)\right] \\ + \mathbb{E}\left[\int_{Z_{k+1}}^\infty e^{-\alpha(w-a)}p(\Delta(w))dw \,\middle|\, \theta(\Delta)\right] \\ = \min_{\pi \in \Pi^{\text{online}}} \mathbb{E}\left[\int_a^{Z_{k+1}} e^{-\alpha(w-a)}p(\Delta(w))dw \,\middle|\, \theta(\Delta')\right]$$

$$+ \mathbb{E}\left[\int_{Z_{k+1}}^{\infty} e^{-\alpha(w-a)} p(\Delta(w)) dw \middle| \theta(\Delta')\right]$$
$$= J_\alpha(\Delta', \ell),$$

for any $\Delta' \geq \Delta$ and $\theta(\Delta) := (Z_k = t_a - \Delta, Z_{k+1} \geq t_a, E(t_a) = \ell)$ where the inequality follows from the fact that $p(\cdot)$ is non-decreasing and the second equality is due to that, given Z_{k+1}, the integrated values are conditionally independent from Z_k.

Accordingly, $J_\alpha(\Delta, \ell) = J_\alpha(0, \ell - 1)$ for any $\ell \in \{0, 1, 2, \cdots, B\}$ and $\Delta \geq \rho_\alpha(\ell)$. Therefore, the stopping rule in (35) is equivalent to:

$$Z_{k+1} = \inf\{t \geq Z_k = a : \Delta(t) \geq \rho_\alpha(E(t))\}, \quad (37)$$

for $\ell \in \{0, 1, 2, \cdots, B\}$.

We showed that the stopping rule in (37) gives the optimal stopping time Z_{k+1} for a policy solving (21). Now, we can start discussing *Step 2* in order to show that the optimal stopping rule with the same structure also gives a solution to (6).

In this part (*Step 2*) of the proof, we will consider the optimal stopping rules in (37) while the discount factor α is vanishing to zero. Notice that the policy solving (21) is identified by $\rho_\alpha(\ell)$ due to (37). Let π_α and $\Delta_{\pi_\alpha}(t)$ be a policy obeying (37) and solving (21) for discount factor α and the age at time t for that policy, respectively. We will show the following

$$\lim_{\beta \downarrow 0} \lim_{t_f \to \infty} \frac{\int_0^{t_f} \mathbb{E}[p(\Delta_{\pi_\beta}(t))] dt}{t_f} =$$
$$\inf_{\pi \in \Pi^{\text{online}}} \limsup_{t_f \to \infty} \frac{\int_0^{t_f} \mathbb{E}[p(\Delta_\pi(t))] dt}{t_f}, \quad (38)$$

which implies that for any $\{\beta_n\}_{n \geq 1} \downarrow 0$ sequence, π_{β_n} converges to the policy solving (6).

To prove the equivalence in (38), we will use Feller's Tauberian theorem [33] (also see the Tauberian theorem in [34]) which can be stated as follows:

Theorem 7: (Feller 1971) Let $f(t)$ be a Lebesgue-measurable, bounded, real function. Then,

$$\liminf_{t_f \to \infty} \frac{\int_0^{t_f} f(t) dt}{t_f} \leq \liminf_{\alpha \downarrow 0} \alpha \int_0^{t_f} e^{-\alpha(t-a)} f(t) dt$$
$$\leq \limsup_{\alpha \downarrow 0} \alpha \int_0^{t_f} e^{-\alpha(t-a)} f(t) dt \leq \limsup_{t_f \to \infty} \frac{\int_0^{t_f} f(t) dt}{t_f}. \quad (39)$$

Moreover, if the central inequality is an equality, then all inequalities are equalities.

This theorem can be applied for the function $f(t) = \mathbb{E}\left[p(\Delta_{\pi_\beta}(t))\right]$ where $\beta > 0$ [4]. To simplify the inequalities for this case, let's define a function $J_{\alpha;\beta}(\Delta(a), E(a))$ for $\beta > 0$

---

[4]Note that the function $\mathbb{E}\left[p(\Delta_{\pi_\beta}(t))\right]$ is Lebesgue-measurable (as $p(\cdot)$ is non-decreasing) and bounded (as X_ks are bounded w.p.1 for a policy obeying (37)).

such that:

$$J_{\alpha;\beta}(\Delta(a), E(a)) :=$$
$$\int_a^{\infty} e^{-\alpha(t-a)} \mathbb{E}\left[p(\Delta_{\pi_\beta}(t)) \middle| \Delta(a), E(a)\right] dt. \quad (40)$$

Note that for $a = 0$:

$$J_{\alpha;\beta}(0, 0) := \int_a^{\infty} e^{-\alpha(t-a)} \mathbb{E}\left[p(\Delta_{\pi_\beta}(t))\right] dt.$$

Accordingly, we can apply Feller's Tauberian theorem for $f(t) = \mathbb{E}\left[p(\Delta_{\pi_\beta}(t))\right]$ when $a = 0$ giving:

$$\liminf_{t_f \to \infty} \frac{\int_0^{t_f} \mathbb{E}[p(\Delta_{\pi_\beta}(t))] dt}{t_f} \leq \liminf_{\alpha \downarrow 0} \alpha J_{\alpha;\beta}(0, 0) \leq$$
$$\limsup_{\alpha \downarrow 0} \alpha J_{\alpha;\beta}(0, 0) \leq \limsup_{t_f \to \infty} \frac{\int_0^{t_f} \mathbb{E}[p(\Delta_{\pi_\beta}(t))] dt}{t_f}. \quad (41)$$

We can show that the inequalities in (41) are satisfied with equality for any π_β with $\beta > 0$ as $\lim_{t_f \to \infty} \frac{\int_0^{t_f} \mathbb{E}[p(\Delta_{\pi_\beta}(t))] dt}{t_f}$ exists for any π_β with $\beta > 0$. To see this, consider the following lemma:

Lemma 4: For $\alpha > 0$ and $\{Z_{k+1}, k \geq 0\}$ with Z_{k+1} as in (37), the following holds:

$$\lim_{t_f \to \infty} \frac{\int_0^{t_f} \mathbb{E}[p(\Delta_{\pi_\alpha}(t))] dt}{t_f} =$$
$$\frac{\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n} \mathbb{E}[p(X_k)]}{\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n} X_k} \quad \text{w.p.1.} \quad (42)$$

Proof: The proof of Lemma 3 showed that for $Z_k = a$ and optimal stopping time solving (24) it is true that $\Pr(X_{k+1} \geq x) \leq \Pr(t_c - t_a + Y_B \geq x)$ where t_c is the deterministic time defined in (31) and Y_B is an Erlang distributed with parameter B which obeys (8). Accordingly, $\mathbb{E}[p(X_{k+1})]$ is finite as $\mathbb{E}[p(\alpha J_\alpha(0, B) + Y_B)]$ is finite for $\alpha > 0$. On the other hand, $\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n} X_k < \infty$ w.p.1 and $\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n} X_k > \frac{1}{\mu_H}$ w.p.1 due to the energy causality constraint. Therefore, we can apply the derivation steps in [35, Theorem 5.4.5] and obtain (42). This completes the proof. $\square$

Lemma 4 and (41) imply the following for for $a = 0$ and $\beta > 0$:

$$\lim_{\alpha \downarrow 0} \alpha J_{\alpha;\beta}(0, 0) = \lim_{t_f \to \infty} \frac{\int_0^{t_f} \mathbb{E}\left[p(\Delta_{\pi_\beta}(t))\right] dt}{t_f}. \quad (43)$$

Now, consider an arbitrary online policy π for which $\mathbb{E}\left[p(\Delta_\pi(t))\right]$ is Lebesgue-measurable and bounded, then apply Feller's Tauberian theorem for $f(t) = \mathbb{E}\left[p(\Delta_\pi(t))\right]$ giving the following inequality when $t_a = 0$:

$$\limsup_{\alpha \downarrow 0} \alpha \int_0^{\infty} e^{-\alpha(t-a)} \mathbb{E}\left[p(\Delta_\pi(t))\right] dt \leq$$
$$\limsup_{t_f \to \infty} \frac{\int_0^{t_f} \mathbb{E}[p(\Delta_\pi(t))] dt}{t_f}. \quad (44)$$

Note that for $\alpha > 0$, $J_{\alpha;\beta}(0, 0)$ is minimized for $\alpha = \beta$,

hence:

$$\lim_{\beta\downarrow0}\lim_{\alpha\downarrow0}\alpha J_{\alpha;\beta}(0,0)=\inf_{\beta>0}\lim_{\alpha\downarrow0}\alpha J_{\alpha;\beta}(0,0)$$
$$\leq\limsup_{\alpha\downarrow0}\alpha\int_0^\infty e^{-\alpha(t-a)}\mathbb{E}\left[p(\Delta_\pi(t))\right]dt. \quad (45)$$

Combining (43), (44) and (45), we get (38). This completes the proof.

APPENDIX B
THE PROOF OF THEOREM 2

Theorem 2 follows from the proof of Theorem 1. To prove the theorem it is sufficient to show that for any $\alpha>0$, $\rho_\alpha(\ell)$ (see (36)) is non-increasing in ℓ as this guarantees that the monotonicity of optimal thresholds holds for any sequence of α values that vanishes to zero. To see this, consider the following lemma and the argument provided below its proof:

Lemma 5: For $J(\cdot,\cdot)$ is the function defined in (20), $J_\alpha(0,\ell)-J_\alpha(0,\ell+1)$ is non-increasing in $\ell\in\{0,1,\cdots,B-1\}$ for any $\alpha\geq0$.

Proof:

First, consider the alternative formulation of $J_\alpha(r,\ell+1)$ in below:

$$J_\alpha(r,\ell+1)=\min_{\pi\in\Pi^{\mathrm{online}}}e^a\mathbb{E}[$$
$$\mathbb{E}\left[\int_a^\infty e^{-\alpha t}p(\Delta(t))dt\bigg|Z_{k+1},\Delta(t_a)=r,E(t_a)=\ell+1\right]],$$

where the outer expectation is taken over Z_{k+1}.

Let
$$K_{r,\ell+1}(z,\sigma):=$$
$$\Pr\left(Z_{k+1}=z,H(z)-H(a)=\sigma|\Delta(t_a)=r,E(t_a)=\ell+1\right)$$

be the joint distribution of $Z_{k+1}\in\mathfrak{M}_\mathfrak{a}$ and the energy harvested during $[a,z]$. Then, we can write $J_\alpha(r,\ell+1)$ as follows:

$$J_\alpha(r,\ell+1)=\min_{Z_{k+1}\in\mathfrak{M}_\mathfrak{a}}\sum_{\sigma=0}^\infty\int_{t_a}^\infty K_{r,\ell+1}(z,\sigma)e^a$$
$$\times\left[\int_a^z e^{-\alpha t}p(\Delta(t))dt+e^{-\alpha z}J_\alpha(0,\min\{\ell+\sigma,B-1\})\right]dz. \quad (46)$$

Similarly,

$$J_\alpha(r,\ell+2)=\min_{Z_{k+1}\in\mathfrak{M}_\mathfrak{a}}\sum_{\sigma=0}^\infty\int_{t_a}^\infty K_{r,\ell+2}(z,\sigma)e^a$$
$$\times\left[\int_a^z e^{-\alpha t}p(\Delta(t))dt+e^{-\alpha z}J_\alpha(0,\min\{\ell+1+\sigma,B-1\})\right]dz. \quad (47)$$

Now, let $K_{r,\ell+2}^*(z,\sigma)$ be the distribution corresponding to the update time $Z_{k+1}\in\mathfrak{M}_\mathfrak{a}$ that is optimal in (47), which means:

$$J_\alpha(r,\ell+2)=\sum_{\sigma=0}^\infty\int_{t_a}^\infty K_{r,\ell+2}^*(z,\sigma)e^a$$
$$\times\left[\int_a^z e^{-\alpha t}p(\Delta(t))dt+e^{-\alpha z}J_\alpha(0,\min\{\ell+1+\sigma,B-1\})\right]dz. \quad (48)$$

Clearly, $K_{r,\ell+2}^*(z,\sigma)$ is not necessarily the joint distribution corresponding the update time $Z_{k+1}\in\mathfrak{M}_\mathfrak{a}$ that is optimal for (46), hence:

$$J_\alpha(r,\ell+1)\leq\sum_{\sigma=0}^\infty\int_{t_a}^\infty K_{r,\ell+2}^*(z,\sigma)[\int_a^z e^{-\alpha(t-a)}p(\Delta(t))dt$$
$$+e^{-\alpha(z-a)}J_\alpha(0,\min\{\ell+\sigma,B-1\})]dz. \quad (49)$$

Combining (50) and (49) gives:

$$J_\alpha(r,\ell+1)-J_\alpha(r,\ell+2)\leq\sum_{\sigma=0}^\infty\int_a^\infty K_{r,\ell+2}^*(z,\sigma)e^{-\alpha(z-a)}$$
$$\times[J_\alpha(0,\min\{\ell+\sigma,B-1\})-J_\alpha(0,\min\{\ell+1+\sigma,B-1\})]dz. \quad (50)$$

which implies :

$$J_\alpha(r,\ell+1)-J_\alpha(r,\ell+2)\leq\max_{\sigma\in\{0,1,\cdots,B-\ell\}}$$
$$J_\alpha(0,\min\{\ell+\sigma,B-1\})-J_\alpha(0,\min\{\ell+1+\sigma,B-1\})$$

Now, consider the case when $r=0$ and $\ell=B-2$ for (50):

$$J_\alpha(0,B-1)-J_\alpha(0,B)\leq$$
$$\sum_{\sigma=0}^\infty\int_a^\infty K_{r,\ell+2}^*(z,\sigma)e^{-\alpha(z-a)}[J_\alpha(0,\min\{B-2+\sigma,B-1\})$$
$$-J_\alpha(0,\min\{B-1+\sigma,B-1\})]dz, \quad (51)$$

which implies:

$$J_\alpha(0,B-1)-J_\alpha(0,B)\leq J_\alpha(0,B-2)-J_\alpha(0,B-1). \quad (52)$$

Suppose that the inequality below is true for $j\geq\ell+1$:

$$J_\alpha(0,j+1)-J_\alpha(0,j+2)\leq J_\alpha(0,j)-J_\alpha(0,j+1). \quad (53)$$

Then, we have:

$$J_\alpha(0,\ell+1)-J_\alpha(0,\ell+2)$$
$$\leq\sum_{\sigma=0}^\infty\int_a^\infty K_{r,\ell+2}^*(z,\sigma)e^{-\alpha(z-a)}\times$$
$$[J_\alpha(0,\min\{\ell+\sigma,B-1\})-J_\alpha(0,\min\{\ell+1+\sigma,B-1\})]dz$$
$$\leq\int_a^\infty K^*(z,0)e^{-\alpha(z-a)}[J_\alpha(0,\ell)-J_\alpha(0,\ell+1)]dz+\sum_{\sigma=1}^\infty$$
$$\int_a^\infty K_{r,\ell+2}^*(z,\sigma)e^{-\alpha(z-a)}[J_\alpha(0,\ell+1)-J_\alpha(0,\ell+2)]dz$$
$$\leq J_\alpha(0,\ell)-J_\alpha(0,\ell+1). \quad (54)$$

This means that the inequality (53) is also true for $j = \ell$ so is for any $j = 0, 1, \cdots, B - 2$ by induction. Combining this and (51):

$$J_\alpha(r, \ell + 1) - J_\alpha(r, \ell + 2) \leq J_\alpha(0, \ell) - J_\alpha(0, \ell + 1), \quad (55)$$

for $\alpha \geq 0$, $r \geq 0$ and $\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

Lemma 5 shows that $\rho_\alpha(\ell)$ is non-increasing in ℓ for $\alpha > 0$. It is sufficient to consider (55) when $r = \rho_\alpha(\ell)$:

$$0 = J_\alpha(0, \ell-1) - J_\alpha(\rho_\alpha(\ell), \ell) \leq J_\alpha(0, \ell-2) - J_\alpha(\rho_\alpha(\ell), \ell-1), \quad (56)$$

which implies $\rho_\alpha(\ell - 1) \geq \rho_\alpha(\ell)$ combining

$$J_\alpha(0, \ell - 2) - J_\alpha(\rho_\alpha(\ell - 1), \ell - 1)$$

and that $J_\alpha(r, \ell - 1)$ is non-decreasing [5] in r. Accordingly, the optimal policies solving (19) are monotone threshold policies, i.e., $\pi_\alpha \in \Pi^{MT}$ for any $\alpha > 0$.

APPENDIX C
THE PROOF OF LEMMA 1

Let $\tau_{B+1} = 0$. Then, consider:

$$\frac{\partial}{\partial \tau_i} \mathbb{E}\left[X^2\right] = \frac{\partial}{\partial \tau_i} \int_0^\infty 2x \Pr(X \geq x) dx$$

$$= \frac{\partial}{\partial \tau_i} \sum_{i=0}^B \int_{\tau_{i+1}}^{\tau_i} 2x \Pr(X \geq x) dx$$

$$= 2\frac{\partial}{\partial \tau_i} \left[\int_{\tau_{i+1}}^{\tau_i} x \Pr(X \geq x) dx + \int_{\tau_i}^{\tau_{i-1}} x \Pr(X \geq x) dx \right]$$

$$= 2\tau_i \frac{\partial}{\partial \tau_i} \int_{\tau_{i+1}}^{\tau_{i-1}} \Pr(X \geq x) dx$$

$$= 2\tau_i \frac{\partial}{\partial \tau_i} \sum_{i=0}^B \int_{\tau_{i+1}}^{\tau_i} \Pr(X \geq x) dx = 2\tau_i \frac{\partial}{\partial \tau_i} \mathbb{E}\left[X\right],$$

for $i = 0, 1, \cdots, B$.

APPENDIX D
USEFUL RESULTS FOR ASYMPTOTIC PROPERTIES

Lemmas 6, 7, and 8 provide some useful results that combine ergodicity properties and renewal-reward theorem for a DTMC with transition probabilities in (10).

Lemma 6: The DTMC with the transition probabilities in (10) is ergodic for a monotone threshold policy where τ_1 is finite.

Proof: Consider an energy state j in $[0, B - 1]$. We will show that any other energy state i is reachable from j in at most $B - 1$ steps with a positive probability. For $i \geq j$, the higher energy state i is reachable from j in one step with a positive probability as for $i = B - 1$, $\Pr(Y_{B-j} \leq \tau_{B-1})$ is strictly

positive and for $j \leq i < B - 1$:

$$\Pr(Y_{1+i-j} \leq \tau_i) - \Pr(Y_{2+i-j} \leq \tau_{i+1}) \geq$$
$$\Pr(Y_{1+i-j} \leq \tau_{i+1}) - \Pr(Y_{2+i-j} \leq \tau_{i+1}) > 0,$$

as $\tau_{i+1} \leq \tau_i$ and $i - j \geq 0$.

Similarly, the energy state $i = j - 1$ for $j = 1, \cdots, B - 1$ can be reached from j with a probability $1 - \Pr(Y_1 \leq \tau_j)$ which is strictly positive as τ_j is finite. This means that any state $i < j$ can be reached from j in at most $B - 1$ steps with a positive probability. $\qquad\square$

Lemma 7: For monotone threshold policies with finite τ_1, the following is true:

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^n X_k = \sum_{j=0}^{B-1} \mathbb{E}\left[X \mid E = j\right] \Pr(E = j) \quad \text{w.p.1.} \tag{57}$$

$$\lim_{n \to +\infty} \frac{1}{2n} \sum_{k=0}^n \mathbb{E}[X_k^2] = \frac{1}{2} \sum_{j=0}^{B-1} \mathbb{E}\left[X^2 \mid E = j\right] \Pr(E = j), \tag{58}$$

where $\Pr(E = j)$ is the steady-state probability for energy state j, $\mathbb{E}\left[X \mid E = j\right] \triangleq \mathbb{E}\left[X_k \mid E(Z_k) = j\right]$ and $\mathbb{E}\left[X^2 \mid E = j\right] \triangleq \mathbb{E}\left[X_k^2 \mid E(Z_k) = j\right]$.

Proof: Consider:

$$\frac{1}{n} \sum_{k=0}^n X_k = \frac{1}{n} \sum_{j=0}^{B-1} \sum_{\substack{k \in [0,n] \\ E(Z_k)=j}} X_k = \frac{1}{n} \sum_{j=0}^{B-1} \sum_{\ell=0}^{L_j} X_{\ell;j},$$

where L_j is the number of ks in $[0, n]$ such that $E(Z_k) = j$ and $X_{\ell;j}$ is a r.v. with the CDF $\Pr(X_{\ell;j} \leq x) = \Pr(X_k \leq x \mid E(Z_k) = j)$ for some k.

Note that the sequence $X_{0;j}, X_{1;j}, \cdots, X_{L_j;j}$ is i.i.d. for any j and their mean is bounded as all thresholds are finite, hence:

$$\lim_{L_j \to \infty} \frac{1}{L_j} \sum_{\ell=0}^{L_j} X_{\ell;j} = \mathbb{E}\left[X \mid E = j\right], w.p.1.$$

Due to the ergodicity of $E(Z_k)$s (Lemma 6):

$$\lim_{n \to \infty} \frac{L_j}{n} = \Pr(E = j), w.p.1.$$

Therefore,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^n X_k = \lim_{n \to \infty} \sum_{j=0}^{B-1} \frac{L_j}{n} \left(\frac{1}{L_j} \sum_{\ell=0}^{L_j} X_{\ell;j} \right),$$

$$= \sum_{j=0}^{B-1} \mathbb{E}\left[X \mid E = j\right] \Pr(E = j), w.p.1.$$

---

[5]This fact is provided in the proof of Theorem 1.

Similarly,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} \mathbb{E}[X_k^2] = \lim_{n\to\infty} \sum_{j=0}^{B-1} \frac{L_j}{n} \left(\frac{1}{L_j} \sum_{\ell=0}^{L_j} X_{\ell;j}^2\right)$$

$$= \sum_{j=0}^{B-1} \mathbb{E}\left[X^2 \mid E = j\right] \Pr(E = j), w.p.1.$$

$\square$

Lemma 8: For a threshold policy where τ_1 is finite, the average age $\bar{\Delta}$ is finite (w.p.1) and given by the following expression.

$$\bar{\Delta} = \frac{\lim_{n\to+\infty} \frac{1}{2n} \sum_{k=0}^{n} \mathbb{E}[X_k^2]}{\lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n} X_k} \quad \text{w.p.1.} \tag{59}$$

Proof: The proof is a generalization of Theorem 5.4.5 in [35] for the case where X_ks are non-i.i.d. but the limits still exist (w.p.1). When X_ks are i.i.d. with $\mathbb{E}[X_k] < \infty$ and $\mathbb{E}[X_k^2] < \infty$, the convergence (w.p.1) of the limits is guaranteed. $\square$

APPENDIX E
THE PROOF OF THEOREM 3

Theorem 3 follows from the proof of Theorem 1. The proof of Lemma 3 shows that given that $Z_k = a$ is the last update time and $E(t') = B$ for some $t' > a$, the condition $S_t = G_t$ is satisfied for the first time when $t \geq \{t', t_c\}$ (see (31)). This means that $\rho_\alpha(B) = \alpha J_\alpha(0, B-1)$ for $\rho_\alpha(E(t))$ in (37). Accordingly,

$$p(\tau_B^*) = \lim_{\alpha\downarrow 0} \rho_\alpha(B) = \lim_{\alpha\downarrow 0} \alpha J_\alpha(0, B-1) =$$

$$\min_{\pi\in\Pi^{\text{online}}} \limsup_{t_f\to\infty} \frac{\int_0^{t_f} \mathbb{E}[p(\Delta_\pi(t)) \mid E(0)=B]dt}{t_f} = \bar{p}_{\pi^*},$$

which follows from the application of Feller's Tauberian theorem (applying Theorem 7 for $f(t) = \mathbb{E}[p(\Delta_\pi(t)) \mid E(0) = B]$). This completes the proof.

APPENDIX F
THE PROOF OF THEOREM 4

By Lemma 8 and Lemma 7, $\bar{\Delta}$ for $B = 1$ can be computed as follows

$$\bar{\Delta} = \frac{1}{2} \frac{\mathbb{E}\left[X^2 \mid E = 0\right] \Pr(E = 0)}{\mathbb{E}[X \mid E = 0] \Pr(E = 0)}, \tag{60}$$

where $\Pr(E = 0) = 1$, $\mathbb{E}[X^2 \mid E = 0] = \tau_1^2 + (\frac{2}{\mu_H^2} + \frac{2}{\mu_H}\tau_1)e^{-\mu_H\tau_1}$ and $\mathbb{E}[X \mid E = 0] = \tau_1 + \frac{1}{\mu_H}e^{-\mu_H\tau_1}$. Accordingly, $\bar{\Delta}$ is given by (13). By Theorem 3, $\tau_1^* = \bar{\Delta}_{\pi^*}$ and combining this with (13) results in

$$\mu_H\tau_1^* = \frac{\frac{1}{2}(\mu_H\tau_1^*)^2 + e^{-\mu_H\tau_1^*}(\mu_H\tau_1 + 1)}{\mu_H\tau_1^* + e^{-\mu_H\tau_1^*}}. \tag{61}$$

Solving (61) gives that $(\tau_1^*)^2 = \frac{2}{\mu_H}e^{-\mu_H\tau_1^*}$ which means $\tau_1^* = \frac{1}{\mu_H}2W(\frac{1}{\sqrt{2}})$.

APPENDIX G

THE PROOF OF THEOREM 5

By Lemma 8 and Lemma 7, $\bar{\Delta}$ for $B = 2$ is the following:

$$\bar{\Delta} =$$

$$\frac{1}{2} \frac{\mathbb{E}\left[X^2 \mid E = 0\right] \Pr(E = 0) + \mathbb{E}\left[X^2 \mid E = 1\right] \Pr(E = 1)}{\mathbb{E}[X \mid E = 0] \Pr(E = 0) + \mathbb{E}[X \mid E = 1] \Pr(E = 1)}. \tag{62}$$

The probability of being in $E = 1$, i.e. $\Pr(E = 1)$ can be solved using:

$$\Pr(E = 1) = \sum_{j=0}^{1} \Pr(E(Z_{k+1}) = 1 \mid E(Z_k) = j) \Pr(E = j). \tag{63}$$

Combining (63) and (9),

$$\Pr(E = 1) = \frac{e^{-\mu_H\tau_1}}{1 - e^{-\mu_H\tau_1}\mu_H\tau_1}. \tag{64}$$

Now, we can obtain $\mathbb{E}\left[X^2 \mid E = j\right]$, $\mathbb{E}[X \mid E = j]$ using (9). Combining these with (64) and substituting in (62) gives (5).

APPENDIX H
THE PROOF OF THEOREM 6

First, we show that $\tau_B \geq \bar{\Delta}_{\pi^*}$ is *necessary* to find a solution to (17) with monotonic non-increasing thresholds. Then, we show that this condition is also *sufficient*.

The *necessity* part of the proof follows from the fact that $\tau_B = \bar{\Delta}_\pi$ for any solution of (17), as $\bar{\Delta}_\pi = m_1(\tau_1, \tau_2, \cdots, \tau_B)/2m_2(\tau_1, \tau_2, \cdots, \tau_B)$ by Lemma 8 and Lemma 7. Therefore, by the optimality of $\bar{\Delta}_{\pi^*}$, $\tau_B \geq \bar{\Delta}_{\pi^*}$ must hold for any solution of (17).

Now, we consider the *sufficiency* part of the proof where it is useful to define a function $\phi : [0,\infty)^B \to \mathbb{R}$ as follows:

$$\phi(\tau_B, \tau_{B-1} - \tau_B, \cdots, \tau_1 - \tau_2) \triangleq$$
$$2\tau_B m_1(\tau_1, \tau_2, \cdots, \tau_B) - m_2(\tau_1, \tau_2, \cdots, \tau_B).$$

Using this definition, (17) can be written as,

$$\phi(\tau_B, \tau_{B-1} - \tau_B, \cdots, \tau_1 - \tau_2) = 0.$$

We need to show that given $\tau_B \geq \bar{\Delta}_{\pi^*}$, one can find a set of non-negative real numbers $d_1, \cdots, d_{B-1}$ such that $\phi(\tau_B, d_{B-1}, \cdots, d_1) = 0$. Accordingly, τ_B and $d_1, \cdots, d_{B-1}$ constitute a solution to (17) with monotonic non-decreasing thresholds where $\tau_i = \tau_{i+1} + d_i$, for $i = 1, \cdots, B - 1$. In order to prove this, let us start with the optimal policy $\pi^* = (\tau_1^*, \tau_2^* \cdots, \tau_B^*)$ where we know that $\tau_B^* = \bar{\Delta}_{\pi^*}$ by Theorem 3. Starting from the optimal policy π^*, the policy will be modified following the procedure below:

- *Phase 1:* Modify the policy $\pi^{(+)} = (\tau_1^{(+)}, \tau_2^{(+)}, \cdots, \tau_B^{(+)})$ from the previous phase to the policy $\pi^{(-)} = (\tau_1^{(-)}, \tau_2^{(-)} \cdots, \tau_B^{(-)})$ so that $\tau_B^{(-)} = \min\{\tau_{B-1}^{(+)}, \tau_B\}$ while $\tau_i^{(-)} = \tau_i^{(+)}$, for $i = 1, \cdots, B - 1$. Then, go to *Phase 2* with policy $\pi^{(-)}$.
- *Phase 2:* Modify the policy $\pi^{(-)} = (\tau_1^{(-)}, \tau_2^{(-)}, \cdots, \tau_B^{(-)})$ from

the previous phase to the policy $\pi^{(+)} = (\tau_1^{(+)}, \tau_2^{(+)} \cdots, \tau_B^{(+)})$ so that $\tau_B^{(+)} = \tau_B^{(-)}$ while $\tau_i^{(+)} = \tau_i^{(-)} + x$ for $i = 1, \cdots, B - 1$ where $x > 0$ is the solution of the following:

$$\phi(\tau_B^{(-)}, \tau_{B-1}^{(-)} - \tau_B^{(-)} + x, \cdots, \tau_1^{(-)} - \tau_2^{(-)} + x) = 0. \quad (65)$$

If $\tau_B^{(-)} = \tau_B$, the procedure stops and (65) gives the solution that $\phi(\tau_B, d_{B-1}, \cdots, d_1) = 0$, otherwise go to *Phase 1* with policy $\pi^{(+)}$.

It can be shown that the procedure always stops with a solution that $\phi(\tau_B, d_{B-1}, \cdots, d_1) = 0$. To see this, first observe that (65) always has a solution as long as:

$$\phi(\tau_B^{(-)}, \tau_{B-1}^{(-)} - \tau_B^{(-)}, \cdots, \tau_1^{(-)} - \tau_2^{(-)}) > 0. \quad (66)$$

This is due to the following facts about the function $\phi(\tau_B^{(-)}, \tau_{B-1}^{(-)} - \tau_B^{(-)} + x, \cdots, \tau_1^{(-)} - \tau_2^{(-)} + x)$: (i) It is a continuous function of x, (ii) it goes to $-\infty$ as x grows.

Next, observe that (66) always holds, i.e.,

$$\phi(\tau_B^{(-)}, \tau_{B-1}^{(-)} - \tau_B^{(-)}, \cdots, \tau_1^{(-)} - \tau_2^{(-)}) =$$

$$\underbrace{\phi(\tau_B^{(+)}, \tau_{B-1}^{(+)} - \tau_B^{(+)}, \cdots, \tau_1^{(+)} - \tau_2^{(+)})}_{=0 \text{ due to the } Phase\ 2 \text{ or the initial/optimal policy}} + \int_{\pi^{(+)}}^{\pi^{(-)}} d\phi,$$

is positive. This can be seen by considering:

$$\frac{\partial \phi}{\partial \tau_B} = 2m_1(\tau_1, \tau_2, \cdots, \tau_B)$$

$$+ \sum_{j=0}^{B-1} \left[2\tau_B \frac{\partial}{\partial \tau_B} \mathbb{E}[X \mid E = j] - \frac{\partial}{\partial \tau_B} \mathbb{E}[X^2 \mid E = j] \right]$$

$$\times \Pr(E = j),$$

which follows from the fact that $\Pr(E = j)$ does not depend on τ_B (see (10)) and can be further simplified by Lemma 1, hence:

$$\frac{\partial \phi}{\partial \tau_B} = 2m_1(\tau_1, \tau_2, \cdots, \tau_B).$$

Accordingly, we have:

$$\phi(\tau_B^{(-)}, \tau_{B-1}^{(-)} - \tau_B^{(-)}, \cdots, \tau_1^{(-)} - \tau_2^{(-)}) = \int_{\pi^{(+)}}^{\pi^{(-)}} d\phi$$

$$= 2 \int_{\tau_B^{(+)}}^{\tau_B^{(-)}} m_1(\tau_1^{(+)}, \tau_2^{(+)}, \cdots, \tau) d\tau > 0,$$

where the inequality follows from the fact that $m_1(\tau_1^{(+)}, \tau_2^{(+)}, \cdots, \tau)$ being the average inter-update time is always positive.

Therefore, (65) can be always satisfied in *Phase 2*. Also, as the second smallest threshold is strictly increased in *Phase 2*, the smallest threshold can be moved toward τ_B in *Phase 1*. Also, it can be shown that the procedure does not converge any policy other than the policy that $\phi(\tau_B, d_{B-1}, \cdots, d_1) = 0$. This can be seen considering the following:

$$\frac{d}{dx} m_2(\tau_1 + x, \tau_2 + x, \cdots, \hat{\tau}_B) \mid_{x=0}$$

$$< \lim_{x \to 0} \lim_{n \to +\infty} \frac{1}{nx} \sum_{k=0}^{n} \left(\mathbb{E}[(X_k + x)^2] - \mathbb{E}[X_k^2] \right)$$

$$= \lim_{n \to +\infty} \frac{2}{n} \sum_{k=0}^{n} \mathbb{E}[X_k] = 2m_1(\tau_1, \tau_2, \cdots, \hat{\tau}_B),$$

hence,

$$\frac{d}{dx} \phi(\hat{\tau}_B, \tau_{B-1} - \hat{\tau}_B + x, \cdots, \tau_1 - \tau_2 + x) \mid_{x=0} \quad (67)$$

$$+ \frac{d}{dx} \phi(\hat{\tau}_B + x, \tau_{B-1} - \hat{\tau}_B - x, \cdots, \tau_1 - \tau_2) \mid_{x=0} \quad (68)$$

$$> 2\hat{\tau}_B \frac{d}{dx} m_1(\tau_1 + x, \tau_2 + x, \cdots, \hat{\tau}_B) \mid_{x=0}, \quad (69)$$

which implies that the procedure cannot converge to a policy with $\tau_B^{(+)} < \tau_B$ as the RHS of (67) is positive [6] and does not vanish for a finite set of thresholds. Therefore, as the smallest threshold of the policies modified by the procedure is increased up to τ_B, a solution that $\phi(\tau_B, d_{B-1}, \cdots, d_1) = 0$ is eventually reached. This completes the proof.

REFERENCES

[1] B. T. Bacinoglu, Y. Sun, E. Uysal-Biyikoglu, and V. Mutlu, "Achieving the age-energy tradeoff with a finite-battery energy harvesting source," in *Proc. IEEE ISIT*, June 2018.

[2] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *Proc. IEEE SECON*, June 2011.

[3] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?," in *Proc. IEEE INFOCOM*, Mar. 2012.

[4] R. Zviedris, A. Elsts, G. Strazdins, A. Mednis, and L. Selavo, "Lynxnet: Wild animal monitoring using sensor networks," in *Proc. REALWSN*, Dec. 2010.

[5] K. R. Chevli *et al.*, "Blue force tracking network modeling and simulation," in *Proc. MILCOM*, Oct. 2006.

[6] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the wiener process over a channel with random delay," in *Proc. ISIT*, June 2017.

[7] T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu, "Age of information under energy replenishment constraints," in *Proc. Info. Theory Appl. Workshop*, Feb. 2015.

[8] T. Bacinoglu and E. Uysal-Biyikoglu, "Scheduling status updates to minimize age of information with an energy harvesting sensor," in *Proc. IEEE ISIT*, June 2017.

[9] X. Wu, J. Yang, and J. Wu, "Optimal status update for age of information minimization with an energy harvesting source," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 1, pp. 193–204, 2017.

[10] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *Proc. IEEE ISIT*, June 2015.

[11] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor, "Age-minimal online policies for energy harvesting sensors with incremental battery recharges," in *Proc. IEEE ITA*, Feb. 2018.

[12] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor, "Age-minimal transmission for energy harvesting sensors with finite batteries: Online policies," [Online]. Available:http://arxiv.org/abs/1806.07271

[13] S. Feng and J. Yang, "Minimizing age of information for an energy harvesting source with updating failures," in *Proc. IEEE ISIT*, June 2018.

[14] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor, "Online timely status updates with erasures for energy harvesting sensors," in *Proc. IEEE Allerton*, Oct. 2018.

[15] J. Cho and H. Garcia-Molina, "Effective page refresh policies for web crawlers," *ACM Trans. Database Syst.*, vol. 28, no. 4, pp. 390–426, 2003.

[16] Y. Sun, E. Uysal-Biyikoglu, R. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," in *Proc. IEEE INFOCOM*, Apr. 2016.

[17] Y. Sun, E. Uysal-Biyikoglu, R. 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, Nov. 2017.

[18] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "Age and value of information: Non-linear age case," in *Proc. IEEE ISIT*, June 2017.

[19] Y. Sun and B. Cyr, "Sampling for data freshness optimization: Nonlinear age functions," *J. Commun. Networks*, in press, 2019.

[20] S. Razniewski, "Optimizing update frequencies for decaying information," in *Proc. ACM CIKM*, Oct. 2016.

[6] This follows from the fact that any increase in thresholds causes an increase in the battery overflow probability which means an increase in the average inter-update duration, i.e, $m_1(\tau_1, \tau_2, \cdots, \hat{\tau}_B)$.

[21] C. Kam, S. Kompella, and A. Ephremides, "Age of information under random updates," in *Proc. IEEE ISIT*, July 2013.

[22] M. Costa, M. Codreanu, and A. Ephremides, "Age of information with packet management," in *Proc. IEEE ISIT*, June 2014.

[23] L. Huang and E. Modiano, "Optimizing age-of-information in a multiclass queueing system," in *Proc. IEEE ISIT*, June 2015.

[24] N. Pappas, J. Gunnarsson, L. Kratz, M. Kountouris, and V. Angelakis, "Age of information of multiple sources with queue management," in *Proc. IEEE ICC*, June 2015.

[25] C. Kam, S. Kompella, G. D. Nguyen, and A. Ephremides, "Effect of message transmission path diversity on status age," *IEEE Trans. Inf. Theory*, vol. 62, no. 3, pp. 1360–1374, 2016.

[26] E. Najm and R. Nasser, "Age of information: The gamma awakening," in *Proc. IEEE ISIT*, July 2016.

[27] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Trans. Inf. Theory*, vol. 65, no. 3, pp. 1807–1827, 2019.

[28] E. Najm, R. Yates, and E. Soljanin, "Status updates through m/g/1/1 queues with harq," in *Proc. IEEE ISIT*, June 2017.

[29] S. Farazi, A. G. Klein, and D. R. Brown, 'Average age of information for status update systems with an energy harvesting server," in *Proc. IEEE INFOCOM*, April 2018.

[30] A. R. Al-Ali, I. Zualkernan, and F. Aloul, "A mobile gprs-sensors array for air pollution monitoring," *IEEE Sensors J.*, vol. 10, no. 10, pp. 1666–1671, 2010.

[31] G. Peskir and A. Shiryaev, Optimal Stopping and Free-Boundary Problems, ser. Lectures in Mathematics. ETH Zurich. Birkhauser Basel, 2006.

[32] X. Guo and O. Hernandez-Lerma, Continuous-Time Markov Decision Processes. Berlin, Springer, Science Pres, 2009.

[33] W. Feller, An Introduction to Probability Theory and Its Applications, Vol. 2, 2nd Edition. John Wiley and Sons, New York, 1971.

[34] D. V. Widder, The Laplace transform. Princeton, Princeton University Press, 1946.

[35] R. Gallager, Stochastic Processes: Theory for Applications. Cambridge University Press, 2013.
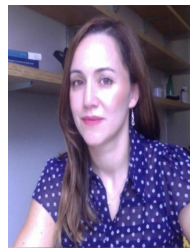
Elif Uysal is a Professor in the Department of Electrical and Electronics Engineering at the Middle East Technical University (METU), in Ankara, Turkey. She received the Ph.D. degree in EE from Stanford University in 2003, the S.M. degree in EECS from the Massachusetts Institute of Technology (MIT) in 1999 and the B.S. degree from METU in 1997. From 2003-05 she was a lecturer at MIT, and from 2005-06 she was an Assistant Professor at the Ohio State University (OSU). Since 2006, she has been with METU, and held visiting positions at OSU and MIT during 2014-2016. Her research interests are at the junction of communication and networking theories, with particular application to energy-efficient wireless networking. Dr. Uysal is a recipient of the 2014 Young Scientist Award from the Science Academy of Turkey, an IBM Faculty Award (2010), the Turkish National Science Foundation Career Award (2006), an NSF Foundations on Communication research grant (2006-2010), the MIT Vinton Hayes Fellowship, and the Stanford Graduate Fellowship. She is an editor for the IEEE/ACM Transactions on Networking served as associate editor for the IEEE Transactions on Wireless Communication (2014-2018), and track 3 co-chair for IEE PIMRC 2019.

Faruk Volkan Mutlu is a predoctoral student at Northeastern University, in Boston, USA. He received the B.S. degree from the Middle East Technical University (METU) in 2018, in Ankara, Turkey. His research interests include data-centric network architectures and optimization of wireless networks.

Baran Tan Bacinoglu received the B.S. and M.Sc. degrees from Middle East Technical University (METU), Ankara, Turkey, in 2009 and 2013, respectively, where he is currently pursuing the Ph.D. degree in Electrical and Electronics Engineering. He has been a Research Assistant with the Department of Electrical and Electronics Engineering, METU, since 2013. His current research interests are optimization in energy constrained communication networks, information-update systems, networking, wireless communication, and information theory.

Yin Sun received his B.Eng. and Ph.D. degrees in Electronic Engineering from Tsinghua University, in 2006 and 2011, respectively. At Tsinghua, he received the Excellent Doctoral Thesis Award of Tsinghua University, among many awards and scholarships. He was a Postdoctoral Scholar and Research Associate at the Ohio State University during 2011-2017. Since fall 2017, He joined Auburn University as an assistant professor in the Department of Electrical and Computer Engineering. His research interests include wireless communications, communication networks, information freshness, information theory, and machine learning. He is the founding co-chair of the first and second Age of Information Workshops, in conjunction with the IEEE INFOCOM 2018 and 2019. The paper he co-authored received the best student paper award at IEEE WiOpt 2013.