

# Age-Minimal Online Policies for Energy Harvesting Sensors with Incremental Battery Recharges

Ahmed Arafa<sup>1</sup>, Jing Yang<sup>2</sup>, Sennur Ulukus<sup>3</sup>, and H. Vincent Poor<sup>1</sup>

<sup>1</sup>Electrical Engineering Department, Princeton University

<sup>2</sup>School of Electrical Engineering and Computer Science, Pennsylvania State University

<sup>3</sup>Department of Electrical and Computer Engineering, University of Maryland

**Abstract**—A sensor node that is sending measurement updates regarding some physical phenomenon to a destination is considered. The sensor relies on energy harvested from nature to transmit its updates, and is equipped with a finite  $B$ -sized battery to save its harvested energy. Energy recharges the battery incrementally in units, according to a Poisson process, and one update consumes one energy unit to reach the destination. The setting is *online*, where the energy arrival times are revealed causally after the energy is harvested. The goal is to update the destination in a timely manner, namely, such that the long term average *age of information* is minimized, subject to energy causality constraints. The age of information at a given time is defined as the time spent since the latest update has reached the destination. It is shown that the optimal update policy follows a *renewal structure*, where the inter-update times are independent, and the time durations between any two consecutive events of submitting an update and having  $k$  units of energy remaining in the battery are independent and identically distributed for a given  $k \leq B - 1$ . The optimal renewal policy for the case of  $B = 2$  energy units is explicitly characterized, and it is shown that it has an *energy-dependent threshold structure*, where the sensor updates only if the age grows above a certain threshold that is a function of the amount of energy in its battery.

## I. INTRODUCTION

An energy harvesting sensor monitors some physical phenomenon and sends measurement updates about it to a destination. Updates are to be sent such that the long term average *age of information* is minimized. The age of information is the time spent since the freshest update has reached the destination. The sensor relies on energy harvested from nature to measure and send its updates, and is equipped with a finite  $B$ -sized battery to save its incoming energy. We characterize optimal *online* policies for this problem, where the sensor has only causal knowledge of the energy harvesting process.

In this work, we connect results from the energy harvesting communication literature and the age of information minimization literature by using the age of information metric as a means to assess the performance of a single-user energy harvesting communication channel. The energy harvesting communication literature is broadly categorized into offline and online settings, depending on whether the energy arrival times/amounts are known prior to the start of communication.

Offline energy management works consider, e.g., single-user channels [1]–[4]; multiuser channels [5]–[9]; and multi hop and relay channels [10]–[14]. Recent online works include the near-optimal results for single-user and multiuser channels [15]–[18], systems with processing costs [19], and systems with general utilities [20].

Age of information minimization is generally studied in a queuing-theoretic framework, including a single source setting [21]; multiple sources [22]; variations of the single source setting such as randomly (out of order) arriving updates [23], update management and control [24], and nonlinear age metrics [25], [26]; multi hop networks [27]; broadcasting, multicasting, and multi streaming [28]–[30]; coding over erasures [31]; and caching systems [32].

Assessing the performance of energy harvesting communication systems by the age of information metric has recently gained some attention [33]–[40]. Except for [36], an underlying assumption in these works is that energy expenditure is normalized, i.e., it takes one energy unit to send an update to the destination. References [33], [34] study a system with an infinite-sized battery, with [33] considering online scheduling with random service times (time for the update to take effect), and [34] considering offline and online scheduling with zero service times. The offline policy in [34] is extended to fixed non-zero service times in [35] for single and multi hop settings, and to energy-controlled service times in [36]. The online policy in [34] is found by dynamic programming in a discrete-time setting, and was shown to be of a threshold structure, where an update is sent only if the age of information is higher than a certain threshold. Motivated by the results in the infinite battery case, [37] then analyzes the performance of threshold policies under a finite-sized battery and varying channel assumptions, yet with no claim of optimality. Reference [38] proves the optimality of threshold policies when the battery size is equal to one unit using tools from renewal theory; it also provides an asymptotically optimal update policy when the battery size grows infinitely large. In our recent work [39], we extend the results of [38] and formally prove the optimality of threshold policies for any finite-sized battery in an online setting where the battery is randomly *fully* recharged over time, i.e., whenever energy is harvested, it completely fills up the battery. An interesting result is recently reported in [40], where status updates send

This research was supported in part by the National Science Foundation under Grants ECCS-1549881, ECCS-1647198, ECCS-1650299, CCF 14-22111, and CNS 15-26608.

information, other than that related to measurements, in an energy harvesting single-user channel.

In this work, we complement our results in [39] and study age-optimal online policies for an energy harvesting sensor with a finite battery with random *incremental* battery recharges; that is, energy is harvested in units as in [37], [38], as opposed to full chunks as in [39]. We extend the unit battery results of [38] and show that for a finite battery of size  $B$ , the optimal status update policy that minimizes the long term average age of information is a *renewal policy*: the times in between the two consecutive events where the sensor sends an update and has  $k$  energy units remaining in its battery, for some  $0 \leq k \leq B - 1$ , are independent and identically distributed (i.i.d.). Further, we show that inter-update times are independent. Based on these results, we explicitly solve the case of  $B = 2$  energy units, and formally prove, using optimization tools, that the optimal policy is an *energy-dependent threshold policy*: the sensor submits an update only if the instantaneous age of information is above a certain threshold that depends on the energy in its battery.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a sensor node that collects measurements from a physical phenomenon and sends updates to a destination over time. The sensor relies on energy harvested from nature to acquire and send its updates, and is equipped with a battery of finite size  $B$  to save its incoming energy. The sensor consumes one unit of energy to measure and send out an update to the destination. We assume that updates are sent over an error-free link with negligible transmission times as in [34], [37]–[39]. Energy arrives (is harvested) one unit at a time, at times  $\{t_1, t_2, \dots\}$  according to a Poisson process of rate 1. Our setting is online in which energy arrival times are revealed causally over time; only the arrival rate is known a priori.

Let  $s_i$  denote the time at which the sensor acquires (and transmits) the  $i$ th measurement update, and let  $\mathcal{E}(t)$  denote the amount of energy remaining in the battery at time  $t$ . We then have the following energy causality constraint [1]

$$\mathcal{E}(s_i^-) \geq 1, \quad \forall i \quad (1)$$

We assume that we begin with an empty battery at time 0, and that the battery evolves as follows over time

$$\mathcal{E}(s_i^-) = \min \{ \mathcal{E}(s_{i-1}^-) - 1 + \mathcal{A}(x_i), B \} \quad (2)$$

where  $x_i \triangleq s_i - s_{i-1}$ , and  $\mathcal{A}(x_i)$  denotes the amount of energy harvested in  $[s_{i-1}, s_i)$ . Note that  $\mathcal{A}(x_i)$  is a Poisson random variable with parameter  $x_i$ . We denote by  $\mathcal{F}$ , the set of feasible transmission times  $\{s_i\}$  described by (1) and (2) in addition to an empty battery at time 0, i.e.,  $\mathcal{E}(0) = 0$ .

The goal is to choose an online feasible transmission policy  $\{s_i\}$  (or equivalently  $\{x_i\}$ ) such that the long term average of the age of information experienced at the destination is minimized. The age of information is defined as the time elapsed since the latest update has reached the destination.

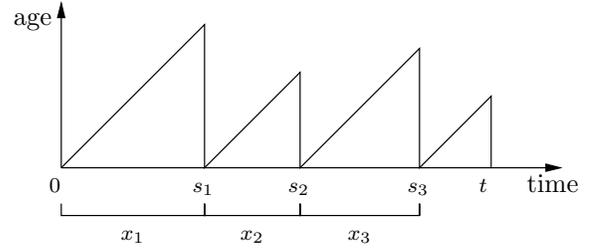


Fig. 1. Example of the age evolution versus time with  $n(t) = 3$ .

The age at time  $t$  is formally defined as

$$a(t) \triangleq t - u(t) \quad (3)$$

where  $u(t)$  is the time stamp of the latest update received before time  $t$ . Let  $n(t)$  denote the total number of updates sent by time  $t$ . We are interested in minimizing the area under the age curve, see Fig. 1 for a possible sample path with  $n(t) = 3$ . At time  $t$ , this area is given by

$$r(t) \triangleq \frac{1}{2} \sum_{i=1}^{n(t)} x_i^2 + \frac{1}{2} (t - s_{n(t)})^2 \quad (4)$$

and therefore the goal is to characterize the following quantity

$$\bar{r} \triangleq \min_{x \in \mathcal{F}} \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E}[r(T)] \quad (5)$$

where  $\mathbb{E}(\cdot)$  is the expectation operator. In the next section, we characterize the structure of the optimal policy.

## III. OPTIMAL SOLUTION STRUCTURE: RENEWAL TYPE POLICIES

In this section, we show that the optimal update policy that solves problem (5) has a renewal structure. Namely, we show that it is optimal to transmit updates in such a way that the inter-update delays are independent over time; and that the time durations in between the two consecutive events of transmitting an update and having  $k \leq B - 1$  units of energy left in the battery are i.i.d., i.e. these events occur at times that constitute a renewal process. We first introduce some notation.

Let the pair  $(\mathcal{E}(t), a(t))$  represent the state of the system at time  $t$ . Fix  $k \in \{0, 1, \dots, B - 1\}$ , and consider the state  $(k, 0)$ , which means that the sensor has just submitted an update and has  $k$  units of energy remaining in its battery. Let  $l_i$  denote the time at which the system visits  $(k, 0)$  for the  $i$ th time. We use the term *epoch* to denote the time in between two consecutive visits to  $(k, 0)$ . Observe that there can possibly be an infinite number of updates occurring in an epoch, depending on the energy arrival pattern and the update time decisions. For instance, in the  $i$ th epoch, which starts at  $l_{i-1}$ , one energy unit may arrive at some time  $l_{i-1} + \tau_{1,i}$ , at which the system goes to state  $(k+1, \tau_{1,i})$ , and then the sensor updates afterwards to get the system state back to  $(k, 0)$  again. Another possibility (if  $k \geq 1$ ) is that the sensor first updates at some time  $l_{i-1} + x_{k,i}$ , at which the system goes to state  $(k-1, 0)$ , and then two consecutive energy units arrive at times  $l_{i-1} + \tau_{1,i}$  and

$l_{i-1} + \tau_{1,i} + \tau_{2,i}$ , respectively, at which the system goes to state  $(k+1, \tau_{1,i} + \tau_{2,i})$ , and then the sensor updates afterwards to get the system state back to  $(k, 0)$  again. Depending on how many energy arrivals occur in the  $i$ th epoch, how far apart from each other they are, and the status update times, one can determine the length of the  $i$ th epoch and how many updates it has. Observe that the update policy in the  $i$ th epoch may depend on the history of events (energy arrivals and transmission updates) that occurred in previous epochs, which we denote by  $\mathcal{H}_{i-1}$ . Our main result in this section shows that this is not the case, under some mild technical conditions, and that epoch lengths should be i.i.d. We first have the following definition.

**Definition 1 (Uniformly Bounded Policy)** *An online policy whose inter-update times, as a function of the energy arrival times, have a bounded second moment.*

We focus on uniformly bounded policies as per Definition 1. Such policies were also considered in [38] in the analysis of the  $B = 1$  case. We now have the following theorem; the proof is in Appendix A.

**Theorem 1** *In the optimal solution of problem (5), any uniformly bounded policy is a renewal policy. That is, the sequence  $\{l_i\}$  denoting the times at which the system visits state  $(k, 0)$  forms a renewal process.*

Based on Theorem 1, the following corollary now follows.

**Corollary 1** *In the optimal solution of problem (5), the inter-update times are independent.*

**Proof:** Observe that whenever an update occurs the system enters state  $(j, 0)$  for some  $j \leq B - 1$ . The system then starts a new epoch with respect to state  $(j, 0)$ . Since the choice of  $k$  energy units in Theorem 1 is arbitrary, the results of the theorem now tell us that the update policy in that epoch, and therefore its length, is independent of the past history, in particular the past inter-update lengths. ■

In the next section, we show how to use the results of Theorem 1 and Corollary 1 to provide an explicit solution for the case of  $B = 2$  energy units.

#### IV. THE CASE $B = 2$

Based on Corollary 1, we now introduce the following notation regarding the update policy in a given epoch. Starting from state  $(0, 0)$  at time  $l_0$ , the sensor has to wait for the first energy arrival in the epoch, which occurs after some time  $\tau_1$ , and at which the system state becomes  $(1, \tau_1)$ . Since the sensor now has energy, it schedules its next update at  $l_0 + y_1(\tau_1)$ , for some function  $y_1(\cdot)$  to be optimally characterized. Now if another energy arrival occurs at time  $l_0 + \tau_1 + \tau_2$ , with  $\tau_2 > y_1(\tau_1) - \tau_1$ , the sensor transmits the update as scheduled at  $l_0 + y_1(\tau_1)$  and the system state returns to  $(0, 0)$  again. On the other hand, if this second energy arrival occurs relatively early, i.e.,  $\tau_2 \leq y_1(\tau_1) - \tau_1$ , the system state becomes  $(2, \tau_1 + \tau_2)$  at  $l_0 + \tau_1 + \tau_2$ , and the sensor *reschedules* its update to be at

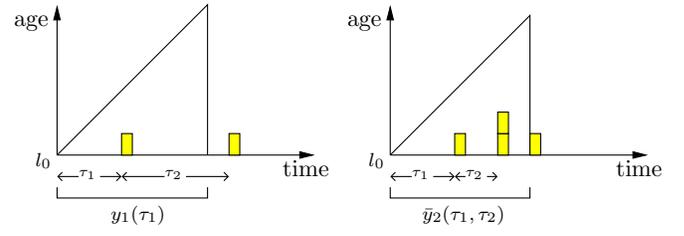


Fig. 2. Age of information versus time under the two possible ways of updating starting from state  $(0, 0)$  at time  $l_0$ . On the left, the second energy arrival occurs late, and hence we have one energy arrival followed by one update, returning to state  $(0, 0)$  again at  $l_0 + y_1(\tau_1)$ . On the right, the second energy arrival occurs early, and hence we have two energy arrivals followed by one update, entering state  $(1, 0)$  at  $l_0 + \bar{y}_2(\tau_1, \tau_2)$ . The yellow boxes represent energy units in the battery.

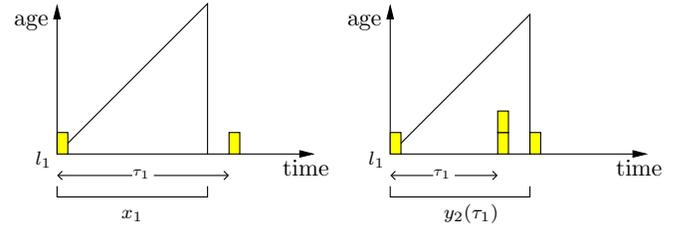


Fig. 3. Age of information versus time under the two possible ways of updating starting from state  $(1, 0)$  at time  $l_1$ . On the left, the first energy arrival occurs late, and hence the sensor updates and enters state  $(0, 0)$  at  $l_1 + x_1$ . On the right, the first energy arrival occurs early, and hence we have one energy arrival followed by an update, returning to state  $(1, 0)$  at  $l_1 + y_2(\tau_1)$ . The yellow boxes represent energy units in the battery.

$l_0 + \bar{y}_2(\tau_1, \tau_2)$  instead of  $l_0 + y_1(\tau_1)$ . Note that it is not clear so far whether  $\bar{y}_2(\tau_1, \tau_2)$  depends only on the age  $\tau_1 + \tau_2$ ; we leave it as a general function of the pair  $(\tau_1, \tau_2)$  for now. The above two cases are illustrated in Fig. 2.

Once the sensor has two energy units in its battery, it will eventually send an update making the system state become  $(1, 0)$  at some time  $l_1$ . The sensor then schedules its next update at  $l_1 + x_1$ , for some  $x_1$  to be optimally characterized. If the first energy arrival after  $l_1$  occurs at time  $l_1 + \tau_1$  with  $\tau_1 > x_1$ , the sensor transmits the update at  $l_1 + x_1$  as scheduled, whence the state becomes  $(0, 0)$ . Note that by energy causality,  $x_1$  cannot depend on  $\tau_1$ , and since it also does not depend on the past history before  $l_1$  (by Corollary 1), it is therefore a constant. On the other hand, if the first energy arrival occurs relatively early, i.e.,  $\tau_1 \leq x_1$ , the state becomes  $(2, \tau_1)$  at  $l_1 + \tau_1$ , and the sensor *reschedules* the update to be at  $l_1 + y_2(\tau_1)$  instead of  $l_1 + x_1$ . Note that it is not clear so far whether  $y_2(\cdot)$  and  $\bar{y}_2(\cdot, \cdot)$  are identical, since the former depends on only one random variable, as opposed to depending on two random variables in the latter; we optimally characterize both functions later on in the analysis. The above two cases are illustrated in Fig. 3.

In summary, the optimal update policy for the case  $B = 2$  in a given epoch is completely characterized by the constant  $x_1$ , and the functions  $y_1(\cdot)$ ,  $y_2(\cdot)$ , and  $\bar{y}_2(\cdot, \cdot)$ . Since these represent the possible inter-update delays, we conclude by Corollary 1 that they do not depend on each other. We denote

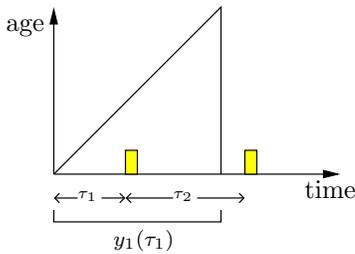


Fig. 4. First possible way to return to state  $(0, 0)$ .

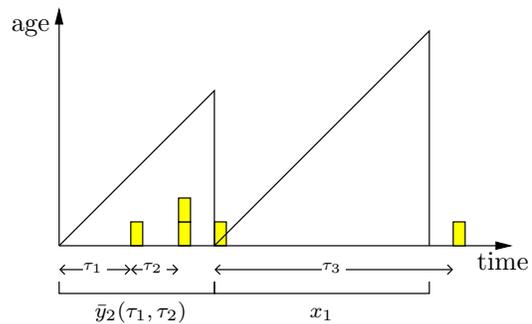


Fig. 5. Second possible way to return to state  $(0, 0)$ .

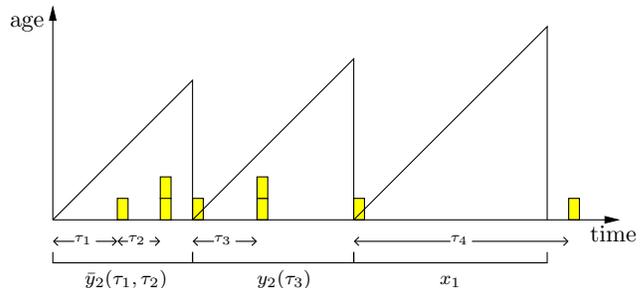


Fig. 6. Third possible way to return to state  $(0, 0)$ .

by  $R(x_1, y_1, y_2, \bar{y}_2)$  and  $L(x_1, y_1, y_2, \bar{y}_2)$  the area under the age curve in the epoch and its length, respectively, as a function of the policy  $(x_1, y_1, y_2, \bar{y}_2)$ . By Theorem 1 (and Corollary 1), one can use the strong law of large numbers of renewal processes [41] to reduce problem (5) to be an optimization over a single epoch as follows

$$\begin{aligned}
 & \min_{x_1, y_1, y_2, \bar{y}_2} \frac{\mathbb{E}[R(x_1, y_1, y_2, \bar{y}_2)]}{\mathbb{E}[L(x_1, y_1, y_2, \bar{y}_2)]} \\
 & \text{s.t.} \quad x_1 \geq 0 \\
 & \quad y_1(\tau) \geq \tau, \quad \forall \tau \\
 & \quad y_2(\tau) \geq \tau, \quad \forall \tau \\
 & \quad \bar{y}_2(\tau_1, \tau_2) \geq \tau_1 + \tau_2, \quad \forall \tau_1, \tau_2 \quad (6)
 \end{aligned}$$

where the expectation is on the energy arrival patterns in the epoch. Note that the constraints on the functions  $y_1$ ,  $y_2$ , and  $\bar{y}_2$ , represent the energy causality constraints. Next, in order to evaluate the expectations in the objective function, one needs to study the different patterns that can occur in a single epoch. We do so in the following subsection.

#### A. Renewal State Analysis

Consider the state  $(0, 0)$  as the renewal state<sup>1</sup>, and without loss of generality assume that we start at time 0. Let us now state the possible ways of returning to that state. Note that the sensor has to wait for at least one energy arrival to update since it starts with no energy at state  $(0, 0)$ .

- The first way to return to  $(0, 0)$  is to receive an energy arrival after  $\tau_1$  time units, and then update at  $y_1(\tau_1)$ . This could happen if and only if the following energy arrival, occurring at  $\tau_2$  time units after the first arrival, arrives after  $y_1(\tau_1) - \tau_1$ . See Fig. 4.
- The second way is to receive another energy arrival after the first one, before using the first energy unit to update. Then, submit the first update at  $\bar{y}_2(\tau_1, \tau_2)$ , which makes the state become  $(1, 0)$ , and then submit another update after  $x_1$  time units. This could happen if and only if the following energy arrival, occurring at  $\tau_3$  time units after the first update, is such that  $\tau_3 > x_1$ . See Fig. 5.
- The third way is exactly as the second way, but with  $\tau_3 \leq x_1$ , and hence the system goes to state  $(2, \tau_3)$  with

the third energy arrival. Then, the sensor updates after  $y_2(\tau_3)$  time units from the first update (as opposed to  $x_1$  in the second way), which makes the state become  $(1, 0)$ , and then finally submit a third update after  $x_1$  time units. As before, this could happen if and only if the following energy arrival, occurring at  $\tau_4$  time units after the second update, is such that  $\tau_4 > x_1$ . See Fig. 6.

- In general, the  $m$ th way,  $m \geq 3$ , begins exactly as in the second way by submitting the first update at  $\bar{y}_2(\tau_1, \tau_2)$ . Then, the second phase of the third way, namely, going from state  $(1, 0)$  to  $(2, \tau_3)$  to  $(1, 0)$  again, keeps repeating for  $m-2$  times. By the end of these repetitions the system will be in state  $(1, 0)$ . This is finally followed by the  $m$ th (and last) update after  $x_1$  time units. See Fig. 7.

Based on the above, one can write the area under the age curve,  $R$ , in a single epoch as in equation (7) at the top of the next page<sup>2</sup>. There,  $\mathbb{1}_A = 1$  if event  $A$  is true, and is 0 otherwise. Taking expectations and simplifying (mainly through using the fact that  $\tau_i$ 's are i.i.d.), we get

$$\begin{aligned}
 \mathbb{E}[R] &= \left( \frac{1}{2} x_1^2 + e^{x_1} \int_0^{x_1} \frac{1}{2} y_2(\tau)^2 e^{-\tau} d\tau \right) \\
 &\quad \times \left( 1 - \int_0^\infty e^{-y_1(\tau)} d\tau \right) + \int_0^\infty \frac{1}{2} y_1(\tau)^2 e^{-y_1(\tau)} d\tau \\
 &\quad + \int_{\tau_1=0}^\infty \int_{\tau_2=0}^{y_1(\tau_1)-\tau_1} \frac{1}{2} \bar{y}_2(\tau_1, \tau_2)^2 e^{-\tau_1} e^{-\tau_2} d\tau_1 d\tau_2 \quad (9)
 \end{aligned}$$

<sup>1</sup>From Theorem 1, we know that both states  $(0, 0)$  and  $(1, 0)$  are renewal states. While we choose to perform our analysis using state  $(0, 0)$ , we note that one can reach the same results if state  $(1, 0)$  is chosen instead.

<sup>2</sup>From now onwards, we drop the dependency on the tuple  $(x_1, y_1, y_2, \bar{y}_2)$  from  $R$  and  $L$  for convenience.

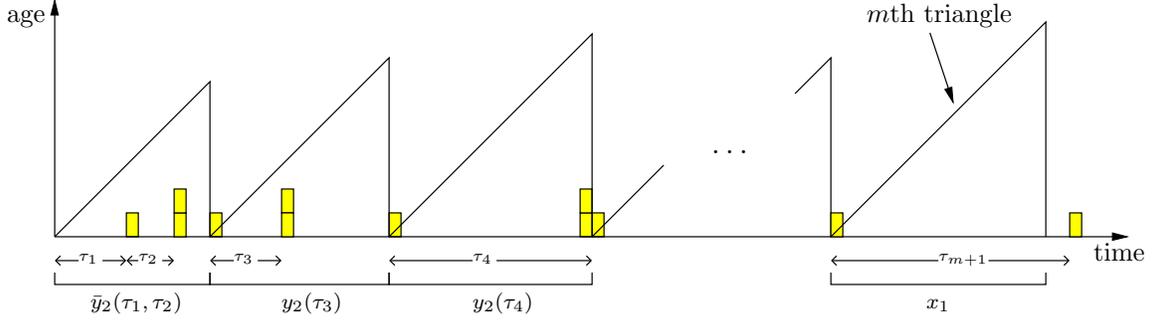


Fig. 7. General  $m$ th possible way to return to state  $(0, 0)$ ,  $m \geq 3$ .

$$\begin{aligned}
R &= \frac{1}{2}y_1(\tau_1)^2 \mathbb{1}_{\tau_2 > y_1(\tau_1) - \tau_1} + \left( \frac{1}{2}\bar{y}_2(\tau_1, \tau_2)^2 + \frac{1}{2}x_1^2 \right) \mathbb{1}_{\tau_2 \leq y_1(\tau_1) - \tau_1} \mathbb{1}_{\tau_3 > x_1} \\
&+ \left( \frac{1}{2}\bar{y}_2(\tau_1, \tau_2)^2 + \frac{1}{2}y_2(\tau_3)^2 + \frac{1}{2}x_1^2 \right) \mathbb{1}_{\tau_2 \leq y_1(\tau_1) - \tau_1} \mathbb{1}_{\tau_3 \leq x_1} \mathbb{1}_{\tau_4 > x_1} \\
&+ \left( \frac{1}{2}\bar{y}_2(\tau_1, \tau_2)^2 + \frac{1}{2}y_2(\tau_3)^2 + \frac{1}{2}y_2(\tau_4)^2 + \frac{1}{2}x_1^2 \right) \mathbb{1}_{\tau_2 \leq y_1(\tau_1) - \tau_1} \mathbb{1}_{\tau_3 \leq x_1} \mathbb{1}_{\tau_4 \leq x_1} \mathbb{1}_{\tau_5 > x_1} \\
&+ \dots
\end{aligned} \tag{7}$$

$$\begin{aligned}
L &= y_1(\tau_1) \mathbb{1}_{\tau_2 > y_1(\tau_1) - \tau_1} + (\bar{y}_2(\tau_1, \tau_2) + x_1) \mathbb{1}_{\tau_2 \leq y_1(\tau_1) - \tau_1} \mathbb{1}_{\tau_3 > x_1} \\
&+ (\bar{y}_2(\tau_1, \tau_2) + y_2(\tau_3) + x_1) \mathbb{1}_{\tau_2 \leq y_1(\tau_1) - \tau_1} \mathbb{1}_{\tau_3 \leq x_1} \mathbb{1}_{\tau_4 > x_1} \\
&+ (\bar{y}_2(\tau_1, \tau_2) + y_2(\tau_3) + y_2(\tau_4) + x_1) \mathbb{1}_{\tau_2 \leq y_1(\tau_1) - \tau_1} \mathbb{1}_{\tau_3 \leq x_1} \mathbb{1}_{\tau_4 \leq x_1} \mathbb{1}_{\tau_5 > x_1} \\
&+ \dots
\end{aligned} \tag{8}$$

Equation (9) is justified in Appendix B. Similarly, the epoch length  $L$  is given by (8) at the top of this page, and its expectation is given by

$$\begin{aligned}
\mathbb{E}[L] &= \left( x_1 + e^{x_1} \int_0^{x_1} y_2(\tau) e^{-\tau} d\tau \right) \\
&\times \left( 1 - \int_0^\infty e^{-y_1(\tau)} d\tau \right) + \int_0^\infty y_1(\tau) e^{-y_1(\tau)} d\tau \\
&+ \int_{\tau_1=0}^\infty \int_{\tau_2=0}^{y_1(\tau_1) - \tau_1} \bar{y}_2(\tau_1, \tau_2) e^{-\tau_1} e^{-\tau_2} d\tau_1 d\tau_2 \tag{10}
\end{aligned}$$

We use the above results to characterize the structure of the optimal policy for problem (6) in the next subsection.

### B. Optimal Solution for Problem (6): Threshold Policies

We define the following parameterized problem to characterize the optimal solution of problem (6)

$$\begin{aligned}
p_2(\lambda) &\triangleq \min_{x_1, y_1, y_2, \bar{y}_2} \mathbb{E}[R] - \lambda \mathbb{E}[L] \\
\text{s.t.} \quad &x_1 \geq 0 \\
&y_1(\tau) \geq \tau, \quad \forall \tau \\
&y_2(\tau) \geq \tau, \quad \forall \tau \\
&\bar{y}_2(\tau_1, \tau_2) \geq \tau_1 + \tau_2, \quad \forall \tau_1, \tau_2 \tag{11}
\end{aligned}$$

where the subscript 2 in  $p_2(\lambda)$  denotes the  $B = 2$  case that we consider here. This approach has also been used in [42]. We now have the following lemma.

**Lemma 1**  $p_2(\lambda)$  is decreasing in  $\lambda$ , and the optimal solution of problem (6) is given by  $\lambda^*$  that solves  $p_2(\lambda^*) = 0$ .

**Proof:** Let  $\lambda_1 > 0$ , and let the solution of problem (11) be given by  $(x_1^{(1)}, y_1^{(1)}, y_2^{(1)}, \bar{y}_2^{(1)})$  for  $\lambda = \lambda_1$ , with the corresponding average area under the age curve in the epoch and the average epoch length given by  $\mathbb{E}[R^{(1)}]$  and  $\mathbb{E}[L^{(1)}]$ , respectively. Now for some  $\lambda_2 > \lambda_1$ , one can write

$$\begin{aligned}
p_2(\lambda_1) &= \mathbb{E}[R^{(1)}] - \lambda_1 \mathbb{E}[L^{(1)}] \\
&> \mathbb{E}[R^{(1)}] - \lambda_2 \mathbb{E}[L^{(1)}] \\
&\geq p_2(\lambda_2). \tag{12}
\end{aligned}$$

where the last inequality follows since  $(x_1^{(1)}, y_1^{(1)}, y_2^{(1)}, \bar{y}_2^{(1)})$  is also feasible in problem (11) for  $\lambda = \lambda_2$ .

Next, note that both problems (11) and (6) have the same feasible set. In addition, if  $p_2(\lambda) = 0$ , then the objective function of (6) satisfies  $\mathbb{E}[R] / \mathbb{E}[L] = \lambda$ . Hence, the objective function of (6) is minimized by minimizing  $\lambda \geq 0$  such that  $p_2(\lambda) = 0$ . Finally, by the first part of lemma, there can only be one such  $\lambda$ , which we denote  $\lambda^*$ . ■

By Lemma 1, one can simply use a bisection method to find  $\lambda^*$  that solves  $p_2(\lambda^*) = 0$ . This  $\lambda^*$  certainly exists since  $p_2(0) > 0$  and  $\lim_{\lambda \rightarrow \infty} p_2(\lambda) = -\infty$ . We focus on problem (11) in the rest of this subsection, for which we introduce the

following Lagrangian [43]

$$\begin{aligned} \mathcal{L} = & \mathbb{E}[R] - \lambda \mathbb{E}[L] - \eta_1 x_1 - \int_0^\infty \gamma_1(\tau) (y_1(\tau) - \tau) d\tau \\ & - \int_0^\infty \gamma_2(\tau) (y_2(\tau) - \tau) d\tau \\ & - \int_0^\infty \int_0^\infty \bar{\gamma}_2(\tau_1, \tau_2) (\bar{y}_2(\tau_1, \tau_2) - \tau_1 - \tau_2) d\tau_1 d\tau_2 \end{aligned} \quad (13)$$

where  $\eta_1$ ,  $\gamma_1(\cdot)$ ,  $\gamma_2(\cdot)$ ,  $\bar{\gamma}_2(\cdot, \cdot)$  are Lagrange multipliers. Using (9) and (10), we take the (functional) derivative of the Lagrangian with respect to  $\bar{y}_2(t_1, t_2)$  and equate it to 0 to get

$$\bar{y}_2(t_1, t_2) = \lambda + \frac{\bar{\gamma}_2(t_1, t_2)}{e^{-(t_1+t_2)}} \quad (14)$$

Now if  $t_1 + t_2 < \lambda$ , then  $\bar{y}_2(t_1, t_2)$  has to be larger than  $t_1 + t_2$ , for if it were equal, the right hand side of the above equation would be larger than the left hand side. By complementary slackness [43], we conclude that in this case  $\bar{\gamma}_2(t_1, t_2) = 0$ , and hence  $\bar{y}_2(t_1, t_2) = \lambda$ . On the other hand, if  $t_1 + t_2 \geq \lambda$ , then  $\bar{y}_2(t_1, t_2)$  has to be equal to  $t_1 + t_2$ , for if it were larger, then by complementary slackness  $\bar{\gamma}_2(t_1, t_2) = 0$  and the right hand side of the above equation would be smaller than the left hand side. In conclusion, we have

$$\bar{y}_2(t_1, t_2) = \begin{cases} \lambda, & t_1 + t_2 < \lambda \\ t_1 + t_2, & t_1 + t_2 \geq \lambda \end{cases} \quad (15)$$

The above result says that starting from state  $(0, 0)$  the sensor has to wait at least for  $\lambda$  time units before submitting an update, provided that it received two consecutive energy units (without using the first one to send an update) in that epoch. If these two energy arrivals occur relatively early, i.e.,  $t_1 + t_2 < \lambda$ , then the sensor updates exactly after  $\lambda$  time units from the beginning of the epoch. Otherwise, if  $t_1 + t_2 \geq \lambda$ , then the sensor updates instantly after receiving the second energy unit. We coin this type of policies  $\lambda$ -threshold policy, where the sensor can only update if the age grows above a certain threshold  $\lambda$ . Such policies were first introduced in the solution of the case of  $B = 1$  energy unit in [38], and have also appeared in the random full battery recharges analysis in [39]. We also note from the result in (15) that  $\bar{y}_2(t_1, t_2)$  only depends on the age at the second energy arrival,  $t_1 + t_2$ .

Next, we take the derivative of the Lagrangian with respect to  $y_2(t)$  and equate to 0 to get

$$y_2(t) = \lambda + \frac{\gamma_2(t)}{q e^{x_1} e^{-t}} \quad (16)$$

where  $q \triangleq 1 - \int_0^\infty e^{-y_1(\tau)} d\tau$ . Note that  $q \in [0, 1]$  since  $y_1(\tau) \geq \tau$ . Following the same arguments as in the  $\bar{y}_2$  case, we get that

$$y_2(t) = \begin{cases} \lambda, & t < \lambda \\ t, & t \geq \lambda \end{cases} \quad (17)$$

That is,  $y_2$  is also a  $\lambda$ -threshold policy, and  $\bar{y}_2(t_1, t_2) = y_2(t_1 + t_2)$ . This settles the earlier question we posed at the

beginning of this section of whether receiving two energy arrivals starting from state  $(0, 0)$  would lead to a different policy than receiving one energy arrival starting from state  $(1, 0)$ ; the optimal policy when the sensor has a full battery is only a function of the age at the time of receiving the second energy unit in the battery. Next, we take the derivative of the Lagrangian with respect to  $x_1$  and equate to 0 to get

$$\begin{aligned} x_1 = & \lambda + e^{x_1} \int_0^{x_1} \left( \lambda y_2(\tau) - \frac{1}{2} y_2(\tau)^2 \right) e^{-\tau} d\tau \\ & + \lambda y_2(x_1^-) - \frac{1}{2} y_2(x_1^-)^2 + \frac{\eta_1}{q} \end{aligned} \quad (18)$$

We now make an assumption that  $x_1 > \lambda$ , and verify that assumption below. Based on that,  $y_2(x_1^-) = x_1$  from (17). One can also use (17) to evaluate the integral in the above equation in terms of  $\lambda$  and  $x_1$ . After some algebraic manipulations, we get that for  $x_1 > 0$ ,  $\eta_1 = 0$  by complementary slackness, and the following holds

$$x_1 = \log \left( \frac{1}{e^{-\lambda} - \frac{1}{2} \lambda^2} \right) \quad (19)$$

where  $\log$  is the natural logarithm. It is direct to see from (19) that  $x_1 > \lambda$  as assumed above.

Finally, we take the derivative of the Lagrangian with respect to  $y_1(t)$  and equate to 0 to get

$$\begin{aligned} y_1(t) = & \lambda + e^{x_1} \int_0^{x_1} \left( \lambda y_2(\tau) - \frac{1}{2} y_2(\tau)^2 \right) e^{-\tau} d\tau + \lambda x_1 - \frac{1}{2} x_1^2 \\ & + \frac{1}{2} y_1(t)^2 - \frac{1}{2} \bar{y}_2 \left( t, (y_1(t) - t)^- \right)^2 \\ & - \lambda y_1(t) + \lambda \bar{y}_2 \left( t, (y_1(t) - t)^- \right) + \frac{\gamma_1(t)}{e^{-y_1(t)}} \end{aligned} \quad (20)$$

We now make another assumption that  $y_1(t) > \lambda$ ,  $\forall t$ , and verify it below. Based on this assumption, we conclude by (15) that  $\bar{y}_2 \left( t, (y_1(t) - t)^- \right) = y_1(t)$ . We substitute this in (20), and use (18) to get

$$y_1(t) = x_1 + \frac{\gamma_1(t)}{e^{-y_1(t)}} \quad (21)$$

which verifies that  $y_1(t) > \lambda$ ,  $\forall t$ , since  $x_1 > \lambda$ . Similar to the arguments used in deriving (15) and (17), we conclude from (21) that  $y_1$  is an  $x_1$ -threshold policy given by

$$y_1(t) = \begin{cases} x_1, & t < x_1 \\ t, & t \geq x_1 \end{cases} \quad (22)$$

Similar to the discussion regarding the equivalence of  $\bar{y}_2$  and  $y_2$ , we conclude from (22) that starting from state  $(0, 0)$  and receiving one energy unit is equivalent to starting from state  $(1, 0)$  and receiving no energy units; in both cases, the sensor has the same threshold  $x_1$  after which it can update. Using (15), (17), (19), and (22) we get that

$$\begin{aligned} p_2(\lambda) = & \frac{1}{2} \lambda^2 + (\lambda + 1) e^{-\lambda} + \lambda \\ & - \left( e^{-\lambda} - \frac{1}{2} \lambda^2 + 1 \right) \log \left( \frac{1}{e^{-\lambda} - \frac{1}{2} \lambda^2} \right) \end{aligned} \quad (23)$$

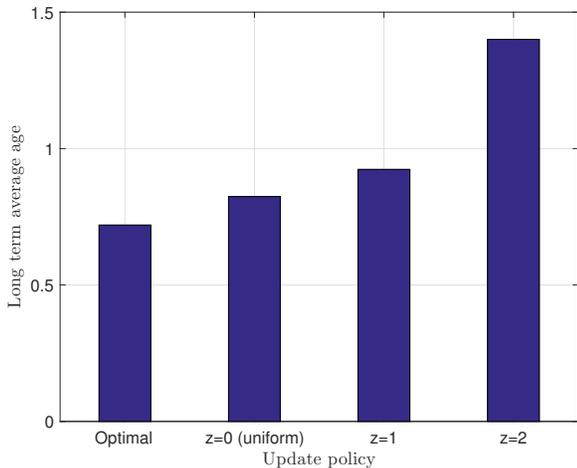


Fig. 8. Comparison of the optimal policy for  $B = 2$  to other policies: uniform updating, and energy-aware adaptive updating of [38].

It now remains to find  $\lambda^*$ . Towards that, we first note that we have an upper bound on  $\lambda^*$  given by 0.9012, the solution of the  $B = 1$  case derived in [38]. We also have a lower bound of 0.5, which is the optimal solution in the case of having an infinite battery, also derived in [38]. Using bisection, we find that the optimal solution at which  $p_2(\lambda^*) = 0$  is given by  $\lambda^* \approx 0.72$ , with the corresponding  $x_1^* \approx 1.48$ . Observe that the fact that  $x_1^*$  is larger than  $\lambda^*$  implies the intuitive behavior that the sensor is less eager to send an update if it has only one energy unit, compared to when it has a full battery of two energy units.

### C. Comparison to Other Policies

We now compare the optimal result derived above with other schemes and system models in the literature. We first compare it to the *energy-aware adaptive status update policy* introduced and analyzed in [38]. In there, the sensor schedules its next update based on the amount of energy in its battery; if the energy is less than  $B/2$ , it schedules the next update after  $1/(1 - \beta)$  time units, for some constant  $\beta < 1$ ; if the energy is larger than  $B/2$ , it schedules the next update after  $1/(1 + \beta)$  time units; and if the energy is exactly equal to  $B/2$ , it schedules the next update after 1 time unit. Then, if the sensor has no energy at its scheduled update time, it stays silent, and reschedules its following update accordingly after  $1/(1 - \beta)$  time units. We note that for  $\beta = 0$ , this energy-aware status update policy transforms into a *best effort uniform update policy*, which is the optimal solution for the infinite battery case [38]. We also note that the constant  $\beta$  is chosen in [38] such that the policy is asymptotically optimal in the battery size. Specifically, it is chosen equal to  $z \log B/B$  for some positive integer  $z$  that controls the policy's asymptotic behavior. We compare our optimal policy to the energy-aware policy above for  $z = 2$ ,  $z = 1$ , and  $z = 0$  (uniform update policy) in Fig. 8. We see that it outperforms all of them.

Finally, we compare the optimal policy to our recent results

on an altered system model of the same problem [39]. There, the battery is *fully* recharged randomly over time, i.e., energy arrives in chunks of  $B$  energy units, as opposed to the incremental unit recharges considered in this work. We consider two situations of this random battery recharges to compare with. The first is when the Poisson arrival process is of unit rate, and the second is when it is of rate  $1/2$ . The second case corresponds to an average recharge rate of  $B/2 = 1$  energy unit per unit time, as considered in this work. From [39], the optimal long term average age for the first situation is given by  $r_1^* = 0.59$ . While the analysis in [39] is done for a Poisson arrival process of unit rate, it can be directly extended to account for that of rate  $1/2$ ; this gives the optimal long term average age for the second situation by  $r_2^* = 1.18$ , which is double  $r_1^*$ , since the average recharge rate is reduced to half. We conclude from this that while it is clearly better to have the battery recharged by 2 energy units, as opposed to only 1, every one time unit on average ( $r_1^* < \lambda^*$ ), it is worse to be recharged by 2 energy units every 2 time units on average, as opposed to 1 energy unit per unit time ( $r_2^* > \lambda^*$ ), although the recharge rate is the same. The latter conclusion for the second situation is due to the fact that the system with 1 energy unit recharge per unit time considered in this work gives more flexibility to the sensor on when to update compared to the system with 2 energy units recharge every 2 time units. This flexibility allows the sensor to submit updates more uniformly over time, which achieves better age by convexity of the square function that governs the areas of the triangles constituting the total area under the age curve to be minimized.

## V. CONCLUSION AND FUTURE DIRECTIONS

We have characterized optimal online policies for energy harvesting sensors with  $B$ -sized batteries that minimize the long term average age of information, subject to energy causality constraints. We have considered a noiseless channel where a transmission update consumes one energy unit and arrives instantaneously at the receiver. Under a Poisson energy arrival process with unit rate, energy units arrive at the sensor's battery in an incremental fashion, i.e., one energy unit per arrival. We first have shown that the optimal status update policy has a renewal structure. Specifically, the times between the two consecutive events of submitting an update and having  $k$  energy units remaining in the battery afterwards,  $0 \leq k \leq B - 1$ , are i.i.d. Then, we have thoroughly studied the specific scenario of  $B = 2$  energy units and further shown that the optimal renewal policy has an *energy-dependent threshold* structure: the sensor submits an update only if the age of information surpasses a certain threshold which is a function of the energy available in its battery.

From the analysis of the  $B = 2$  case, it is amenable to show that threshold policies are also optimal for any  $B \geq 3$ . One main difficulty in showing that is the combinatorial nature of how the different  $B$  random variables that govern the energy arrivals in between inter-updates are related. Similar to the approaches in [15]–[20], it is therefore of interest to study near-optimal renewal-type policies that provably perform

within a constant gap from the optimal solution of problem (5) in future works.

## APPENDIX

### A. Proof of Theorem 1

We prove this by showing that any given status update policy that is uniformly bounded according to Definition 1 is outperformed by a renewal policy as defined in the theorem. Let us consider the  $i$ th epoch (time between two consecutive visits to state  $(k, 0)$ ); we introduce the following notation regarding the energy arrivals occurring in it. Let  $\tau_{1,i}$  denote the time until the first energy arrival after the epoch starts, and let there be  $j_1$  status updates after that energy arrival *before* a second energy arrival occurs. If  $j_1 \geq 1$ , then let  $\tau_{2,i}$  denote the time until the first energy arrival *after* the  $j_1$ th update. Otherwise, if  $j_1 = 0$ , then let  $\tau_{2,i}$  denote the inter-arrival time between the first and the second energy arrivals in the epoch. Similarly, let there be  $j_2$  status updates after the second energy arrival *before* a third energy arrival occurs. If  $j_2 \geq 1$ , then let  $\tau_{3,i}$  denote the time until the first energy arrival *after* the  $j_2$ th update. Otherwise, if  $j_2 = 0$ , then let  $\tau_{3,i}$  denote the inter-arrival time between the second and the third energy arrivals in the epoch. We continue defining  $\tau_{j,i}$ 's,  $j = 1, 2, \dots$ , until the epoch ends by retuning back to state  $(k, 0)$  again. Finally, in the event that the  $j$ th energy arrival in the epoch makes the battery full, then we wait until the first status update occurs after that event and denote by  $\tau_{j+1,i}$  the time until the first energy arrival *after* that update, i.e., we do not account for energy arrivals that cause battery overflows.

As noted before Theorem 1, there can possibly be an infinite number of updates before the system returns back to state  $(k, 0)$ , depending on the energy arrival pattern and the update time decisions. For a given status update policy, one can enumerate all such patterns. For instance, following the above notation, the first pattern could be when the system goes from state  $(k, 0)$  to state  $(k+1, \tau_{1,i})$  and then to state  $(k, 0)$  again; the second pattern could be when the system goes through the following sequence of states:  $(k, 0) - (k+1, \tau_{1,i}) - (k+2, \tau_{1,i} + \tau_{2,i}) - (k+1, 0) - (k, 0)$ ; and so on. Let the vector  $\tau_{m,i}$  contain all the  $\tau_{j,i}$ 's in the  $m$ th pattern. Note that this vector's length varies with the pattern. For instance, we have  $\tau_{1,i} = \tau_{1,i}$  and  $\tau_{2,i} = [\tau_{1,i}, \tau_{2,i}]$  for the above two pattern examples, respectively. For a given status update policy, one can also compute the probability of occurrence of the  $m$ th pattern in the  $i$ th epoch, denoted by  $p_{m,i}$ , with  $\sum_{m=1}^{\infty} p_{m,i} = 1$ . Let us also denote by  $R_{m,i}$  the area under the age curve in that epoch, given that it went through the  $m$ th pattern.

Next, for a fixed history  $\mathcal{H}_{i-1}$  and a pattern  $m$ , let us group all the status updating sample paths that have the same  $\tau_{m,i}$  and perform a statistical averaging over all of them to get the following average age in the  $i$ th epoch given that it went through the  $m$ th pattern

$$\hat{R}_{m,i}(\gamma_m, \mathcal{H}_{i-1}) \triangleq \mathbb{E}[R_{m,i} | \tau_{m,i} = \gamma_m, \mathcal{H}_{i-1}] \quad (24)$$

Now for a given time  $T$ , let  $N_T$  denote the number of epochs

that have already *started* by time  $T$ . Then, we have

$$\begin{aligned} & \mathbb{E}[R_{m,i} \cdot \mathbb{1}_{i \leq N_T}] \\ &= \mathbb{E}_{\mathcal{H}_{i-1}} \left[ \mathbb{E}_{\tau_{m,i}} \left[ \hat{R}_{m,i}(\gamma_m, \mathcal{H}_{i-1}) \right] \cdot \mathbb{1}_{i \leq N_T} \middle| \mathcal{H}_{i-1} \right] \end{aligned} \quad (25)$$

where equality follows since  $\mathbb{1}_{i \leq N_T}$  is independent of  $\tau_{m,i}$  given  $\mathcal{H}_{i-1}$ . Similarly, let  $x_{k,m,i}$  denote the length of the  $i$ th epoch under the  $m$ th pattern, and define its (conditional) average as

$$\hat{x}_{k,m,i}(\gamma_m, \mathcal{H}_{i-1}) \triangleq \mathbb{E}[x_{k,m,i} | \tau_{m,i} = \gamma_m, \mathcal{H}_{i-1}] \quad (26)$$

Finally, we denote by  $R_i$  and  $x_{k,i}$  the area under the age curve in the  $i$ th epoch and its length, respectively, irrespective of which pattern it went through.

Next, note that by (4), the following holds

$$\frac{1}{T} \sum_{i=1}^{\infty} R_i \mathbb{1}_{i \leq N_T-1} \leq \frac{r(T)}{T} \leq \frac{1}{T} \sum_{i=1}^{\infty} R_i \mathbb{1}_{i \leq N_T} \quad (31)$$

Following similar analysis as in [38, Appendix C-1], one can show that

$$\lim_{T \rightarrow \infty} \frac{\mathbb{E}[R_{N_T}]}{T} = 0 \quad (32)$$

for any uniformly bounded policy as in Definition 1. Hence, the expected values of the upper and lower bounds in (31) are equal as  $T \rightarrow \infty$ . Hence, in the sequel, we derive a lower bound on  $\frac{1}{T} \mathbb{E}[\sum_{i=1}^{\infty} R_i \mathbb{1}_{i \leq N_T}]$  and use the above note to conclude that it is also a lower bound on  $\frac{\mathbb{E}[r(T)]}{T}$  as  $T \rightarrow \infty$ . Towards that end, note that  $\mathbb{E}[\sum_{i=1}^{\infty} x_{k,i} \mathbb{1}_{i \leq N_T}] \geq T$ . Then, we have

$$\frac{1}{T} \mathbb{E} \left[ \sum_{i=1}^{\infty} R_i \mathbb{1}_{i \leq N_T} \right] \geq \frac{\mathbb{E}[\sum_{i=1}^{\infty} R_i \mathbb{1}_{i \leq N_T}]}{\mathbb{E}[\sum_{i=1}^{\infty} x_{k,i} \mathbb{1}_{i \leq N_T}]} \quad (33)$$

We now proceed by lower bounding the right hand side of the above equation through a series of equations at the top of the next page. In there, (27) follows from (25) and the monotone convergence theorem, together with the fact that  $\mathbb{E}[R_i] = \sum_{m=1}^{\infty} p_{m,i} \mathbb{E}[R_{m,i}]$ ;  $R^*(\mathcal{H}_{i-1})$  is the minimum value of  $\frac{\sum_{m=1}^{\infty} p_{m,i} \mathbb{E}_{\tau_{m,i}}[\hat{R}_{m,i}(\gamma_m, \mathcal{H}_{i-1})]}{\sum_{m=1}^{\infty} p_{m,i} \mathbb{E}_{\tau_{m,i}}[\hat{x}_{k,m,i}(\gamma_m, \mathcal{H}_{i-1})]}$ ; and  $R_{\min}$  is the minimum value of  $R^*(\mathcal{H}_{i-1})$  over all possible epochs and their corresponding histories, i.e., the minimum over all  $i$  and  $\mathcal{H}_{i-1}$ . This, together with the fact that  $\mathbb{E}[x_{k,i}] = \sum_{m=1}^{\infty} p_{m,i} \mathbb{E}[x_{k,m,i}]$ , gives the last inequality.

Observe that a policy achieving  $R^*(\mathcal{H}_{i-1})$  is a policy which is a function of the possible energy arrival patterns in the  $i$ th epoch  $\tau_{m,i}$ 's only, since the history  $\mathcal{H}_{i-1}$  is fixed. Since the energy arrival process is Poisson with rate 1, it follows that the random vector  $\tau_{m,i}$  consists of i.i.d. exponential random variables with parameter 1, and that  $\{\tau_{m,i}\}$  are also independent across epochs. Therefore, if we repeat the policy that achieves  $R_{\min}$  over all epochs, we get a renewal policy where the epoch lengths are also i.i.d., and  $\{l_i\}$  forms a renewal process. This completes the proof.

$$\frac{\mathbb{E} \left[ \sum_{i=1}^{\infty} R_i \mathbb{1}_{i \leq N_T} \right]}{\mathbb{E} \left[ \sum_{i=1}^{\infty} x_{k,i} \mathbb{1}_{i \leq N_T} \right]} = \frac{\sum_{i=1}^{\infty} \mathbb{E}_{\mathcal{H}_{i-1}} \left[ \sum_{m=1}^{\infty} p_{m,i} \mathbb{E}_{\tau_{m,i}} \left[ \hat{R}_{m,i}(\gamma_m, \mathcal{H}_{i-1}) \right] \cdot \mathbb{1}_{i \leq N_T} \middle| \mathcal{H}_{i-1} \right]}{\mathbb{E} \left[ \sum_{i=1}^{\infty} x_{k,i} \mathbb{1}_{i \leq N_T} \right]} \quad (27)$$

$$= \frac{\sum_{i=1}^{\infty} \mathbb{E}_{\mathcal{H}_{i-1}} \left[ \sum_{m=1}^{\infty} p_{m,i} \mathbb{E}_{\tau_{m,i}} \left[ \hat{x}_{k,m,i}(\gamma_m, \mathcal{H}_{i-1}) \right] \cdot \frac{\sum_{m=1}^{\infty} p_{m,i} \mathbb{E}_{\tau_{m,i}} \left[ \hat{R}_{m,i}(\gamma_m, \mathcal{H}_{i-1}) \right]}{\sum_{m=1}^{\infty} p_{m,i} \mathbb{E}_{\tau_{m,i}} \left[ \hat{x}_{k,m,i}(\gamma_m, \mathcal{H}_{i-1}) \right]} \cdot \mathbb{1}_{i \leq N_T} \middle| \mathcal{H}_{i-1} \right]}{\mathbb{E} \left[ \sum_{i=1}^{\infty} x_{k,i} \mathbb{1}_{i \leq N_T} \right]} \quad (28)$$

$$\geq \frac{\sum_{i=1}^{\infty} \mathbb{E}_{\mathcal{H}_{i-1}} \left[ \sum_{m=1}^{\infty} p_{m,i} \mathbb{E}_{\tau_i} \left[ \hat{x}_{k,m,i}(\gamma, \mathcal{H}_{i-1}) \right] \cdot R^*(\mathcal{H}_{i-1}) \cdot \mathbb{1}_{i \leq N_T} \middle| \mathcal{H}_{i-1} \right]}{\mathbb{E} \left[ \sum_{i=1}^{\infty} x_{k,i} \mathbb{1}_{i \leq N_T} \right]} \quad (29)$$

$$\geq R_{\min} \quad (30)$$

## B. Justification of (9)

First, we have

$$\begin{aligned} & \mathbb{E} \left[ \frac{1}{2} y_1(\tau_1)^2 \mathbb{1}_{\tau_2 > y_1(\tau_1) - \tau_1} \right] \\ &= \int_{\tau_1=0}^{\infty} \int_{\tau_2=y_1(\tau_1) - \tau_1}^{\infty} \frac{1}{2} y_1(\tau_1)^2 e^{-\tau_1} e^{-\tau_2} d\tau_2 d\tau_1 \\ &= \int_{\tau_1=0}^{\infty} \frac{1}{2} y_1(\tau_1)^2 e^{-\tau_1} e^{-(y_1(\tau_1) - \tau_1)} d\tau_1 \\ &= \int_{\tau_1=0}^{\infty} \frac{1}{2} y_1(\tau_1)^2 e^{-y_1(\tau_1)} d\tau_1 \end{aligned} \quad (34)$$

Next, let us define

$$\alpha \triangleq \mathbb{E} \left[ \left( \frac{1}{2} \bar{y}_2(\tau_1, \tau_2)^2 + \frac{1}{2} x_1^2 \right) \mathbb{1}_{\tau_2 \leq y_1(\tau_1) - \tau_1} \right] \quad (35)$$

Since  $\tau_m$  is independent of  $\tau_1$  and  $\tau_2$  for  $m \geq 3$ , and they are all i.i.d., we have that the term  $\alpha$  gets multiplied by  $\sum_{i=1}^{\infty} (1 - e^{-x_1})^{i-1} e^{-x_1} = 1$  when we compute  $\mathbb{E}[R]$ . Note that

$$\begin{aligned} \mathbb{E} \left[ \mathbb{1}_{\tau_2 \leq y_1(\tau_1) - \tau_1} \right] &= 1 - \mathbb{P}[\tau_2 > y_1(\tau_1) - \tau_1] \\ &= 1 - \int_{\tau_1=0}^{\infty} \int_{\tau_2=y_1(\tau_1) - \tau_1}^{\infty} e^{-\tau_1} e^{-\tau_2} d\tau_2 d\tau_1 \\ &= 1 - \int_{\tau_1=0}^{\infty} e^{-\tau_1} e^{-(y_1(\tau_1) - \tau_1)} d\tau_1 \\ &= 1 - \int_{\tau_1=0}^{\infty} e^{-y_1(\tau_1)} d\tau_1 \end{aligned} \quad (36)$$

and hence the term  $\alpha$  can be expanded to

$$\begin{aligned} \alpha &= \frac{1}{2} x_1^2 \left( 1 - \int_{\tau_1=0}^{\infty} e^{-y_1(\tau_1)} d\tau_1 \right) \\ &+ \int_{\tau_1=0}^{\infty} \int_{\tau_2=0}^{y_1(\tau_1) - \tau_1} \frac{1}{2} \bar{y}_2(\tau_1, \tau_2)^2 e^{-\tau_1} e^{-\tau_2} d\tau_1 d\tau_2 \end{aligned} \quad (37)$$

Next, let us define

$$\beta_m \triangleq \int_0^{x_1} \frac{1}{2} y_2(\tau_m)^2 e^{-\tau_m} d\tau_m, \quad m \geq 3 \quad (38)$$

Now observe that, again by the fact that  $\tau_i$ 's are i.i.d., the terms  $\beta_m$ 's appear as follows when we take  $\mathbb{E}[R]$

$$\begin{aligned} & \beta_3 \mathbb{E} \left[ \mathbb{1}_{\tau_2 \leq y_1(\tau_1) - \tau_1} \right] + \sum_{m=4}^{\infty} \beta_m \mathbb{E} \left[ \mathbb{1}_{\tau_2 \leq y_1(\tau_1) - \tau_1} \right] \prod_{i=3}^m \mathbb{E} \left[ \mathbb{1}_{\tau_i \leq x_1} \right] \\ &= \beta_3 \mathbb{E} \left[ \mathbb{1}_{\tau_2 \leq y_1(\tau_1) - \tau_1} \right] \sum_{i=0}^{\infty} (1 - e^{-x_1})^i \\ &= e^{x_1} \beta_3 \left( 1 - \int_{\tau_1=0}^{\infty} e^{-y_1(\tau_1)} d\tau_1 \right) \end{aligned} \quad (39)$$

where the second equality follows since  $\beta_m$  is the same for all  $m$ . Equations (34), (37), and (39) yield  $\mathbb{E}[R]$  in (9).

## REFERENCES

- [1] J. Yang and S. Ulukus. Optimal packet scheduling in an energy harvesting communication system. *IEEE Trans. Commun.*, 60(1):220–230, January 2012.
- [2] K. Tutuncuoglu and A. Yener. Optimum transmission policies for battery limited energy harvesting nodes. *IEEE Trans. Wireless Commun.*, 11(3):1180–1189, March 2012.
- [3] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener. Transmission with energy harvesting nodes in fading wireless channels: Optimal policies. *IEEE JSAC*, 29(8):1732–1743, September 2011.
- [4] C. K. Ho and R. Zhang. Optimal energy allocation for wireless communications with energy harvesting constraints. *IEEE Trans. Signal Process.*, 60(9):4808–4818, September 2012.
- [5] J. Yang, O. Ozel, and S. Ulukus. Broadcasting with an energy harvesting rechargeable transmitter. *IEEE Trans. Wireless Commun.*, 11(2):571–583, February 2012.
- [6] O. Ozel, J. Yang, and S. Ulukus. Optimal broadcast scheduling for an energy harvesting rechargeable transmitter with a finite capacity battery. *IEEE Trans. Wireless Commun.*, 11(6):2193–2203, June 2012.
- [7] M. A. Antepi, E. Uysal-Biyikoglu, and H. Erkal. Optimal packet scheduling on an energy harvesting broadcast link. *IEEE JSAC*, 29(8):1721–1731, September 2011.
- [8] J. Yang and S. Ulukus. Optimal packet scheduling in a multiple access channel with energy harvesting transmitters. *Journal of Commun. Networks*, 14(2):140–150, April 2012.
- [9] K. Tutuncuoglu and A. Yener. Sum-rate optimal power policies for energy harvesting transmitters in an interference channel. *Journal Commun. Networks*, 14(2):151–161, April 2012.
- [10] C. Huang, R. Zhang, and S. Cui. Throughput maximization for the Gaussian relay channel with energy harvesting constraints. *IEEE JSAC*, 31(8):1469–1479, August 2013.
- [11] D. Gunduz and B. Devillers. Two-hop communication with energy harvesting. In *Proc. IEEE CAMSAP*, December 2011.
- [12] B. Gurakan and S. Ulukus. Cooperative diamond channel with energy harvesting nodes. *IEEE JSAC*, 34(5):1604–1617, May 2016.
- [13] B. Varan and A. Yener. Delay constrained energy harvesting networks with limited energy and data storage. *IEEE JSAC*, 34(5):1550–1564, May 2016.

- [14] A. Arafa, A. Baknina, and S. Ulukus. Energy harvesting two-way channels with decoding and processing costs. *IEEE Trans. Green Commun. and Networking*, 1(1):3–16, March 2017.
- [15] D. Shaviv and A. Ozgur. Universally near optimal online power control for energy harvesting nodes. *IEEE JSAC*, 34(12):3620–3631, December 2016.
- [16] H. A. Inan and A. Ozgur. Online power control for the energy harvesting multiple access channel. In *Proc. WiOpt*, May 2016.
- [17] A. Baknina and S. Ulukus. Energy harvesting multiple access channels: Optimal and near-optimal online policies. *IEEE Trans. Commun.* To appear.
- [18] A. Baknina and S. Ulukus. Optimal and near-optimal online strategies for energy harvesting broadcast channels. *IEEE JSAC*, 34(12):3696–3708, December 2016.
- [19] A. Baknina and S. Ulukus. Online scheduling for energy harvesting channels with processing costs. *IEEE Trans. Green Commun. and Networking*, 1(3):281–293, September 2017.
- [20] A. Arafa, A. Baknina, and S. Ulukus. Energy harvesting networks with general utility functions: Near optimal online policies. In *Proc. IEEE ISIT*, June 2017.
- [21] S. Kaul, R. Yates, and M. Gruteser. Real-time status: How often should one update? In *Proc. IEEE Infocom*, March 2012.
- [22] R. Yates and S. Kaul. Real-time status updating: Multiple sources. In *Proc. IEEE ISIT*, July 2012.
- [23] C. Kam, S. Kompella, and A. Ephremides. Age of information under random updates. In *Proc. IEEE ISIT*, July 2013.
- [24] M. Costa, M. Codreanu, and A. Ephremides. On the age of information in status update systems with packet management. *IEEE Trans. Inf. Theory*, 62(4):1897–1910, April 2016.
- [25] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis. Age and value of information: Non-linear age case. In *Proc. IEEE ISIT*, June 2017.
- [26] 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*, 63(11):7492–7508, November 2017.
- [27] A. M. Bedewy, Y. Sun, and N. B. Shroff. Age-optimal information updates in multihop networks. In *Proc. IEEE ISIT*, June, 2017.
- [28] Y. Hsu, E. Modiano, and L. Duan. Age of information: Design and analysis of optimal scheduling algorithms. In *Proc. IEEE ISIT*, June 2017.
- [29] J. Zhong, E. Soljanin, and R. D. Yates. Status updates through multicast networks. In *Proc. Allerton*, October 2017.
- [30] E. Najm and E. Telatar. Status updates in a multi-stream M/G/1/1 preemptive queue. Available Online: arXiv1801.04068.
- [31] R. Yates, E. Najm, E. Soljanin, and J. Zhong. Timely updates over an erasure channel. In *Proc. IEEE ISIT*, June 2017.
- [32] R. D. Yates, P. Ciblat, A. Yener, and M. A. Wigger. Age-optimal constrained cache updating. In *Proc. IEEE ISIT*, June 2017.
- [33] R. D. Yates. Lazy is timely: Status updates by an energy harvesting source. In *Proc. IEEE ISIT*, June 2015.
- [34] B. T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu. Age of information under energy replenishment constraints. In *Proc. ITA*, February 2015.
- [35] A. Arafa and S. Ulukus. Age-minimal transmission in energy harvesting two-hop networks. In *Proc. IEEE Globecom*, December 2017.
- [36] A. Arafa and S. Ulukus. Age minimization in energy harvesting communications: Energy-controlled delays. In *Proc. Asilomar*, October 2017.
- [37] B. 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.
- [38] X. Wu, J. Yang, and J. Wu. Optimal status update for age of information minimization with an energy harvesting source. *IEEE Trans. Green Commun. and Networking*. To appear.
- [39] A. Arafa, J. Yang, and S. Ulukus. Age-minimal online policies for energy harvesting sensors with random battery recharges. In *Proc. IEEE ICC*, May 2018.
- [40] A. Baknina, O. Ozel, J. Yang, S. Ulukus, and A. Yener. Sending information through status updates. Available Online: arXiv:1801.04907.
- [41] S. M. Ross. *Stochastic Processes*. Wiley, 1996.
- [42] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu. Remote estimation of the wiener process over a channel with random delay. In *Proc. IEEE ISIT*, June 2017. Longer version available: arXiv:1701.06734.
- [43] S. P. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.