Minimizing Age of Information in an Energy-Harvesting Scheduler With Rateless Codes

Subhankar Banerjee Sennur Ulukus
Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742
sbanerje@umd.edu *ulukus@umd.edu*

***Abstract*—We consider a status update system consisting of a source, a sampler, a monitor, and an energy harvesting battery at the sampler. The capacity of the battery is 1 unit of energy and the energy generation time follows an independent and identically distributed (i.i.d.) Bernoulli process. The sampler aims to sample fresh update packets from the source and transmit them to the monitor. The communication channel between the sampler and the monitor is erroneous, with the probability of successful transmission being q. Each update packet consists of 2 sub-packets. At time slot t, the sampler can transmit only one sub-packet. The transmitter employs a rateless coding scheme, for which the monitor successfully decodes an update packet if and only if it receives both of the sub-packets. Transmission of a sub-packet requires 1 unit of energy. When there is no update packet available to transmit, the sampler has to decide whether to sample a fresh update packet or stay idle. When the monitor has received the first sub-packet, the sampler has to decide whether to sample a fresh packet and transmit the first sub-packet of the fresher update packet or transmit the second sub-packet of the staler update packet. We formulate this problem as a Markov decision process (MDP), and identify several structures of an optimal policy. We propose an optimal policy based on the studied structure of an optimal policy. Finally, we propose a low-complexity heuristic policy, by leveraging some of the structures of an optimal policy, and compare its performance with the actual optimal policy obtianed by the MDP.**

I. Introduction

We consider a communication system that consists of a source, a sampler, a monitor, and an energy harvesting battery at the sampler. The source governs a stochastic process that the monitor aims to track. The monitor aims to receive fresh update packets from the source so that it can track the source efficiently. We quantify the freshness of the system with the well-known metric of age of information [1]. We consider that the communication channel is erroneous. There are examples in the literature where the transmitter employs standard automatic repeat request (ARQ), hybrid ARQ (HARQ) protocols, or employs different kinds of coding schemes such as rateless codes, maximum distance separable (MDS) codes, to mitigate the effects of the erasure channel, see e.g., [2]–[13]. Also, in the age of information literature, there are several works that consider an energy harvesting transmitter, see e.g., [13]–[18].

In this work, we consider an age minimization problem, where an energy harvesting sampler employs a rateless code to transmit update packets. We assume that the energy arrival to the battery follows an i.i.d. Bernoulli process, where the size of the battery is one unit. We assume that each update packet consists of ℓ sub-packets. To combat the channel errors, the transmitter employs a rateless code. The monitor can only decode a packet if it receives ℓ coded sub-packets. [13] studies the average age performance of a rateless code for an energy harvesting transmitter under different transmission schemes. However, [13] assumes that when the transmission of a packet starts it cannot stop the transmission in between and has to finish the transmission of all ℓ sub-packets. However, that may not be the optimal thing to do. As we want to minimize the average age of the system, the transmitter may want to drop a packet if it takes a long amount of time to transmit all ℓ sub-packets successfully. In this paper, we consider that the transmitter can discard a stale packet and sample a new packet for the transmission. [7] also optimizes the age of information for rateless codes and it considers a similar transmitter. However, [7] does not consider an energy harvesting battery in its model. With the introduction of an energy harvesting battery, the problem becomes more challenging, as it imposes a constraint on how often we can discard a packet for which the receiver already received one or more sub-packets.

We formulate this problem with an MDP formulation. We focus on the case where $\ell = 2$, i.e., that each update (full packet) can be decoded at the monitor, if two sub-packets are received successfully. We show that there exists a stationary policy that is optimal for the considered problem. We identify several structures of an optimal policy. We show that for an optimal policy, as the age of the monitor increases, the interval between discarding a packet for which the receiver has one sub-packet also increases. Under some assumptions, the optimal policy waits till the age of the monitor reaches a threshold, after that, whenever the transmitter has a unit of energy, it either discards a packet and samples a new packet or keeps transmitting a stale packet for which the receiver has a sub-packet. We use these structures to propose an optimal policy. We also propose a low-complexity heuristic policy by employing some of the structures of an optimal policy. Due to space restrictions here, we omit proofs for some of the results, which we will provide in the longer journal version.

II. System Model and Problem Formulation

We consider a slotted time communication system that comprises a source, an energy harvesting sampler/scheduler, and a monitor. We denote the probability of successful transmission

ISBN 978-3-903176-65-2 © 2024 IFIP

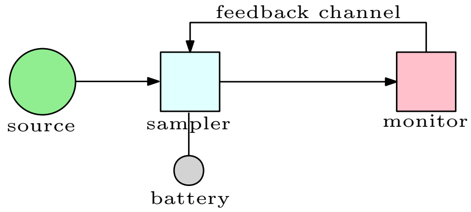

Fig. 1. The sampler can sample an update packet from the source any time it wants, i.e., the generate at-will model. The sampler is equipped with an energy-harvesting battery. The sampler can spend a unit of stored energy in the battery for transmission of a fresh update packet to the monitor.

of a packet with $q > 0$. In this work, we assume $\ell = 2$, a general ℓ can be an interesting future work. The sampler is equipped with an energy harvesting battery. In every time slot, the battery generates a unit of energy with probability $p > 0$. We assume that the capacity of the battery is one unit of energy. At a given time slot t, if the battery has one unit of energy, then the sampler can either transmit a sub-packet of a previously sampled packet, or can drop a previously sampled packet from the service, sample a new packet and put one coded sub-packet into the service, or can stay idle. For every transmission of a sub-packet, the sampler takes one unit of energy from the battery. Staying idle does not cost any energy for the sampler. We denote a valid policy of the sampler with π. We denote the set of all valid causal policies with Π. After a successful reception of an update packet, the user's age drops to the packet's generation time. We denote the instantaneous age of the system at time t, corresponding to a policy π, with $v^{\pi}(t)$. The average age of the system corresponding to a policy π with an initial age v_0, is

$$\Delta^{\pi}(v_0) = \limsup_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[v^{\pi}(t)|v^{\pi}(1) = v_0]. \qquad (1)$$

Thus, we are interested in the following optimization problem,

$$\inf_{\pi\in\Pi} \Delta^{\pi}(v_0). \qquad (2)$$

The feedback channel is instantaneous and error-free.

III. Optimal Policy

First, we re-formulate the problem in (1) as a Markov decision process (MDP), and then we optimally solve that MDP. Now, we state important components of this MDP.

State: A state of this system is defined with a vector of size four, $s = (v_1, v_2, c, d)$. We denote the state of the system at time t, with $s(t)$. The variable v_1 corresponds to the age of the monitor, and v_2 corresponds to the age of the recently sampled packet in the system. If there is a successful transmission of an update packet, i.e., if the monitor receives both sub-packets of a packet, then we assume that the sampler has no available packet and we denote this state with v_2 taking ∞. The variable c indicates the number of sub-packets of the available packet successfully received by the monitor. Thus, c can take only two values 0 and 1. The variable d indicates if the battery has an energy or not with $d = 1$ or $d = 0$, respectively. We denote the collection of all valid states with state space $\mathcal{S}$. For all states $s \in \mathcal{S}$, the age of a packet at the sampler is less than or equal to the age of the monitor, i.e., $v_2 \le v_1$. Thus,

$$\begin{aligned}\mathcal{S} =& \big\{(v_1, v_2, c, d) \mid (v_1, v_2) \in \mathbb{N}^2, v_2 \le v_1, (c, d) \in \{0, 1\}^2\big\} \\ &\cup \big\{(v_1, \infty, 0, d) \mid v_1 \in \mathbb{N}, d \in \{0, 1\}\big\}. \qquad (3)\end{aligned}$$

For simplicity of the presentation, we define the following states which we will refer to in several results in this paper:

$$\begin{aligned}
&\bar{s}_1 = (v_1, v_2, 1, 1),\ \bar{s}_2 = (v_1 + 1, v_2 + 1, 1, 1),\\
&\bar{s}_3 = (v_1 + 1, v_2 + 1, 1, 0),\ \bar{s}_4 = (v_2 + 1, \infty, 0, 1),\\
&\bar{s}_5 = (v_2 + 1, \infty, 0, 0),\ \bar{s}_6 = (v_1 + 1, 1, 0, 0),\\
&\bar{s}_7 = (v_1 + 1, 1, 0, 1),\ \bar{s}_8 = (v_1 + 1, 1, 1, 1),\\
&\bar{s}_9 = (v_1 + 1, 1, 1, 0),\ \tilde{s}_1 = (v_1, v_2, 1, 0),\ \tilde{s}_2 = (v_1, v_2, 0, 1),\\
&\tilde{s}_3 = (v_1 + 1, v_2 + 1, 0, 0),\ \tilde{s}_4 = (v_1 + 1, v_2 + 1, 0, 1),\\
&\tilde{s}_5 = (v_1, v_2, 0, 0),\ \tilde{s}_6 = (v_1, \infty, 0, 1),\ \tilde{s}_7 = (v_1, \infty, 0, 0),\\
&\tilde{s}_8 = (v_1 + 1, \infty, 0, 1),\ \tilde{s}_9 = (v_1 + 1, \infty, 0, 0),\\
&\hat{s}_1 = (v_1, 1, 1, 1),\ \hat{s}_2 = (v_1, 1, 1, 0),\ \hat{s}_3 = (v_1, 1, 0, 1),\\
&\hat{s}_4 = (v_1, 1, 0, 0),\ \hat{s}_5 = (v_1 + 1, 2, 1, 1),\ \hat{s}_6 = (v_1 + 1, 2, 1, 0),\\
&\hat{s}_7 = (v_1 + 1, 2, 0, 1),\ \hat{s}_8 = (v_1 + 1, 2, 0, 0),\\
&\hat{s}_9 = (v_1 + 2, 2, 1, 1),\ \check{s}_1 = (v_1 + 2, 2, 1, 0),\\
&\check{s}_2 = (v_1 + 2, 2, 0, 1),\ \check{s}_3 = (v_1 + 2, 2, 0, 0). \qquad (4)
\end{aligned}$$

Action space: We denote action of the sampler at time t with $a(t)$, which can take values from set $\mathcal{A} = \{0, 1, 2\}$. Action $a(t) = 0$, implies that the sampler stays idle at time t. Similarly, action $a(t) = 1$ implies that at time t, the sampler drops the packet in the service and samples a new packet. Finally, action $a(t) = 2$ implies that at time t, the sampler keeps transmitting sub-packets of the currently available packet at the sampler.

Cost: We define cost corresponding to a cost-action pair (s, a), $s \in \mathcal{S}$, $a \in \mathcal{A}$, with $c(s, a) = v_1$.

Transition probability: The probability with which the system evolves from a state s to a state s', under an action a, with $P_a(s, s')$. Next, we explicitly state transition probabilities for all possible state pairs under all possible actions:

$$\begin{aligned}
&P_0(\bar{s}_1, \bar{s}_2) = 1,\ P_0(\tilde{s}_1, \bar{s}_2) = p,\ P_0(\tilde{s}_1, \bar{s}_3) = (1 - p),\\
&P_0(\tilde{s}_2, \tilde{s}_4) = 1,\ P_0(\tilde{s}_5, \tilde{s}_4) = p,\ P_0(\tilde{s}_5, \tilde{s}_3) = (1 - p),\\
&P_0(\tilde{s}_6, \tilde{s}_8) = 1,\ P_0(\tilde{s}_7, \tilde{s}_8) = p,\ P_0(\tilde{s}_7, \tilde{s}_9) = (1 - p),\\
&P_1(\bar{s}_1, \bar{s}_8) = pq,\ P_1(\bar{s}_1, \bar{s}_9) = (1 - p)q,\ P_1(\bar{s}_1, \bar{s}_7) = p(1 - q),\\
&P_1(\bar{s}_1, \bar{s}_6) = (1 - p)(1 - q),\ P_1(\tilde{s}_2, \bar{s}_8) = pq,\\
&P_1(\tilde{s}_2, \bar{s}_9) = (1 - p)q,\ P_1(\tilde{s}_2, \bar{s}_7) = p(1 - q),\\
&P_1(\tilde{s}_2, \bar{s}_6) = (1 - p)(1 - q),\ P_1(\tilde{s}_6, \bar{s}_8) = pq,\\
&P_1(\tilde{s}_6, \bar{s}_9) = (1 - p)q, P_1(\tilde{s}_6, \bar{s}_7) = p(1 - q),\\
&P_1(\tilde{s}_6, \bar{s}_6) = (1 - p)(1 - q),\ P_2(\bar{s}_1, \bar{s}_2) = (1 - p)(1 - q),\\
&P_2(\bar{s}_1, \bar{s}_3) = pq,\ P_2(\bar{s}_1, \bar{s}_5) = (1 - p)q,\\
&P_2(\bar{s}_1, \bar{s}_2) = p(1 - q),\ P_2(\tilde{s}_2, \tilde{s}_3) = (1 - p)(1 - q),\\
&P_2(\tilde{s}_2, \tilde{s}_4) = p(1 - q),\ P_2(\tilde{s}_2, \bar{s}_3) = q(1 - p),
\end{aligned}$$

$$P_2(\tilde{s}_2, \bar{s}_2) = pq. \quad (5)$$

We denote the MDP with the above mentioned state space, action space, transition probability, and cost, with Δ.

Now, we reformulate the average age of the system, under a policy π and an initial state $s_0 \in \mathcal{S}$ as,

$$\Delta^\pi(s_0) = \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_\pi\big[C(s(t), a(t)) \mid s(1) = s_0\big]. \quad (6)$$

For a given initial state s_0, we study in the following problem,

$$\inf_{\pi\in\Pi} \Delta^\pi(s_0). \quad (7)$$

Note that, there exist some states for which if we do not choose those states as the initial state, then under any policy we would never visit those states. For example, consider state $(1, \infty, 0, 1)$, this particular state can never be visited by any policy, unless we choose this state to be the initial state for the system. In this work, we consider that state space $\mathcal{S}$ does not contain any such state, i.e., we assume that the system's initial state cannot be any such state.

Now, for the completeness of this work, we state some well-known results from the MDP literature [19], [20]. Let us assume that $0 < \alpha < 1$ and $s \in \mathcal{S}$, we denote the discounted cost corresponding to a policy π with $V_\alpha^\pi(s)$ where

$$V_\alpha^\pi(s) = \mathbb{E}_\pi\left[\sum_{t=1}^{\infty} \alpha^t C(s(t), a(t))\right]. \quad (8)$$

We denote the optimal value function with $V_\alpha(s)$, where $V_\alpha(s) = \inf_\pi V_\alpha^\pi(s)$. We define $V_\alpha(s; a)$, $a \in \{0, 1, 2\}$ as,

$$V_\alpha(s; a) = C(s, a) + \alpha \sum_{s'\in\mathcal{S}} P_a(s, s') V_\alpha(s'). \quad (9)$$

From [21], we know that if for all $s \in \mathcal{S}$, $V_\alpha(s)$ is finite then, $V_\alpha(s) = \min_{a\in\{0,1,2\}} V_\alpha(s; a)$. Now, let us consider the following iteration with $V_{\alpha,0}(s) = 0$, $s \in \mathcal{S}$ and $n \geq 1$,

$$V_{\alpha,n}(s) = \min_{a\in\{0,1,2\}} \left\{C(s, a) + \alpha \sum_{s'\in\mathcal{S}} P_a(s, s') V_{\alpha,n-1}(s')\right\}. \quad (10)$$

From [21], we say that,

$$\lim_{n\to\infty} V_{\alpha,n}(s) = V_\alpha(s). \quad (11)$$

We call a policy π stationary, if action taken by the sampler at time t depends only on state s and is independent of the time index t. In the next theorem, we study the existence of a stationary policy that is optimal for the problem in (7).

Theorem 1. *There exists a stationary policy π which is an optimal policy for the optimization problem,*

$$\inf_{\pi\in\Pi} \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_\pi\big[C(s(t), a(t)) \mid s(1) = s_0\big]. \quad (12)$$

From Theorem 1 and [21], we know that for any arbitrary initial state s_0, the optimization problem in (7) provides the same average age. Thus, without the loss of generality, we assume that the initial state of the system is $(2, \infty, 0, 1)$. For the simplicity of presentation, in the rest of this paper, we omit the initial state notation from all expressions.

Observation 1. *Consider any two arbitrary states, $s = (v_1, v_2, c, 1)$ and $s' = (v_1, v_2', c', 1)$, then $V_{\alpha,n}(s; 1) = V_{\alpha,n}(s'; 1)$ and $V_\alpha(s; 1) = V_\alpha(s'; 1)$, where v_2 or/and v_2' can be ∞ as well.*

In the next lemma, we study some monotonicity properties of the discounted value function.

Lemma 1. *Consider a state $s = (v_1, v_2, c, d)$, $c, d \in \{0, 1\}$, then $V_\alpha(s)$ is an increasing function of v_1 as well as of v_2. Similarly, for a state $s' = (v_1, \infty, c, d)$, $c, d \in \{0, 1\}$, $V_\alpha(s')$ is an increasing function of v_1.*

In the next lemma, we continue to study some further monotonicty properties of the discounted value function.

Lemma 2. *Let us consider a state $s = (v_1, v_1, c, d)$, $c, d \in \{0, 1\}$, then*

$$V_\alpha(s) = V_\alpha((v_1, \infty, 0, d)). \quad (13)$$

Moreover,

$$V_\alpha(\tilde{s}_2) \geq V_\alpha(\bar{s}_1), \quad (14)$$
$$V_\alpha(\tilde{s}_1) > V_\alpha(\bar{s}_1), \quad (15)$$
$$V_\alpha(\tilde{s}_5) \geq V_\alpha(\tilde{s}_2), \quad (16)$$
$$V_\alpha(\tilde{s}_5) \geq V_\alpha(\tilde{s}_1). \quad (17)$$

In the next lemma, we show that if no first sub-packet of a sampled packet is delivered to the monitor, then continuing transmission of that packet is not optimal for the sampler.

Lemma 3. *For state $\tilde{s}_2$ action $a = 2$, can be omitted to find an optimal action, i.e., either action $a = 2$ is not an optimal action or there exists another action which is as good as action $a = 2$.*

From Lemma 3, we know that for a state $\tilde{s}_2$, we can find an optimal action between actions $a = 1$ and $a = 0$. In the next lemma, we show that an optimal action between actions $a = 1$ and $a = 0$, for state $\tilde{s}_2$, depends on the age of the monitor v_1 and is independent of the age of the packet v_2.

Lemma 4. *If action $a = 1$ ($a = 0$) is an optimal action for a state $(v_1, v_2, 0, 1)$, then for any other state $(v_1, v_2', 0, 1)$, action $a = 1$ ($a = 0$) is also optimal. Moreover,*

$$V_\alpha((v_1, v_2, 0, 1)) = V_\alpha((v_1, v_2', 0, 1)). \quad (18)$$

Corollary 1. *For $0 < \alpha < 1$, $n \in \mathbb{N}$, the following relation holds,*

$$V_{\alpha,n}(\tilde{s}_5) = V_{\alpha,n}(\tilde{s}_7). \quad (19)$$

In the next lemma, we study a structural property of the discounted value function, which will be useful to prove a structural property of an optimal policy.

Lemma 5. *Between action $a = 2$ and $a = 0$, if action $a = 0$ is strictly better or as good as action $a = 2$, corresponding to the function $V_{\alpha,n}(\bar{s}_1)$, then action $a = 2$ cannot be optimal, for the function $V_{\alpha,n-1}(\bar{s}_2)$.*

Proof: We prove this lemma by contradiction. Thus, we assume that,

$$V_{\alpha,n}(\bar{s}_1; 2) \geq V_{\alpha,n}(\bar{s}_1; 0). \tag{20}$$

However,

$$V_{\alpha,n-1}(\bar{s}_2; 2) \leq V_{\alpha,n-1}(\bar{s}_2; a), \quad a \in \{0, 1\}. \tag{21}$$

If we expand both sides of (20), we get,

$$\begin{aligned} &pqV_{\alpha,n-1}(\bar{s}_4) + (1-p)qV_{\alpha,n-1}(\bar{s}_5) + (1-p)(1-q) \\ &V_{\alpha,n-1}(\bar{s}_3) \geq (1 - p(1-q))V_{\alpha,n-1}(\bar{s}_2) \end{aligned} \tag{22}$$

From the assumption in (21), we know that action $a = 2$ is optimal for $V_{\alpha,n-1}(\bar{s}_2)$. Thus, from (22) we have,

$$\begin{aligned} &pqV_{\alpha,n-1}(\bar{s}_4; 0) + (1-p)qV_{\alpha,n-1}(\bar{s}_5; 0) + (1-p)(1-q) \\ &V_{\alpha,n-1}(\bar{s}_3; 0) \geq (1 - p(1-q))V_{\alpha,n-1}(\bar{s}_2; 2). \end{aligned} \tag{23}$$

For the simplicity of the presentation, we define the following states,

$$\begin{aligned} &\bar{\bar{s}}_1 = (v_2 + 2, \infty, 0, 1),\ \bar{\bar{s}}_2 = (v_2 + 2, \infty, 0, 0), \\ &\bar{\bar{s}}_3 = (v_1 + 2, v_2 + 2, 1, 1),\ \bar{\bar{s}}_4 = (v_1 + 2, v_2 + 2, 1, 0). \end{aligned} \tag{24}$$

Simplifying (23), we have,

$$\begin{aligned} &(1 - pq)V_{\alpha,n-2}(\bar{\bar{s}}_1) \geq q(1-p)V_{\alpha,n-2}(\bar{\bar{s}}_2) \\ &+ (1-q)pV_{\alpha,n-2}(\bar{\bar{s}}_3) + (1-p)(1-q)V_{\alpha,n-2}(\bar{\bar{s}}_4). \end{aligned} \tag{25}$$

Now, from Lemma 1, Lemma 2, we have,

$$\begin{aligned} &V_{\alpha,n-2}(\bar{\bar{s}}_2) > V_{\alpha,n-2}(\bar{\bar{s}}_1),\ V_{\alpha,n-2}(\bar{\bar{s}}_3) \geq V_{\alpha,n-2}(\bar{\bar{s}}_1), \\ &V_{\alpha,n-2}(\bar{\bar{s}}_4) \geq V_{\alpha,n-2}(\bar{\bar{s}}_1). \end{aligned} \tag{26}$$

From (25) and (26), we have,

$$V_{\alpha,n-2}(\bar{\bar{s}}_1) > V_{\alpha,n-2}(\bar{\bar{s}}_1), \tag{27}$$

which is a contradiction. ■

In the next lemma, we show two important properties of the discounted cost function, leading us to a result about two properties of an optimal policy. First, we show that if for state $\bar{s}_1$, action $a = 0$ ($a = 1$) is optimal corresponding to the function $V_{\alpha,n}(\cdot)$, then for any state $s_1' = (v_1, {v_2}', 1, 1)$, action $a = 0$ ($a = 1$) is optimal, ${v_2}' \geq v_2$ and $n \in \mathbb{N}$. Moreover, $V_{\alpha,n}(\bar{s}_1) = V_{\alpha,n}(s_1')$.

Lemma 6. *For all $n \in \mathbb{N}$, if*

$$V_{\alpha,n}(\bar{s}_1; a) \leq V_{\alpha,n}(\bar{s}_1; 2), \tag{28}$$

then

$$V_{\alpha,n}(s_1'; a) \leq V_{\alpha,n}(s_1'; 2), \tag{29}$$

where $a \in \{0, 1\}$. Moreover,

$$V_{\alpha,n}(\bar{s}_1) = V_{\alpha,n}(s_1'). \tag{30}$$

Furthermore, action $a \in \{0, 1\}$ which satisfies (28) and (29), belongs to the set,

$$\underset{a \in \{0,1\}}{\arg\max}\, V_{\alpha,n}(\tilde{s}_6; a). \tag{31}$$

In the next lemma, we show that if the receiver has a sub-packet and if the age of the packet is 1, then $a = 1$ cannot be optimal.

Lemma 7. *Under an optimal stationary policy π, $a = 1$ cannot be optimal for state $\bar{s}_8$.*

Corollary 2. *For large enough $\alpha < 1$, the following relation holds,*

$$V_\alpha(\bar{s}_8; 1) \geq V_\alpha(\bar{s}_8; a), \quad a \in \{0, 2\}. \tag{32}$$

To prove the future results of this paper, we extend the countable state space $\mathcal{S}$, to an uncountable state space. Specifically, we assume that the age of the user v_1 and the age of the update packet v_2 can take values from the set of real numbers. We denote this new state space with $\bar{\mathcal{S}}$. Thus,

$$\begin{aligned} \bar{\mathcal{S}} = \big\{&(v_1, v_2, c, d) \cup (v_1, \infty, 0, d'), (v_1, v_2) \in \mathbb{R}^2, \\ &(c, d, d') \in \{0, 1\}^3\big\}. \end{aligned} \tag{33}$$

We denote a new MDP $\bar{\Delta}$, with the state space $\bar{\mathcal{S}}$, the same action space, cost, and transition probability of the MDP Δ. We denote the α optimal value function on $\bar{\mathcal{S}}$ with $\bar{V}_\alpha(s)$, $s \in \bar{\mathcal{S}}$. Note that, because of the structure of the transition probability of $\bar{\Delta}$, if $\bar{\Delta}$ starts with any arbitrary state $s \in \bar{\mathcal{S}}$, the evolution of states belongs to a countable subset of $\bar{\mathcal{S}}$. Depending on the initial state this sub-set changes accordingly. For example, consider a state $s = (4.5, 3.5, 1, 1) \in \bar{\mathcal{S}}$. Now, consider the following sets of states,

$$\mathcal{S}_1 = \{(4.5 + x, 3.5 + x, 1, d), x \in \mathbb{N}, d \in \{0, 1\}\}, \tag{34}$$

$$\mathcal{S}_2 = \{(3.5 + x, \infty, 0, d), x \in \mathbb{N}, d \in \{0, 1\}\}, \tag{35}$$

$$\mathcal{S}_3 = \{(4.5 + x, y, c, d), (x, y) \in \mathbb{N}^2, (c, d) \in \{0, 1\}^3\}. \tag{36}$$

Note that under any policy the evolution of state s belongs to the set $\mathcal{S} \cup \mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3$, which is a countable subset of $\bar{\mathcal{S}}$. Thus, for any $s \in \bar{\mathcal{S}}$, and $0 < \alpha < 1$, if $\bar{V}_{\alpha,n}(s)$ follows the similar iteration of (10), then the following holds,

$$\lim_{n \to \infty} \bar{V}_{\alpha,n}(s) = \bar{V}_\alpha(s). \tag{37}$$

Also note that for any $s \in \mathcal{S}$, the following relation holds,

$$\bar{V}_\alpha(s) = V_\alpha(s). \tag{38}$$

Lemma 8. *For $0 < \alpha < 1$, the function $\bar{V}_\alpha(s)$, $s = (v_1, v_2, c, d) \in \bar{\mathcal{S}}$, is a concave function of v_1.*

As, $\bar{V}_\alpha(s)$ is a concave function of v_1, the left sided derivative of $\bar{V}_\alpha(s)$ with respect to v_1 exists, which is

$$\frac{\partial \bar{V}_{\alpha,n}(s)^-}{\partial v_1} = \lim_{h\to 0+} \frac{\bar{V}_{\alpha,n}((v_1,v_2,c,d)) - \bar{V}_{\alpha,n}((v_1-h,v_2,c,d))}{h}. \quad (39)$$

The following remark will be useful to prove future results.

Remark 1. *If $F(x)$ is an increasing concave function of x, then the following relation holds true,*

$$\int_{x_1}^{x_2} \frac{\partial F(x)^-}{\partial x} dx = F(x_2) - F(x_1). \quad (40)$$

Theorem 2. *For $0 < \alpha < 1$ and $n \in \mathbb{N}$, the following relations hold,*

$$\frac{\partial \bar{V}_{\alpha,n}(\tilde{s}_6)^-}{\partial v_1} \geq \frac{\partial \bar{V}_{\alpha,n}(\bar{s}_1)^-}{\partial v_1}, \quad (41)$$

$$\frac{\partial \bar{V}_{\alpha,n}(\tilde{s}_7)^-}{\partial v_1} \geq \frac{\partial \bar{V}_{\alpha,n}(\tilde{s}_1)^-}{\partial v_1}, \quad (42)$$

$$\frac{\partial \bar{V}_{\alpha,n}(\tilde{s}_6)^-}{\partial v_1} \geq (1-q)\frac{\partial \bar{V}_{\alpha,n}(\bar{s}_1)^-}{\partial v_1}. \quad (43)$$

Proof: We prove (41)-(43) together with mathematical induction on n. Note that, for $n = 1$, for any state $s \in \bar{\mathcal{S}}$,

$$\frac{\partial \bar{V}_{\alpha,1}(s)^-}{\partial v_1} = 1. \quad (44)$$

Now, we assume that (41)-(43) hold for the nth induction step. First, we show that (41) holds for the $(n+1)$th induction step. First, let us assume that $a = 1$ is optimal for state $\bar{s}_1$. Thus, from Lemma 6, we say that $a = 1$ is also optimal for state $\tilde{s}_6$. Thus,

$$\frac{\partial \bar{V}_{\alpha,n}(\tilde{s}_6;1)^-}{\partial v_1} = 1+\alpha\Big(pq\frac{\partial \bar{V}_{\alpha,n}(\bar{s}_8)^-}{\partial v_1} + (1-p)q \frac{\partial \bar{V}_{\alpha,n}(\bar{s}_9)^-}{\partial v_1} + p(1-q)\frac{\partial \bar{V}_{\alpha,n}(\bar{s}_7)^-}{\partial v_1} + (1-p)(1-q)\frac{\partial \bar{V}_{\alpha,n}(\bar{s}_6)^-}{\partial v_1}\Big) = \frac{\partial \bar{V}_{\alpha,n+1}(\bar{s}_1;1)^-}{\partial v_1}. \quad (45)$$

Now, assume that $a = 0$ is optimal for state $\bar{s}_1$, then from Lemma 6, we say that $a = 0$ is also optimal for state $\tilde{s}_6$,

$$\frac{\partial \bar{V}_{\alpha,n+1}(\bar{s}_1;0)^-}{\partial v_1} = 1 + \alpha\frac{\partial \bar{V}_{\alpha,n}(\bar{s}_2)^-}{\partial v_1}, \quad (46)$$

$$\frac{\partial \bar{V}_{\alpha,n+1}(\tilde{s}_6;0)^-}{\partial v_1} = 1 + \alpha\frac{\partial \bar{V}_{\alpha,n}(\tilde{s}_8)^-}{\partial v_1}. \quad (47)$$

Thus, from the nth induction step of (41), (46) and (47), we have,

$$\frac{\partial \bar{V}_{\alpha,n+1}(\bar{s}_1;0)^-}{\partial v_1} \geq \frac{\partial \bar{V}_{\alpha,n+1}(\tilde{s}_6;0)^-}{\partial v_1}. \quad (48)$$

Note that, only $a = 0$ or $a = 1$, can be optimal for state $\tilde{s}_6$. Now, we assume that $a = 2$ is optimal for $\bar{s}_1$ and $a = 0$ is optimal for $\tilde{s}_6$,

$$\frac{\partial \bar{V}_{\alpha,n+1}(\bar{s}_1;2)^-}{\partial v_1} = 1 + \alpha\Big((1-q)(1-p)\frac{\partial \bar{V}_{\alpha,n}(\bar{s}_3)^-}{\partial v_1} + (1-q)p\frac{\partial \bar{V}_{\alpha,n}(\bar{s}_2)^-}{\partial v_1}\Big). \quad (49)$$

As for all $s \in \bar{\mathcal{S}}$, $\frac{\partial \bar{V}_{\alpha,n}(\bar{s})^-}{\partial v_1} > 0$, from the nth induction step of (41) and (43), and from (47) and (49), we say that,

$$\frac{\partial \bar{V}_{\alpha,n+1}(\tilde{s}_6;0)^-}{\partial v_1} \geq \frac{\partial \bar{V}_{\alpha,n+1}(\bar{s}_1;2)^-}{\partial v_1}. \quad (50)$$

Now, assume that $a = 1$ is optimal for $\tilde{s}_6$ and $a = 2$ is optimal for $\bar{s}_1$. From (45), (49), and from the nth induction step of (41) and (43), we have,

$$\frac{\partial \bar{V}_{\alpha,n+1}(\tilde{s}_6;1)^-}{\partial v_1} \geq \frac{\partial \bar{V}_{\alpha,n+1}(\bar{s}_1;2)^-}{\partial v_1}. \quad (51)$$

Thus, from (45), (48), (50) and (51), we have the $(n+1)$th induction step of (41). Now, we show that (42) holds for the $(n+1)$th induction stage. Note that, only $a = 0$ is feasible for both $\tilde{s}_7$ and $\tilde{s}_1$. Thus,

$$\frac{\partial \bar{V}_{\alpha,n+1}(\tilde{s}_7)^-}{\partial v_1} = 1 + \alpha\Big(p\frac{\partial \bar{V}_{\alpha,n}(\tilde{s}_8)^-}{\partial v_1} + (1-p)\frac{\partial \bar{V}_{\alpha,n}(\tilde{s}_9)^-}{\partial v_1}\Big), \quad (52)$$

$$\frac{\partial \bar{V}_{\alpha,n+1}(\tilde{s}_1)^-}{\partial v_1} = 1 + \alpha\Big(p\frac{\partial \bar{V}_{\alpha,n}(\bar{s}_2)^-}{\partial v_1} + (1-p)\frac{\partial \bar{V}_{\alpha,n}(\bar{s}_3)^-}{\partial v_1}\Big). \quad (53)$$

From (52), (53), and the nth induction step of (41) and (42), we have,

$$\frac{\partial \bar{V}_{\alpha,n+1}(\tilde{s}_7)^-}{\partial v_1} \geq \frac{\partial \bar{V}_{\alpha,n+1}(\tilde{s}_1)^-}{\partial v_1}, \quad (54)$$

which proves the $(n+1)$th induction stage of (42). Finally, we show that the relation in (43) holds for $(n+1)$. From (45), (53), and from the nth induction step of (41) and (42), we have,

$$\frac{\partial \bar{V}_{\alpha,n+1}(\tilde{s}_6;1)^-}{\partial v_1} \geq (1-q)\frac{\partial \bar{V}_{\alpha,n+1}(\bar{s}_1)^-}{\partial v_1}. \quad (55)$$

Similarly, from (53), and the nth induction step of (41) and (43), we have,

$$\frac{\partial \bar{V}_{\alpha,n+1}(\tilde{s}_6;0)^-}{\partial v_1} \geq (1-q)\frac{\partial \bar{V}_{\alpha,n+1}(\bar{s}_1)^-}{\partial v_1}. \quad (56)$$

Thus, from (55) and (56), we have (43) for $(n+1)$. ■

In the next theorem, we show a threshold structure of an optimal policy.

Theorem 3. *Under an optimal policy π^*, if action $a = 2$ is as good or better than action $a = 1$ for a state $\bar{s}_1$, then for any state $s = (v_1 + x, 1, 1, 1)$, action $a = 2$ would be as good or better than action $a = 1$, where $x \geq 0$.*

Proof: First, we show that the statement of this lemma holds for the discounted cost value function, i.e., for any $\bar{s}_1 \in \mathcal{S}$, we show the following,

$$V_\alpha(\bar{s}_1; 2) - V_\alpha(\bar{s}_1; 1) \geq V_\alpha(s; 2) - V_\alpha(s; 1). \quad (57)$$

Then, with a similar argument as [21], we will have the statement of this lemma. From Theorem 2, and Remark 1, we have,

$$\bar{V}_{\alpha,n}(\bar{s}_1; 2) - \bar{V}_{\alpha,n}(\bar{s}_1; 1) \geq \bar{V}_{\alpha,n}(s; 2) - V_{\alpha,n}(s; 1). \quad (58)$$

Now, from (37) and (38), we have (57). ■

For the rest of this paper, we assume that the age of the monitor is bounded by an integer K. We denote this bounded state space with $\mathcal{S}_K$. Thus,

$$\mathcal{S}_K = \big\{(v_1, v_2, c, d) \mid (v_1, v_2) \in N^2, v_1 \leq K, v_2 \leq v_1, (c, d) \in \{0,1\}^2\big\} \cup \big\{(v_1, \infty, 0, d) \mid v_1 \in N, v_1 \leq K, d \in \{0,1\}\big\}. \quad (59)$$

We denote the transition probability from a state $s = (v_1, v_2, c, d)$ to $s' = (v'_1, v'_2, c, d)$, $(s, s') \in \mathcal{S}_K^2$, under an action a with $P_a(s, s'; K)$, which we define as follows,

$$P_a(s, s'; K) = \begin{cases} P_a(s, s''), & v''_2 = \infty, \\ & v'_1 = \min\{K, v''_1\}, \\ P_a(s, s''), & v''_2 \neq \infty, \\ & v'_1 = \min\{K, v''_1\}, \\ & v'_2 = \min\{K, v''_2\}, \end{cases} \quad (60)$$

where $s'' = (v''_1, v''_2, c, d) \in \mathcal{S}$. We denote the MDP with state space of (59), transition probability of (60), cost and action space of Δ, with Δ_K. Now, we consider the following iteration on $n \in \mathbb{N}$, for $s \in \mathcal{S}$,

$$h^{n+1}(s) = (1-\gamma)h^n(s) + \min_a\{C(s,a) + \gamma \sum_{s' \in \mathcal{S}} P_a(s, s')h^n(s')\} - \min_a\{C(t,a) + \gamma \sum_{s' \in \mathcal{S}} P_a(t, s')h^n(s')\}, \quad (61)$$

where $t = (2, \infty, 0, 1)$, $0 < \gamma < 1$, and $h^0(s) = 0$, for all $s \in \mathcal{S}$. From [19], we say that,

$$\lim_{n\to\infty} h^n(s) = \frac{h^*(s)}{\gamma}, \quad (62)$$

$$\lim_{n\to\infty} \min_a\{C(t,a) + \gamma \sum_{s' \in \mathcal{S}} P_a(t, s')h^n(s')\} = \lambda, \quad (63)$$

where λ is the optimal average cost for Δ_K, and $h^*(s)$ follows the following relation,

$$h^*(s) + \lambda = \min_a\{C(s,a) + \sum_{s' \in \mathcal{S}} P_a(s, s')h^*(s')\}. \quad (64)$$

We define $h^*(s; a)$ as,

$$h^*(s; a) = C(s,a) + \sum_{s' \in \mathcal{S}} P_a(s, s')h^*(s'). \quad (65)$$

Also, note that a stationary policy that satisfies the right-hand side of (64) is optimal.

Remark 2. *Similar to Lemma 1, we can show that $h^*(s)$ is an increasing function of v_1.*

Note that, for every v_1, there exists at least one $x > 0$ such that for state $(v_1 + x, \infty, 0, 1)$, $a = 1$ is optimal, as otherwise, the average age of the system will be unbounded, as $q < 1$. Now, we show that, if the age of the monitor is large enough, then action $a = 1$ is always better than action $a = 0$, when there is no available packet for transmission. Note that, for a given p and q, the optimal average age of the system λ is fixed. Thus, we choose v_1 such that $v_1 \geq \lambda$. Now, we assume a state $(v_1, \infty, 0, 1)$. Let us assume that $x > 0$ is the minimum integer such that $a = 1$ is optimal for state $(v_1 + x, \infty, 0, 1)$. Then, from (64), we have,

$$\begin{aligned} h^*((v_1+1, \infty, 0, 1)) &= \sum_{k=1}^{x}(v_1 + k) - \lambda x + pq \\ & h^*((v_1 + x + 1, 1, 1, 1)) + p(1-q)h^*((v_1 + x + 1, 1, 0, 1)) \\ & + (1-p)qh^*((v_1 + x + 1, 1, 1, 0)) \\ & + (1-p)(1-q)h^*((v_1 + x + 1, 1, 0, 0)). \end{aligned} \quad (66)$$

As $v_1 \geq \lambda$, and as $h^*(s)$ is an increasing function of v_1, we say that action $a = 1$ is optimal for state $(v_1, \infty, 0, 1)$. We present this result in the following theorem.

Theorem 4. *For a large enough age of the monitor, an optimal policy follows a threshold property on the age of the monitor when there is no available packet for transmission. In other words, there exists v_1, such that if action $a = 1$ is optimal for a state $(v_1, \infty, 0, 1)$, then action $a = 1$ is also optimal for a state $(v'_1, \infty, 0, 1)$, where $v'_1 \geq v_1$.*

From Theorem 4, we see that a threshold structure exists on the age of the monitor for an optimal policy, when the age of the monitor is large enough, however, it may not be true in general. Next, we assume that the threshold structure in Theorem 4 holds for any v_1, and propose a corresponding optimal policy.

Assumption 1. *If action $a = 1$ is optimal for a state $(v_1, \infty, 0, 1)$, then for any state $(v_1 + x, \infty, 0, 1)$, action $a = 1$ is also optimal, $x \geq 0$.*

Now, we consider the following policy $\hat{\pi}$. For this policy: Recall that $\ell = 2$, and an update (a packet) has two sub-packets, a first sub-packet and a second sub-packet. An update will be received at the monitor and will decrease the age, when both of the sub-packets are received.

- If there is no packet to transmit, then the sampler does not sample a new packet till the age of the monitor reaches a threshold, $v_{th,1}$.
- After the age of the monitor reaches or crosses $v_{th,1}$, the sampler keeps sampling and transmitting new update packet whenever there is energy available in the battery, till there is a successful transmission of a first sub-packet.
- When there is a successful transmission of a first sub-packet, the sampler keeps transmitting the second sub-

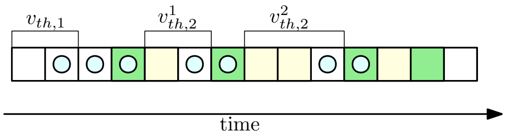

Fig. 2. A pictorial representation of a policy from the set $\hat{\Pi}$. We consider that the initial age of the system is 1 unit. We consider that $p = 1$, i.e., at every time instant energy is available in the battery. The blue circles represent the sampling of a fresh update packet. Green boxes represent a successful transmission of a sub-packet, whereas the yellow boxes represent the failed transmission of a second sub-packet. Note that, for this policy we assume that $v_{th,1} = 2$, $v_{th,2}^1 = 2$, $v_{th,2}^2 = 3$ and $v_{th,2}^3 = 4$.

packet whenever there is energy available in the battery, till the age of the packet reaches a threshold, $v_{th,2}^1$.

- When the age of the packet reaches or crosses $v_{th,2}^1$, the sampler drops the packet, samples a new packet and transmits its first sub-packet, if there is energy available in the battery.
- The sampler keeps sampling new update packets whenever there is energy available in the battery, till there is a successful transmission of a first sub-packet.
- After that the sampler keeps transmitting the second sub-packet, whenever there is energy available in the battery till, the age of the packet reaches or crosses a threshold $v_{th,2}^2$.
- When the age of the packet reaches or crosses $v_{th,2}^2$, the sampler drops the packet, samples a new packet and transmits its first sub-packet, if there is energy available in the battery.
- This process goes on till there is a successful transmission of both sub-packets, and there is no packet to transmit.
- The threshold on the age of the packet when it is dropped increases with the first successful sub-packet index, i.e., $v_{th,2}^1 \leq v_{th,2}^2 \leq \cdots$.

We denote the set of all policies with the structure of $\hat{\pi}$ with $\hat{\Pi}$, and denote an optimal policy in the set $\hat{\Pi}$ with $\hat{\pi}^*$.

Theorem 5. *Under Assumption 1, $\hat{\pi}^*$ is optimal for Δ_K.*

Proof: Let us assume that the initial state of the system is $(2, \infty, 0, 1)$. Now, we set $v_{th,1}$ to be v_1 in Assumption 1. Then, from Assumption 1, for any state $(v_{th,1} + x, \infty, 0, 1)$, $x \geq 0$, $a = 1$ is optimal. From Lemma 6, we say that for any state $(v_{th,1} + x, \infty, 0, 1)$, $x > 0$, either action $a = 2$ or action $a = 1$ is optimal. Furthermore, from Theorem 3, we say that,

$$v_{th,2}^1 \leq v_{th,2}^2 \leq \cdots , \tag{67}$$

which proves this theorem. ■

Similar to before, we extend state space $\mathcal{S}_K$ to $\bar{\mathcal{S}}_K$, where v_1 and v_2 can take values from the set of real numbers and bounded by K. Similar to Lemma 8, we can show that the function $h^*(s)$ is a concave function of v_1 on state space $\bar{\mathcal{S}}_K$. Thus, $\frac{\partial h^*(s)^-}{\partial v_1}$ exists. Next, we present a sufficient condition for which Assumption 1 holds.

Theorem 6. *If the following relation*

$$1 + \frac{\partial h^*(\tilde{s}_8)^-}{\partial v_1} \geq \frac{\partial h^*(\tilde{s}_6)^-}{\partial v_1}, \tag{68}$$

holds for all $K \geq v_1 \geq v_1'$, then an optimal policy has a threshold structure on the age of the monitor, when there is no packet to transmit. Specifically, if action $a = 1$ is optimal for a state $\tilde{s}_6$, then for any state $(v_1 + x, \infty, 0, 1)$ also $a = 1$ is optimal, where $x > 0$, and

$$\begin{aligned} v_1' = \inf \big\{ v_1 \mid pqh^*(\bar{s}_8) + p(1-q)h^*(\bar{s}_7) + (1-p)qh^*(\bar{s}_9) \\ + (1-p)(1-q)h^*(\bar{s}_6) \leq h^*(\bar{s}_9) \big\}. \end{aligned} \tag{69}$$

Proof: From (69), we say that v_1' is the least age of the monitor for which the following holds,

$$h^*((v_1', \infty, 0, 1); 1) \leq h^*((v_1', \infty, 0, 1); 0). \tag{70}$$

Thus,

$$\frac{\partial h^*(\tilde{s}_6)^-}{\partial v_1}\Big|_{v_1 = v_1'} = \frac{\partial h^*(\tilde{s}_6; 1)^-}{\partial v_1}\Big|_{v_1 = v_1'} \tag{71}$$

Similarly,

$$\frac{\partial h^*(\tilde{s}_6; 0)^-}{\partial v_1}\Big|_{v_1 = v_1'} = 1 + \frac{\partial h^*(\tilde{s}_8)^-}{\partial v_1}\Big|_{v_1 = v_1'} \tag{72}$$

From (68), (71) and (72), we have,

$$\frac{\partial h^*(\tilde{s}_6; 1)^-}{\partial v_1}\Big|_{v_1 = v_1'} \leq \frac{\partial h^*(\tilde{s}_6; 0)^-}{\partial v_1}\Big|_{v_1 = v_1'}. \tag{73}$$

Thus, from (71) and Remark 1, we have the statement of this theorem. ■

Next, we propose a low-complexity heuristic policy. We consider a policy $\bar{\pi}$ with a similar structure to $\hat{\pi}$. However, we assume that for policy $\bar{\pi}$ the following relation holds,

$$v_{th,2}^1 = v_{th,2}^2 = \cdots = v_{th,2}. \tag{74}$$

That is, we omit the superscript from the second threshold, and keep it constant. Thus, policy $\bar{\pi}$ is completely characterized by two thresholds, $v_{th,1}$ and $v_{th,2}$. We denote the set of all policies with the structure of $\bar{\pi}$ with $\bar{\Pi}$. We denote an optimal policy in the set $\bar{\pi}$ with $\bar{\pi}^*$. Next, we find the average age expression for the policy $\bar{\pi}$ for a fixed $v_{th,1}$ and $v_{th,2}$. We can divide the whole time horizon into several frames. We assume that a frame ends when the monitor successfully decodes a packet. We denote the length of a frame with a random variable $\bar{\ell}$. Now, we find the first and second moments of $\bar{\ell}$. Note that, the initial age of a frame depends on the delivery time of the last packet of the previous frame, which we denote with a random variable $\bar{Y}$, which follows the following distribution,

$$\mathbb{P}[\bar{Y} = i] = pq(1-pq)^{i-1}, \quad i \leq v_{th,2} - 1. \tag{75}$$

Note that, at the start of a frame, we do not sample a new packet till the age of the monitor reaches or exceeds the threshold $v_{th,1}$. We denote the initial period during which the sampler stays idle, with a random variable $\tilde{Y}$. Note that,

$\tilde{Y} = \max\{v_{th,1} - \bar{Y} - 2, 0\}$, and

$$\mathbb{P}(\tilde{Y} = 0) = \mathbb{P}(\bar{Y} \geq v_{th,1} - 2), \quad (76)$$

$$\mathbb{P}(\tilde{Y} = i) = \mathbb{P}(\bar{Y} = v_{th,1} - 2 - i), 1 \leq i \leq (v_{th,1} - 3). \quad (77)$$

We denote the time interval needed for the successful transmission of the first sub-packet after the random wait time $\tilde{Y}$, with a random variable $\hat{X}$. The random variable $\hat{X}$ for a given realization of $\tilde{Y}$ is,

$$\hat{X} = \mathbf{1}\{E_1 \cap E_2\} + \mathbf{1}\{E_1 \cap E_2^c\}(1 + \bar{X}) + \mathbf{1}\{E_1^c\}\bar{X}, \quad (78)$$

where E_1 is the event that a unit of energy arrives at the battery during $\tilde{Y}$, E_2 is an event of getting a head in a coin toss with probability q, and $\mathbf{1}\{A\}$, is the indicator function of an event A. Let us consider any arbitrary t, such that there is no sub-packet at the receiver and no energy at the battery till the end of time slot t. We denote the duration from time slot $t+1$ to the time slot in which the receiver receives a sub-packet for the first time after time t, with the random variable $\bar{X}$ in (78). Thus, $\bar{X}$ is geometrically distributed with the parameter $\frac{1}{pq}$. In a frame, we denote the number of times the sampler drops a packet for which there is a successful transmission of the first sub-packet, with a random variable N. Note that, N can take value 0, and is geometrically distributed with a parameter, $1 - (1 - pq)^{v_{th,2} - 1}$. Thus,

$$\mathbb{E}[\bar{\ell}] = \mathbb{E}[\tilde{Y} + \hat{X} + \bar{Y}] + \mathbb{E}[N](\mathbb{E}[\bar{X}] + (v_{th,2} - 1)), \quad (79)$$

$$\begin{aligned}\mathbb{E}[\bar{\ell}^2] &= \mathbb{E}[\tilde{Y}^2 + \hat{X}^2 + \bar{Y}^2 + N^2(v_{th,2} - 1)^2 \\ &+ 2N^2\bar{X}(v_{th,2} - 1) + N\bar{X}^2 + \mathbb{E}[\bar{X}]^2(N^2 - N) + 2(\tilde{Y}\hat{X} + \tilde{Y}\bar{Y} \\ &+ \hat{X}\bar{Y} + (\bar{X} + v_{th,2} - 1)(\tilde{Y}N + \hat{X}N + \bar{Y}N))]. \quad (80)\end{aligned}$$

From (80), we see that $\mathbb{E}[\bar{\ell}^2] < \infty$. Note that, in every frame the age process is i.i.d., thus we use the renewal reward theorem to find the average age. Thus,

$$\Delta^{\bar{\pi}}(s_0) = \frac{1}{2} + \mathbb{E}[\bar{Y}] + \frac{1}{2}\frac{\mathbb{E}\bar{\ell}^2}{\mathbb{E}[\bar{\ell}]}. \quad (81)$$

Note that, (81) is independent of the initial state s_0.

IV. Numerical Result

We compare the performance between an optimal policy π^* and an optimal policy $\bar{\pi}^*$ in the set $\bar{\Pi}$. We find the average age of π^* from (62). We bound the instantaneous age of the monitor with 1000, i.e., we assume that $K = 1000$. We consider the policy set $\bar{\Pi}$, with $v_{1,th}$ and $v_{2,th}$ taking values from 1 to 100. To find the average age of $\bar{\pi}^*$, we use (81), and compare the average age for all the policies in $\bar{\pi}$. We see from Fig. 3, that the performance of $\bar{\pi}^*$ is very close to the performance of π^*.

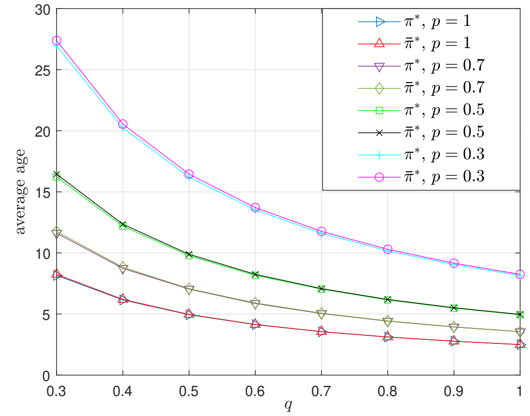

Fig. 3. Average age comparison between the low-complexity policy $\bar{\pi}^*$ and an optimal policy π^* .

References

[1] R. D. Yates, Y. Sun, R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus. Age of information: An introduction and survey. *IEEE Journal on Selected Areas in Communications*, 39(5):1183–1210, May 2021.

[2] E. Najm, R. D. Yates, and E. Soljanin. Status updates through M/G/1/1 queues with HARQ. In *IEEE ISIT*, June 2017.

[3] R. D. Yates, E. Najm, E. Soljanin, and J. Zhong. Timely updates over an erasure channel. In *IEEE ISIT*, June 2017.

[4] A. Arafa, K. Banawan, K. G. Seddik, and H. V. Poor. On timely channel coding with hybrid ARQ. In *IEEE Globecom*, December 2019.

[5] E. T. Ceran, D. Gündüz, and A. György. A reinforcement learning approach to age of information in multi-user networks with HARQ. *IEEE Journal on Selected Areas in Comm.*, 39(5):1412–1426, May 2021.

[6] E. T. Ceran, D. Gündüz, and A. György. Average age of information with hybrid ARQ under a resource constraint. *IEEE Transactions on Wireless Communications*, 18(3):1900–1913, March 2019.

[7] S. Feng and J. Yang. Age-optimal transmission of rateless codes in an erasure channel. In *IEEE ICC*, May 2019.

[8] S. Farazi, A. G. Klein, and D. R. Brown. Average age of information in update systems with active sources and packet delivery errors. *IEEE Wireless Communications Letters*, 9(8):1164–1168, August 2020.

[9] Y. Wang, S. Wu, J. Jiao, W. Wu, Y. Wang, and Q. Zhang. Age-optimal transmission policy with harq for freshness-critical vehicular status updates in space–air–ground-integrated networks. *IEEE Internet of Things Journal*, 9(8):5719–5729, April 2020.

[10] A. Arafa, K. Banawan, K. G. Seddik, and H. V. Poor. Timely estimation using coded quantized samples. In *IEEE ISIT*, June 2020.

[11] A. Arafa, K. Banawan, K. G. Seddik, and H. V. Poor. Sample, quantize, and encode: Timely estimation over noisy channels. *IEEE Transactions on Communications*, 69(10):6485–6499, October 2021.

[12] A. Li, S. Wu, J. Jiao, N. Zhang, and Q. Zhang. Age of information with hybrid-ARQ: A unified explicit result. *IEEE Transactions on Communications*, 70(12):7899–7914, December 2022.

[13] A. Baknina and S. Ulukus. Coded status updates in an energy harvesting erasure channel. In *CISS*, March 2018.

[14] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor. Age-minimal online policies for energy harvesting sensors with incremental battery recharges. In *ITA*, February 2018.

[15] A. Arafa and S. Ulukus. Timely updates in energy harvesting two-hop networks: Offline and online policies. *IEEE Transactions on Wireless Communications*, 18(8):4017–4030, August 2019.

[16] A. Arafa and S. Ulukus. Age minimization in energy harvesting communications: Energy-controlled delays. In *Asilomar*, October 2017.

[17] A. Arafa, J. Yang, and S. Ulukus. Age-minimal online policies for energy harvesting sensors with random battery recharges. In *IEEE ICC*, May 2018.

[18] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor. Age-minimal transmission for energy harvesting sensors with finite batteries: Online policies. *IEEE Transactions on Information Theory*, 66(1):534 – 556, September 2019.

[19] D. Bertsekas. *Dynamic Programming and Optimal Control: Volume II*. Athena Scientific, 2012.

[20] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 2014.

[21] L. I. Sennott. Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs. *Operations Research*, 37(4):626–633, July 1989.