Pull or Wait: How to Optimize Query Age of Information

M. Emrullah Ildiz, Orhan T. Yavascan, Elif Uysal, and O. Tugberk Kartal

Abstract—We study a pull-based status update communication model where a source node submits update packets to a channel with random transmission delay, at times requested by a remote destination node. The objective is to minimize the average query-age-of-information (QAoI), defined as the average age-of-information (AoI) measured at query instants that occur at the destination side according to a stochastic arrival process. In reference to a push-based problem formulation defined in the literature where the source decides to update or wait at will, with the objective of minimizing the time average AoI at the destination, we name this problem the Pull-or-Wait (PoW) problem. We identify the PoW problem in the case of a single query as a stochastic shortest path (SSP) problem with uncountable state and action spaces, which has not been solved in previous literature. We derive an optimal solution for this SSP problem and use it as a building block for the solution of the PoW problem under periodic query arrivals.

Index Terms—Age of information, Internet of Things, pull-based communication, query age of information, status updates, stochastic shortest path problem, update or wait, energy efficiency, energy efficient sensor communication.

I. INTRODUCTION

THE INTERNET of Things (IoT) paradigm has been gaining wide use in various sectors such as environmental monitoring [1], health and wellness [2], vehicular networks [3], smart cities [4]. In many applications, a destination node seeks to have accurate information about a remote process measured by a sensor to utilize toward a computation. Often, the value of new update packets to the computation at the destination side is reduced as the packet ages, or gets stale.

As a metric to measure timeliness of update packets, the age-of-information (AoI), or simply age, has been studied in many different environments [5], [6], [7]. AoI of a flow is defined as the elapsed time since the generation of the latest received update packet. This definition makes it possible to measure the freshness of information for every time point at the destination node in contrast to the traditional metric, latency, that corresponds to individual packet delay [8].

Moreover, in many applications, the goal-oriented significance of the next update to can we well captured by a nonlinear monotone function of age (e.g., remote estimation of an unstable process, location tracking). Therefore, AoI has been a useful performance metric toward effective communication. The generate-at-will model was first introduced in [9]. In this mode, the source decides to generate an update packet in order to control the ageing process [10], [11], [12], [13], [14], [15], [16], [17], [18], [19]. The optimization of this method (under general age penalty functions, and for delay with memory) was done in [11]. We will refer to the formulation in [11] as Update-or-Wait (UoW) problem. In the UoW problem, the source controls the age by determining the submission times of the update packets to the channel.

The UoW model is a good fit for the case when it is desirable to keep the age at the destination refreshed continuously. Yet in many IoT scenarios the application running at the destination side will utilize the information updates at certain times, rather than continuously [20], [21]. A policy that strives to keep the overall time average age at a minimum will not necessarily maintain minimal age at those utilization times.

In this paper, we define an extension of the UoW problem, which is referred to as the Pull-or-Wait (PoW) problem. In the PoW the destination node requests an update packet from the source node in an effort to keep a low AoI “at the next query instant”. The query-age-of-information (QAoI) is defined as the average age measured over query instants. The goal of the destination is to determine optimal request points to minimize QAoI, knowing only the statistics of the channel delay and the query arrival process. The following simple example illustrates the difference between the UoW and PoW problems, and why UoW may run into inefficiency.

Example 1: Consider an IoT monitoring system that requires an update packet every 4 milliseconds. Hence, the query instants are at times 4, 8, 12, . . . . Suppose that the transmission delay in the channel from the source to the destination is 1.5 msec, while pull requests from the destination to the source are subject to no delay. As transmission delays are constant, it follows from [11] that the zero-wait policy, which issues a new update once the previous one is received, is optimal for the UoW problem. The evolution of the age of information under the zero-wait policy is shown in Figure 1. This policy results in a time average age of information equal to 2.25 and performs one packet transmission per 1.5 msec. On the other hand, a reasonable policy, which is later shown to be an optimal policy, for the PoW problem is that the destination node requests update packets at times
The rest of the paper is organized as follows: In Section II, we discuss related work. In Section III, we define the PoW problem, and present its analysis in Section IV. In Section V, we present numerical results to illustrate the behavior of the solution in the PoW problem under different transmission delay statistics. We conclude in VI by discussing results and future directions.

II. RELATED WORK

AoI has been studied in enqueue-and-forward models [25], [26], [27], [28], [29], [30], generate-at-will models [9], [12], [13], [14], [15], [16], [17], [18], [19], [23], as well as random access environments [31], [32], [33], [34], [35]. Optimizing average age captures the timeliness aspect, but this does not suffice for ensuring effective communication in some cases. For example, the optimal policy that minimizes the MSE in the remote estimation of a Wiener process over a random delay channel is distinct from the age optimal policy as shown in [36]. As a result, various suggestions other metrics capturing goal-oriented significance of information have emerged [37], [38], [39], [40]: the Age of Incorrect Information (AoII) extends the notion of fresh updates to that of fresh “informative” updates in [41], [42], [43], [44]. Other metrics such as the Urgency of Information (UoI) and the Age of Changed Information (AoCl) have been proposed in [45] and [46], respectively.

Query Age of Information (QAoI) is defined as the average of the age measured at query instants, where query instants represent the utilization times of the destination node in the application. This notion has been introduced in an independent set of works with different names such as Age upon Decision (AuD), Age of Effective Information (AoEI) [20], [21], [47], [48], [49], [50], [51]. The first works that suggest a pull-based communication model in the context of AoI are [47], [48], where a user proactively requests update packets from multiple servers, but the authors minimize the plain AoI and do not take utilization time into account. A series of works [20], [21], [49], [50] suggests AuD and studies a special case of the enqueue-and-forward model where a user utilizes upcoming update packets under a stochastic arrival process. This model leads the authors to measure the AoI at the utilization times. The works that are most relevant to this paper are [11], [24] [52], and [53]. In [11], the authors formulated the generate-at-will model where packet generation is controlled by the sender to minimize time average age. We refer to this problem as the UoW problem. In contrast, PoW problem was defined in our previous work [52], which extended [11] to a pull-based communication model and modified the objective function with respect to QAoI. This was extended in [53]. These two works compared the PoW and UoW problems under periodic query arrivals, general transmission delays, and under general query arrivals constant transmission delays. It was shown that the query age achieved by solving the PoW problem is never above that of the optimal time average age in the UoW problem under the above scenarios. The present paper is both a full version of the shorter treatments in [52] and [53], and extends them by obtaining the solution to the PoW problem. In [24], the authors suggest the QAoI and study a similar pull-based communication model. In this paper, in contrast to the packet erasure channel that is considered in [24], we study more general channels that can have discrete, continuous, or mixed distributed transmission
delays. In addition, we define an age penalty function \( g(\Delta) \) to characterize the level of dissatisfaction for data staleness, where \( g(.) \) can be any nonnegative, continuous, and nondecreasing function. This age penalty function enables us to simulate model-specific applications. Furthermore, our minimization does not require a discount factor.

### III. SYSTEM MODEL AND PROBLEM DEFINITION

We consider a pull-based information update system depicted in Figure 2, where a destination node is interested in information updates generated by a source node. The destination node requests an update packet from the source node according to an update policy. The request arrives at the source node without any delay. When a request occurs, the source node immediately generates an update packet and submits it to the channel. The channel induces a random delay between the source node and the destination node. The destination node should not request a new update packet when the previously requested update packet has not arrived at the destination node, because this will incur an unnecessary waiting time in the queue.

The update packets delivered to the destination node are utilized toward a computation. In this information update system, we assume that the destination node possesses a query arrival process that represents the utilization time of the upcoming update packets received from the source node. The destination node aims to minimize the average AoI at the query instants. As the destination node can recognize past states of the query arrival process, it requests update packets from the source node by taking account of not only the random delays induced by the channel but also the past states of the query process.

Let the time that Update \( j \), \( j = 1, 2, \ldots \) is requested from the source, and submitted to the communication channel be denoted by \( R_j \). Update \( j \) is delivered to the destination node after a random transmission delay \( Y_j \) at time \( D_j = R_j + Y_j \). Then, the destination node requests Update \( j+1 \) at time \( R_{j+1} \) after a waiting period \( Z_j \in [0, M] \). This implies that \( R_{j+1} = D_j + Z_j \). We assume that the transmission delay process, \( \{Y_j\}_{j=0}^{\infty} \), is i.i.d. and takes values in a bounded range such that \( \Pr(Y_j \in [B_L, B_U]) = 1 \) where \( B_L > 0 \). On the other side, the query arrival process based on which the destination node utilizes the upcoming update packets is denoted as \( \{Q_k, k = 1, 2, \ldots \} \).

At any time \( t \), let \( U(t) \) denote the generation time of the update packet that has been most recently received by the destination node. Consequently,

\[
U(t) = \max \{R_j : D_j \leq t\}
\]

The age of information corresponding to this flow in the destination node at time \( t \) is denoted by \( \Delta(t) \), and is defined as:

\[
\Delta(t) = t - U(t)
\]

We also introduce an age penalty function, \( g(\Delta) \), that represents the level of dissatisfaction for data staleness or the need for a new information update. This function is defined as \( g : [0, \infty) \rightarrow [0, \infty) \) and it is continuous, nonnegative, and nondecreasing. Our goal is to minimize the average age penalty at the time of queries by controlling the sequence of waiting periods, \( (Z_0, Z_1, \ldots) \). Let \( \pi \triangleq (Z_0, Z_1, \ldots) \) denote an update policy. A causal update policy determines the waiting period \( Z_j \) based on the sequence \( (Z_i)_{i=0}^{j-1} \), the random processes \( \{Y_i\}_{i=0}^{\infty}, \{Q_i\}_{i=0}^{\infty} \) and their realizations before \( D_j \). Let \( \Pi \) be the set of all causal update policies. Then, the objective function is defined as the following:

\[
\tilde{h}_{\text{opt}} = \min_{\pi \in \Pi} \limsup_{n \to \infty} \frac{E\left[ \sum_{k=1}^{n} g(\Delta(Q_k)) \right]}{n} \tag{3}
\]

Throughout the paper, we refer to this problem as the Pull or Wait (PoW) problem. We refer to the objective function of the PoW problem as the query average age penalty.

#### 1) UoW Problem

In the system model that was studied in [11] and that is depicted in Figure 3, the source node generates update packets and sends them directly to the destination node through the channel. Different from the system model of the PoW problem, the destination node does not request an update packet in an effort to minimize age penalty at the queries as there is no query in this system model. Instead, the source node submits update packets to the channel seeking to minimize the time average age penalty at the destination node. Therefore, the objective function is the following:

\[
\tilde{g}_{\text{opt}} = \min_{\pi \in \Pi} \limsup_{n \to \infty} \frac{E\left[ \int_0^T g(\Delta(t)) \, dt \right]}{E[D_M]} \tag{4}
\]

Throughout the paper, we refer to this problem as the Update or Wait (UoW) problem. We refer to the objective function of the UoW problem as the time average age penalty.

### IV. PROBLEM FORMULATION AND ANALYSIS

In this section, we first analyze the PoW problem under a specific case of single query. Let \( Q > 0 \) be the time at which the query occurs. For this case, Problem (3) reduces to:

\[
\tilde{h}_{\text{one}}(Q) = \min_{\pi \in \Pi} E\left[ g(\Delta(Q)) \right] \tag{5}
\]

Henceforth, we will refer to Problem (5) as the “single query problem”. As we will show in the rest of this section, the solution of the single query problem will be a building block of the solution of the PoW problem, given in (3), under periodic query arrivals.

The single query problem belongs to the class of stochastic shortest path problems with uncountable state and action spaces. The state of the problem at stage \( j \) is the pair of the
remaining time from the delivery point of Update \( j \) until the query and the current age at the delivery point of Update \( j, (Q - D_j, \Delta(D_j)) \). \(^1\) The random disturbance and the control action at stage \( j \) are \( Y_j \) and \( Z_j \), respectively. The absorbing state occurs at stage \( j \) when \( Q - D_j \leq 0 \). State transitions that do not end in the absorbing state are costless. The cost of reaching the absorbing state from a state \( (Q - D_j, \Delta(D_j)) \) where \( Q - D_j > 0 \) is \( g(Q - D_j + \Delta(D_j)) \). This problem class is introduced in [54] for a finite state space, compact action space, a transition kernel that is continuous for all actions, under the assumption that an optimal policy must be proper (i.e. reachability of the termination state in a finite expected time). Reference [55] relaxes the assumptions of [54] such that the state and action spaces are arbitrary, the transition kernel does not need to be continuous, but the space of the random disturbance is countable. A related problem class is introduced by [56] as transient Markov decision problems with solutions that are transient policies (similar, but not identical, to proper policies), general state and action spaces, and continuous transition kernel. Reference [57] further relaxes the assumptions of [56] to the existence of non-transient policies, but keeps the assumption about the continuity of the transition kernel [57, Assumption 1b]. None of these results are directly applicable to the single query problem because in our problem the random disturbance \( Y_j \) may not come from a countable set and the transition kernel is not restricted to be continuous especially when the random disturbance \( Y_j \) has a mixed distribution.

In the rest of this section, we will show the existence of a deterministic optimal policy for the single query problem, and characterize its first request point in Section IV-A. With the help of this characterization, we will! reformulate the PoW problem under periodic query arrivals in terms of the single query problem in Section IV-B. Finally, we will provide a complete solution of the single query problem in Section IV-C, which concludes the solution of the PoW problem in (3) under periodic query arrivals.

A. Existence of a Deterministic Optimal Policy for the Single Query Problem

In this subsection, we first show that there exists an optimal policy, \( \pi_1^{opt} \), for the single query problem, that is a deterministic policy. Then, we define the border point of \( \pi_1^{opt} \) for a query arriving at time \( Q \), denoted as \( Q^{BP} \in [Q - 3BU, Q - B_U] \). We prove that \( Q^{BP} \) is an optimal request point under the policy \( \pi_1^{opt} \) for every delivery point \( D_j \) satisfying \( D_j < Q - 3BU \). This property will help us transform the solution of the single query problem into a solution of the PoW problem under periodic query arrivals. At any delivery point \( D_j \), an optimal update policy seeks to find a request point \( R_j+1 \) to minimize the expected age penalty at the query. To express the expected age penalty at the query in terms of a request point \( R_j \), we define the \( G_R^2 \) function. In addition to the \( G_R^2 \) function, we define the \( G_D^2 \) function to express the expected age penalty at the query in terms of a delivery point \( D_j \) as the following:

Definition I: For a given query \( Q \), let \( R_j \) and \( D_j \) be any request and delivery points, respectively. \( G_R^2 : [0, \infty) \times [0, \infty) \rightarrow [0, \infty) \) and \( G_D^2 : [0, \infty) \times [0, \infty) \rightarrow [0, \infty) \) are defined as follows:

\[
G_R^2(Q - R_j, \Delta(R_j)) \equiv E \left[ g(\Delta(Q)) \mid \pi \text{ is applied}, \right.
\]

\( R_j \) is a request point, AoI at \( R_j \) is \( \Delta(R_j) \) \hspace{1cm} (6)

\[
G_D^2(Q - D_j, \Delta(D_j)) \equiv E \left[ g(\Delta(Q)) \mid \pi \text{ is applied}, \right.
\]

\( D_j \) is a delivery point, AoI at \( D_j \) is \( \Delta(D_j) \) \hspace{1cm} (7)

These expectations are taken over the possible transmission delays and the waiting period decisions by the policy \( \pi \in \Pi \). It will be shown in Proposition 1 that the information of the remaining time until the query \( Q - D_j \) and the AoI at the delivery point \( \Delta(D_j) \) are sufficient statistics to determine an optimal waiting period. This implies that the minimization of the single query problem can be performed by only considering the set of causal policies that determines the waiting period \( Z_j \) based on \( Q - D_j \) and \( \Delta(D_j) \). Therefore, there is no need to explicitly provide the sequences of \( (Y_i)_{i=0}^{\infty} \) and \( (Z_i)_{i=0}^{\infty} \) for the functions \( G_R^2 \) and \( G_D^2 \). The two functions have a chain relationship with each other. When the destination node requests an update packet from the source node at \( R_j \), Update \( j \) is delivered to the destination node after a random transmission delay \( Y_j \) at time \( D_j = R_j + Y_j \). Hence, \( \Delta(D_j) = Y_j \). If the delivery occurs before the query, i.e., \( Q - R_j - Y_j \geq 0 \), the expected age penalty can be represented with the function \( G_R^2 \). If \( Q - R_j - Y_j < 0 \), the AoI at the query is \( Q - R_j + \Delta(R_j) \) for sure. This relationship can be written as follows:

\[
G_R^2(Q - R_j, \Delta(R_j)) = E \left[ G_D^2(Q - R_j - Y_j, Y_j) \right] \left. \right| Y_j \leq Q - R_j \right.
\]

\[
\times \Pr(Y_j \leq Q - R_j) + g(Q - R_j + \Delta(R_j)) \times \Pr(Y_j > Q - R_j) \] (8)

This expectation is taken over possible transmission delays. On the other hand, when the update packet is delivered to the destination node at \( D_j \), the destination node waits for a duration \( Z_j \) to request a new update packet. Hence, the request point is \( Q - D_j - Z_j \), and the AoI at the request point is \( \Delta(D_j) + Z_j \). When the request point is before the query, i.e., \( Q - D_j - Z_j \geq 0 \), the expected age penalty at the query can be represented with the function \( G_R^2 \). When \( Q - D_j - Z_j < 0 \), the AoI at the query is \( Q - D_j + \Delta(D_j) \) for sure. This relationship can be written as follows:

\[
G_R^2(Q - D_j, \Delta(D_j)) = E \left[ G_R^2(Q - D_j - Z_j, \Delta(D_j) + Z_j) \right] \left. \right| Z_j \leq Q - D_j \right.
\]

\[
\times \Pr(Z_j \leq Q - D_j) + g(Q - D_j + \Delta(D_j)) \times \Pr(Z_j > Q - D_j) \] (9)

\(^1\)It is shown in Proposition 1 that there exists an optimal policy of the single query problem in which \( Z_j \) is determined as a function of \( Q - D_j \) and \( \Delta(D_j) \). As a result, the single query problem can be minimized in the set of deterministic policies. When \( Z_j \) is determined as a function of \( Q - D_j \) and \( \Delta(D_j) \), the pair \( (Q - D_j, \Delta(D_j)) \) \( j \geq 0 \) forms a Markov chain because \( \Delta(D_j) = Y_j \), \( Y_j \)'s are i.i.d., and \( Q - D_{j+1} = Q - D_j - Y_j - Z_j \).

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
This expectation is taken over possible waiting periods that are determined by the policy $\pi$ in order to take randomized policies into account. Now, we move on to obtain a deterministic optimal policy of the single query problem. The optimal age penalty in this problem can be achieved in a special subset of $\Pi$. In the next proposition, we prove this in detail.

**Definition 2:** A policy $\pi \in \Pi$ is said to be a stationary and deterministic policy if there exists decision function $z : [0, \infty) \rightarrow [0, M]$ such that $Z_j = z(Y_j, Q - D_j)$ for $j = 0, 1, \ldots$

- The set of all stationary and deterministic policies is denoted as $\Pi_{SD}$.

**Proposition 1:** If the transmission delay process $\{Y_j\}_{j=0}^\infty$ is i.i.d. such that $\Pr(Y_1 \in [B_L, B_U]) = 1, M < \infty$, and the penalty function $g$ is continuous, non-negative, and non-decreasing, then there exists a deterministic update policy that is optimal for the single query problem.

**Proof:** In the proof, we need to use the extended version of the functions $G_R^\pi$ and $G_D^\pi$ that must include the sequences $(Y_i^\pi)_{i=0}^\infty$ and $(Z_i^\pi)_{i=0}^\infty$ in order to cover all possible causal update policies. Hence, they are $G_R^\pi(Q - D_j, \Delta(D_j), (Y_i^\pi)_{i=0}^j, (Z_i^\pi)_{i=0}^j)$ and $G_D^\pi(Q - R_j, \Delta(R_j), (Y_i^\pi)_{i=0}^j, (Z_i^\pi)_{i=0}^j)$. Let us map each $Q - D_j$ to a natural number $n$ satisfying $(n-1)B_L \leq Q - D_j < nB_L$. We perform discrete induction on $n$. The proposition is first proved for every $j$, $(Y_i^\pi)_{i=0}^j$, and $(Z_i^\pi)_{i=0}^j$ that satisfy $(n-1)B_L \leq Q - D_j < nB_L$ when $n = 1$. Then, the proposition is assumed to be correct when $n = 2, 3, \ldots, K$ where $K$ is an arbitrary natural number. Finally, it is proved when $n = K + 1$. The details are available in Appendix A.

According to the previous proposition, there exists a deterministic optimal update policy $\pi_1^{opt} \in \Pi_{SD}$ that decides waiting periods based on the values of $Q - D_j$ and $\Delta(D_j)$ for every $j$, $(Y_i^\pi)_{i=0}^j$, and $(Z_i^\pi)_{i=0}^j$. Interestingly, for some specific values of $Q - R_j$, the expected age penalty at the query may not depend on the value of $\Delta(R_j)$. For example, when the destination node is supposed to request an update packet from the source node before $Q - B_U$, the requested update packet must reach the destination node before the query. This is because the transmission delay can be at most $B_U$. Therefore, the AoI at the request point cannot affect the expected age penalty at the query. The next proposition proves this in detail.

**Proposition 2:** If the elapsed time since a request point until the query is greater than $B_U$, then $Y_j \leq Q - R_j$ for sure. Therefore, (8) becomes

$$G_R^\pi(Q - R_j, \Delta(R_j)) = E[G_D^\pi(Q - R_j - Y_j, Y_j)]$$

(10)

As the transmission delay process is i.i.d. and $\Delta(R_j) = Y_{j-1} + Z_{j-1}$, $\Delta(R_j)$ does not affect $G_D^\pi(Q - R_j - Y_j, Y_j)$ when $Q - R_j$ is given. Hence, the proof is completed. Note that this property is valid for every $\pi \in \Pi_{SD}$.

As a result of previous proposition, we can modify the function $G_R^{\pi_1^{opt}}$ when $\pi_1^{opt}$ is a deterministic optimal policy and $Q - R_j$ is greater than or equal to $B_U$. Hence, for every request point $R_j$ and its AoI $\Delta(R_j)$ satisfying $Q - R_j \geq B_U$, we redefine the $G_R^{\pi_1^{opt}}$ function with one argument as the following:

$$G_R^{\pi_1^{opt}}(Q - R_j) = G_R^{\pi_1^{opt}}(Q - R_j, \Delta(R_j))$$

(11)

For a given query $Q$ and a deterministic optimal policy $\pi_1^{opt}$, let us define its border point $Q^{BP}$ that satisfies the following:

$$Q^{BP} = \inf_{R_j : R_j \leq Q - B_U} G_R^{\pi_1^{opt}}(Q - R_j)$$

(12)

In the next proposition, we show the existence of a border point. Then, we specify one of these points as the border point.

**Proposition 3:** Let $D^*_j$ be a specific delivery point satisfying $D^*_j = Q - 3B_U$ and $Y^*_j = B_L$. The request point $R^*_j$ determined by a deterministic optimal policy $\pi_1^{opt}$ is a border point for the query $Q$. We designate $R^*_j$ as the “selected” border point.

**Proof:** In the proof, we first prove that the request must occur by the time $Q - B_U$, i.e., $R^*_j \leq Q - B_U$. This ensures that $R^*_j$ is in the intended interval of (12). Then, we show that the optimal request point, $R^*_j$, attains the infimum in (12). The details are in Appendix B.

In the rest, for brevity, we will refer to the selected border point as the border point. The exact location, $Q^{BP}$, of the border point depends on the exact time of the query and the optimal policy $\pi_1^{opt}$. This is because the border point is specified as the request point that is determined by $\pi_1^{opt}$ when the delivery point is $Q - 3B_U$ and the age at the delivery point is $B_L$. Hence, the border point can be considered as a function of a query $Q$ and a deterministic optimal policy $\pi_1^{opt}$. Nevertheless, there is a special property of the border point concerning the relation between $Q$ and $Q^{BP}$, proved in the following corollary:

**Corollary 1:** The time duration between a query and its border point does not depend on the exact time of the query for a given deterministic optimal policy.

**Proof:** This corollary is an immediate result of Proposition 1 and the definition of $R^*_j$. The request point $R^*_j$ determined by a deterministic optimal policy $\pi_1^{opt}$ is the border point when $D^*_j = Q - 3B_U$ and $\Delta(D^*_j) = Y^*_j = B_L$ regardless of the exact time of the query. The optimal waiting period at $D^*_j$ is solely determined by $\pi_1^{opt}$ based on $Q - D^*_j$ and $\Delta(D^*_j)$ by Proposition 1. As $Q$ changes, $Q - D^*_j$ and $\Delta(D^*_j)$ do not change. Hence, $Z^*_j$ does not change. As $R^*_j = Q - 3B_U + Z^*_j$, the proof is completed.

We next prove in Lemma 1 that if a delivery point occurs before $Q - 3B_U$, then it is optimal to wait until the border point to place a request.

**Lemma 1:** Let $Q^{BP}$ be the border point of a query $Q$ and a deterministic optimal policy $\pi_1^{opt}$. Then, for any delivery point $D^*_j$ satisfying $D^*_j < Q - 3B_U$, the border point $Q^{BP}$ is an optimal request point under the policy $\pi_1^{opt}$.

**Proof:** To reach contradiction, suppose that the claim is false. Then, there exists a delivery point $D^*_j \in [0, Q - 3B_U)$ and
an AoI at the delivery $\Delta(D_j)$ such that the request point $R_{j+1}$ determined by a deterministic optimal policy $\pi_1^{opt}$ satisfies the following: $G_{R_1}^{opt}(Q - Q^{BP}) > G_{R_1}^{opt}(Q - R_{j+1}, \Delta(R_{j+1}))$. By (12), $R_{j+1}$ cannot be in the interval $[Q - B_U, Q]$. By Lemma 3 that is given in Appendix B, $R_{j+1}$ cannot be in the interval $[Q - B_U, Q]$ as well. This completes the proof. ■

Corollary 2: There exists a deterministic optimal policy $\pi_1^{opt}$ for a given query $Q$ satisfying $Q > 3B_U$ such that the first request point is the border point.

Proof: This is an immediate result of Lemma 1 and the designation of the border point in Proposition 3.

Corollary 3: If $Q > 3B_U$, $h_{opt}^{one}(Q)$ is independent of the exact time of the query $Q$. In other words, we can define $h_{opt}^{one}$ as the following:

$$h_{opt}^{one}(Q) = G_{R_1}^{opt}(Q - Q^{BP})$$

(13)

where $\pi_1^{opt}$ is a deterministic optimal policy and $Q^{BP}$ is their border.

Proof: From Corollary 2, there exists a deterministic optimal policy $\pi_1^{opt}$ whose first request point is the border for a given query $Q$ satisfying $Q > 3B_U$. This means that $h_{opt}^{one}(Q) = G_{R_1}^{opt}(Q - Q^{BP})$. Furthermore, the time duration between $Q - Q^{BP}$ does not change when $Q$ is shifted by Corollary 1. Hence, the expected age penalty at the border point for any $Q > 3B_U$ is the same because the destination node can request an update packet at the border point under an optimal policy. As a result, we can define $h_{opt}^{one} = h_{opt}^{one}(Q)$. This completes the proof. ■

Thus far, we have shown the existence of an optimal policy $\pi_1^{opt}$ that has two important properties:

- $\pi_1^{opt}$ is a deterministic optimal policy that decides the waiting period at $D_j$ solely based on the $Q - D_j$ and $\Delta(D_j)$.
- The first request point of the policy $\pi_1^{opt}$ is in the interval $[Q - 3B_U, Q - B_U]$. These two properties enable us to transform the optimal update policy of the single query problem into an optimal update policy of the PoW problem under periodic query arrivals.

B. Periodic Sequence of Queries

In this subsection and next subsection, we assume that the query arrival process $\{Q_k\}_{k=1}^{\infty}$ is deterministic and periodic with $T$. Let $Q_k = kT$, for $k = 1, 2, \ldots$ Furthermore, we assume that $T > 4B_U$. Based on these assumptions, we construct an optimal update policy $\pi^{opt}$ for a periodic sequence of queries in the next proposition. Then, we point out the properties of the update policy $\pi^{opt}$ based on the next proposition.

Proposition 4: If the transmission delay process $\{Y_j\}_{j=0}^{\infty}$ is i.i.d. such that $\Pr(Y_j \in [B_U, B_U]) = 1$ and the query arrival process $\{Q_k\}_{k=1}^{\infty}$ is deterministic and periodic with $T > 4B_U$, then $h_{opt}^{one}$ is equal to $h_{opt}^{one}$.

Proof: It is clear that $h_{opt}^{one} \leq h_{opt}$. Otherwise, it would contradict the optimal solution of the single query problem. Therefore, it is enough to construct an update policy $\pi^{opt}$ achieving $h_{opt}^{one}$ of expected age penalty for the periodic sequence of queries. Let $\pi_1^{opt}$ be the optimal policy of the single query problem characterized in Corollary 2. Let $Q^{BP}$ be the border point of $Q$ and $\pi_1^{opt}$. From the starting point, $\pi^{opt}$ can follow $\pi_1^{opt}$ between $[0, Q_1]$. This can be performed because $\pi_1^{opt}$ decides to wait until $Q^{BP}$ and $Q^{BP} \geq Q - 3B_U > 0$. From Corollary 3, the expected age penalty at $Q_1$ is $G_{R_1}^{opt}(Q_1 - Q^{BP})$. As the policy $\pi^{opt}$ follows $\pi_1^{opt}$ until the point $Q_1$, the channel must be idle before $Q_1 + B_U$ as the transmission delay can be at most $B_U$. When the channel is idle, $\pi^{opt}$ can follow $\pi_1^{opt}$ again, but this time the policy is performed for the query $Q_2$. The act of following the policy $\pi_1^{opt}$ is possible because $Q_2^{BP} \geq Q_2 - 3B_U > Q_1 + B_U$. Hence, the expected age penalty at $Q_2$ is $G_{R_1}^{opt}(Q_2 - Q^{BP})$ by Corollary 3. For the remaining queries $Q_3, Q_4, \ldots$, it can be replicated similar to $Q_2$. Then, the expected age penalty at every query $Q_k$ is $G_{R_1}^{opt}(Q_k - Q^{BP})$. From Corollary 1, all of the expected age penalties are equal to $h_{opt}^{one}$.

The previous proposition allows us to decouple the immediate next query from the set of all the queries while constructing an optimal policy $\pi^{opt}$ for the PoW problem under periodic query arrivals. As a result, the update policy $\pi^{opt}$ takes only the immediate next query into account. This decoupling property enables us to solve the PoW problem without a discount factor. The next corollary presents another result of the decoupling property.

Corollary 4: Let $A_j \triangleq Q - T \lfloor \frac{D_j}{T} \rfloor$ that represents the remaining time until the next query at a delivery point $D_j$. The update policy $\pi^{opt}$ constructed in Proposition 4 is a stationary and deterministic policy, which is a function of $\Delta(D_j) = Y_j$ and $A_j$.

Proof: The update policy $\pi^{opt}$ is a repetitive employment of the update policy $\pi_1^{opt}$, that is characterized in Corollary 2. Therefore, $\pi^{opt}$ possesses all the properties of $\pi_1^{opt}$. As $\pi_1^{opt}$ is solely determined based on $Q - D_j$ and $\Delta(D_j)$ by Proposition 1, $\pi^{opt}$ is stationary and deterministic function of $A_j$ and $\Delta(D_j) = Y_j$.

Note that we prove in Corollary 4 that the constructed policy $\pi^{opt}$ is a stationary and deterministic policy, which is a function of $A_j$ and $Y_j$. We also show in Proposition 4 that the optimal update policy for the PoW problem under periodic query arrivals turns out to myopic in the sense that at any delivery point, the decision about the optimal waiting time does not depend on future queries other than the immediate next one. Therefore, what remains to solve the PoW problem is to find an optimal policy for the single query problem, and apply it at each consecutive query interval.

C. Explicit Solution of PoW Problem

In the previous subsection, we exploited the decoupling property Proposition 4 to show that one can construct a solution of the PoW problem under periodic query arrivals through employing a sequence of deterministic policies that
each solve the single query problem. In this subsection, we provide an explicit solution of the single query problem by generating a sequence of update policies that are solutions of stochastic shortest path problems with finite state and action spaces obtained by quantization. Then, we show that the sequence of update policies converges to an optimal policy of the single query problem with increasingly fine quantization. The quantization argument is given next. We divide the real line interval [0, Q] into N equal sub-intervals, and define two new transmission delay processes: 1) Upper Quantized Transmission Delay Process: If a transmission delay \( Y_j \) occurs with a probability in a transmission delay process, the transmission delay is quantized to \( \frac{mQ}{N} \) with the same probability in its upper quantized transmission delay process. In other words, for every \( m \in \mathbb{N} \), we have the following:

\[
\Pr(Y_j^{upp} = m\frac{Q}{N}) = \Pr(Y_j \in \left( (m-1)\frac{Q}{N}, m\frac{Q}{N} \right])
\]

2) Lower Quantized Transmission Delay Process: If a transmission delay \( Y_j \) occurs with a probability in a transmission delay process, the transmission delay is quantized to \( \frac{mQ}{N} \) with the same probability in its lower quantized transmission delay process. In other words, for every \( m \in \mathbb{N} \), we have the following:

\[
\Pr(Y_j^{low} = (m-1)\frac{Q}{N}) = \Pr(Y_j \in \left( (m-1)\frac{Q}{N}, m\frac{Q}{N} \right])
\]

Even though the transmission delays are quantized, an optimal policy can determine waiting periods in the real interval [0, M]. Hence, the state space is still an uncountable set. The next proposition allows us to restrict the state space to a finite set.

**Proposition 5:** When a quantization on the transmission delay is performed for any number of sub-intervals \( N \), there exists an optimal update policy whose request points are in the set \( \{0, \frac{Q}{N}, \frac{2Q}{N}, \ldots, Q \} \).

**Proof:** The proof is provided in Appendix C.

The state and action spaces for lower and upper quantizations of a transmission delay process becomes finite because the ages at the delivery points are quantized and the possible delivery points form a finite set as a result of Proposition 5. Then, we can define the spaces of \( A_j, Y_j \), and \( Z_j \) as follows:

**Definition 3:** For a given query \( Q \), let us define the following sets:

- \( A^N \triangleq \{0, \frac{Q}{N}, \frac{2Q}{N}, \ldots, Q \} \)
- \( Z^N \triangleq \{0, \frac{Q}{N}, \frac{2Q}{N}, \ldots, \frac{M\cdot Q}{N} \} \)
- \( Y^N \triangleq \{\frac{Q}{N}, \frac{2Q}{N}, \ldots, \frac{(M+1)Q}{N}, \frac{M\cdot Q}{N} \} \)

Up to now, we have only analyzed the optimal update policy for quantized transmission delays. The next proposition puts an upper and a lower bound to the optimal expected age penalty for an unquantized transmission delay process. Furthermore, it proposes an update policy whose expected age penalty lays between the upper and lower bounds with the help of characterization in Section IV-A.

**Proposition 6:** For any given transmission delay process and the number of sub-intervals \( N \), the following hold:

1) There exists an update policy for an unquantized transmission delay process whose expected age penalty is less than or equal to the optimal age penalty for the upper quantized transmission delay process.

2) The optimal expected age penalty for lower quantization of a transmission delay process is less than or equal to the optimal expected age penalty for the unquantized transmission delay process.

**Proof:** For the proof of (i), we construct an update policy for an unquantized transmission delay process whose expected age penalty is less than or equal to the optimal expected age penalty for the upper quantized transmission delay process. There exists an optimal update policy for the upper quantized transmission delay process by Proposition 1. Let \( \pi_1^{opt} \) be a deterministic optimal policy that is characterized in Corollary 2. Let \( z^{opt}(\cdot, \cdot) \) be the decision function of the update policy \( \pi_1^{opt} \). The constructed optimal policy determines \( R_{j+1} \) for \( j \geq 0 \) as the following:

\[
R_{j+1} = Q - \frac{A_j}{Q/N} - z^{opt}(Q, (\frac{\Delta(Y_j)}{Q/N}, \frac{A_j}{Q/N}))
\]

Now, let us prove that this constructed policy gives the desired expected age penalty. Let \( (Y_j)_{j=1}^{\infty} \) where \( J \) is an arbitrary natural number be a transmission delay sequence from the unquantized transmission delay process when an update packet at the border point is requested. The correspondence of the transmission delay sequence on the upper quantized transmission delay process is \( (\frac{Q}{N} + Y_j)_{j=1}^{\infty} \). If the constructed policy follows the steps above, then the request points are the same for \( (Y_j)_{j=1}^{\infty} \) and \( (\frac{Q}{N} + Y_j)_{j=1}^{\infty} \). Thus, for any \( D_j \) where \( 1 \leq j \leq J \), the AoI in the interval \( [Q - D_{j}, Q + (\frac{Q - D_{j}}{Q/N})] \) is smaller for the unquantized transmission delay process. For every point outside this interval, the AoI will be the same for both of the transmission delay processes. This is valid for every transmission delay sequence \( (Y_j)_{j=1}^{\infty} \), hence the expected age penalty for the unquantized transmission delay process is less than or equal to the optimal expected age penalty for the upper quantized transmission delay process.

The proof of (ii) is similar to the previous part. Let \( \pi_1^{opt} \) be a deterministic optimal policy for the unquantized transmission delay process. By Proposition 1, \( \pi_1^{opt} \) can find the optimal waiting period for every \( \Delta(D_{j}) \) and \( Q - D_{j} \). If the destination nodes follow the same update policy \( \pi_1^{opt} \), for the lower quantized transmission delay process, the obtained expected age penalty is less than or equal to the optimal expected age penalty for the unquantized transmission delay process. This completes the proof.

The optimal update policy for the upper quantization of a transmission delay process enables us to construct an update policy for the transmission delay process. The expected age penalty resulting from this constructed update policy is proved to lay between the optimal expected age penalties of the upper and lower quantized transmission delay processes. Furthermore, we show in the next proposition that the upper and lower bounds converge to each other as \( N \) increases. Thus, we can find an update policy whose expected age penalty is
arbitrarily close to the optimal expected age penalty for any transmission delay process and age penalty function.

Proposition 7: For $\epsilon > 0$, there exists $N_1 \in \mathbb{N}$ such that the difference between optimal expected age penalties of upper and lower quantized transmission delay processes is less than $\epsilon$ if the quantization is performed with $N \geq N_1$ sub-intervals.

Proof: The proof is provided in Appendix D.

Propositions 6 and 7 employ optimal solutions of the upper and lower quantized transmission delay processes while constructing an update policy for the unquantized transmission delay process. Hence, the remaining part of this subsection is to solve the stochastic shortest path problem for quantized transmission delay processes. When the transmission delay process is quantized, the problem turns out to be a stochastic shortest path problem with finite state and action spaces as a result of Proposition 5. This problem class can be solved by the value iteration method given the explicit cost of each action in each state [58], To provide an explicit cost of each action in each state, we again use the function $G$. To provide an explicit cost of each action in each state [58], this problem class can be solved by the value iteration method given the explicit cost of each action in each state [58]. To provide an explicit cost of each action in each state, we again use the function $G$. We prove in Proposition 1 that there exists a deterministic policy $\pi_{opt} = z(Y_j, A_j)$ that is optimal for a given transmission delay process $\{Y_j\}$. Then, the following can be obtained by incorporating (9) into (8):

$$
G_{\pi_{opt}}^{opt}(Q - R_j, \Delta(R_j)) = \mathbb{E}[G_{\pi_{opt}}^{opt}(Q - R_j - Y_j - Z_j, \Delta(R_j)) | Y_j + Z_j \leq Q - R_j] \\
\times \Pr(Y_j + Z_j \leq Q - R_j) \\
+ g(Q - R_j + \Delta(R_j)) \times \Pr(Y_j + Z_j > Q - R_j)
$$

(17)

where $Z_j = z(Y_j, A_j)$.

The single query problem is explicitly solved in Algorithm 1. In this algorithm, the functions $upperG_{\pi_{opt}}^{opt}$ and $lowerG_{\pi_{opt}}^{opt}$ denote the expected age penalties for the upper and lower quantized transmission delays, respectively. These functions are recursively calculated by using (17) similar to the value iteration method. This calculation is performed through the loop in $A^N$ with ascending order. The optimal waiting time for a pair $(Y_j, A_j) \in Y^N \times A^N$ is determined by minimizing the function $G_{\pi_{opt}}^{opt}$ in the set $Z^N$. Note that the set $A^N$, $Y^N$, and $Z^N$ is employed in the algorithm as if they are arrays. The output of Algorithm 1 is a decision function of $Y_j$ and $A_j$ that characterizes an optimal update policy of the single query problem for the upper quantized transmission delay process. An optimal policy of the single query problem for the unquantized transmission delay process is constructed by an optimal update policy for the upper quantized transmission delay process as it is shown in Proposition 6(2). Then, the constructed update policy is applied to each consecutive query interval, which is optimal for the PoW problem under periodic query arrivals.

V. NUMERICAL RESULTS

Throughout the section, we exhibit the behavior of the average age penalties for the PoW and UoW problems under different transmission delay processes. To be consistent with our system model which assumes finite valued transmission delay, we will utilize truncated versions of certain transmission delay distributions such as exponential and log-normal distributions. Specifically, we truncate the values to start at 0.01 and go up to a maximum value chosen such that the cumulative distribution of the transmission delay at this value is 0.95. We choose $T = 4B_j$.

We compare three different update policies: the zero-wait policy, the optimal policy of the UoW problem found in [11], and the optimal policy of the PoW problem found in Algorithm 1. The optimal solutions of the UoW problem and the PoW problem are referred to as UoW-optimal policy and PoW-optimal policy, respectively. The average age penalty of the PoW-optimal policy is calculated by averaging the age penalties at the query instants. The average age penalties of the zero-wait policy and UoW-optimal policy are calculated as time average age penalties. Perhaps surprisingly, in all of our simulations, the time-average AoI and QAoI are identical for the zero-wait and UoW-optimal policies. The reason is, in all of our examples $X_j = Y_j + Z_j$ obeys the “Case I i.i.d.” random variable definition in [59]. Case I random variables are all the random variables except the cases that there exists $\beta \in \mathbb{R}$ such that $\Pr(X_j \in \{k\beta; k \in \mathbb{N}\}) = 1$ or $\Pr(X_j = 0) = 1$. The proof the equivalence of the time-average AoI and QAoI is subject to our future works. Note that the random variable $X_j$ under the PoW-optimal policy may not be an i.i.d. random variable, that is why the PoW-optimal policy can result in

Algorithm 1 Solution of the Single Query Problem

1: given tolerance $\epsilon$ and sufficiently large $N$
2: repeat
3: for $i = 1$ to length($A^N$) do
4: for $j = 1$ to length($Y^N$) do
5: for $k = 1$ to length($Z^N$) do
6: Calculate
7: $upperG_{\pi_{opt}}^{opt}(A^N(i) - Z^N(k), Y^N(j) + Z^N(k))$
8: $lowerG_{\pi_{opt}}^{opt}(A^N(i) - Z^N(k), Y^N(j) + Z^N(k))$
9: by using (17)
10: end for
11: end for
12: $N = 2N$
13: until $upperG_{\pi_{opt}}^{opt}(Q) - lowerG_{\pi_{opt}}^{opt}(Q) < \epsilon$
14: return $upperG_{\pi_{opt}}^{opt}(Q)$
in a lower age than the time-average age of the UoW-optimal policy.

Figures 4 and 5 illustrate the behavior of the average ages under i.i.d. beta distributed service times, i.i.d. truncated log-normal distributed service times, i.i.d. truncated Pareto distributed service times, respectively. When $\alpha = \beta = 1$, the Beta distribution becomes a uniform distribution between 0 and 1. As $\alpha = \beta$ approaches 0, it approaches a bimodal distribution concentrated around 0 and 1 with probability close to 0.5 each. Interestingly, as $\alpha$ and $\beta$ increase, the average ages of the zero-wait policy and the UoW-optimal policy decrease whereas the average age of the PoW-optimal policy increases even though the mean of the beta distribution is constant, $\frac{\alpha}{\alpha + \beta} = \frac{1}{2}$. The benefit of using the PoW-optimal policy is pronounced when the transmission delay is bi-modal distributed. The log-normal distribution is a heavy-tailed distribution, especially for large $\sigma$. As $\alpha$ goes to $\infty$, the Pareto distribution converges to the Dirac delta function $\delta(t - x_m)$, similar to the example 1 in Section I. We choose $x_m = 1$ which leads that UoW-optimal policy is equivalent to the zero wait policy for $\alpha \geq 3$ [11, Th. 5]. We observe in Figure 5 that PoW-optimal policy performs well as the transmission delay distribution approaches the Dirac delta function.

Figure 6 exhibits the behavior of the average age penalties for different $\alpha$ when the age penalty function $g(t) = e^{\lambda t} - 1$ and i.i.d. truncated exponential distributed service times where $\lambda = 1$. This nonlinear age penalty function represents destination nodes that demand very fresh update packets and harshly penalize stale update packets. In the figure, we observe that the PoW-optimal policy works much better than the other policies especially for high $\alpha$ values. It means that the pull-based communication model is beneficial to utilize when the destination node demands very fresh update packets.

Up to now, we have not put any constraint on the number of transmissions for the policies. Figures 7 and 8 illustrate the behavior of the average ages under truncated i.i.d. exponential distributed service times and Pareto distributed service times, respectively, when the number of transmissions in the UoW-optimal policy is constrained by the number of transmissions made by the PoW-optimal policy. We observe that the average age of the PoW-optimal policy is much lower than the average age of the UoW-optimal policy for an equal number of transmissions. This implies that in a practical situation, applying the PoW solution can be significantly more energy-efficient, for the same age performance.
VI. CONCLUSION AND FUTURE DIRECTIONS

We studied the optimal control of a status update system in which the destination node requests the source node to submit an update packet to a channel, which has delay. We defined a continuous, non-decreasing, and non-negative penalty function to represent the level of dissatisfaction on data staleness. We defined the PoW problem and first solved it for a single query, which we showed to be a stochastic shortest path problem with an uncountable state and action spaces. For this SSP problem, we devised an optimal policy for the PoW problem and first solved it for a single query, which we showed to be a stochastic shortest path problem with a possibly continuous function as well.

APPENDIX A

PROOF OF PROPOSITION 1

We first prove that \( Q - D_j \) and \( \Delta(D_j) \) are sufficient statistics to obtain an optimal \( Z_j \) for each \( j \), \((Y_j)_{i=0}^{j-1} \), and \((Z_j)_{i=0}^{j-1} \). We perform induction on \( Q - D_j \). Let us map each \( Q - D_j \) to a natural number \( n \) such that \( (n - 1)B_L \leq Q - D_j < nB_L \). If \( n = 1 \), then \( Q - D_j \leq B_L \). For every waiting period \( Z_j \), the age penalty at the query is constant because a new update cannot arrive until the query. Then, the age penalty at the query is:

\[
g(Q - D_j + \Delta(D_j)) \]

Thus, if \( n = 1 \), \( Q - D_j \) and \( \Delta(D_j) \) are sufficient statistics to obtain an optimal \( Z_j \) for each \( j \), \((Y_j)_{i=0}^{j-1} \), and \((Z_j)_{i=0}^{j-1} \). Let us assume that \( Q - D_j \) and \( \Delta(D_j) \) are sufficient statistics to obtain an optimal \( Z_j \) for \( n = 2, 3, \ldots K \) where \( K \) is an arbitrary natural number. Let \( \Pi_K \) be the set of all causal waiting policies such that if \( \pi \in \Pi_K \), then \( \pi \) determines waiting times at delivery points \( D_j; Q - D_j < KB_L \) solely based on \( Q - D_j \) and \( \Delta(D_j) \), for the delivery points \( D_j; Q - D_j \geq KB_L \), the waiting policy may not determine the waiting time based on \( Q - D_j \) and \( \Delta(D_j) \). Due to the induction assumption, the single query problem can be minimized in the set of \( \Pi_K \). Let us prove that the single query problem can be minimized in the set of \( \Pi_{K+1} \) as well. For every \( \pi \in \Pi_K \), we can obtain the following:

\[
G^\pi_R(Q - D_j - Z_j, \Delta(D_j) + Z_j, (Y_i)_{i=0}^{j}, (Z_i)_{i=0}^{j}) \]

\[
= E[G^\pi_D(Q - D_j - Z_j - Y_{j+1}, Y_{j+1}, (Y_i)_{i=0}^{j+1}, (Z_i)_{i=0}^{j+1})]
\]

\[
\times Pr(Y_{j+1} + Z_j \leq Q - D_j)
\]

\[
+ g(Q - D_j + \Delta(D_j)) \times Pr(Y_{j+1} + Z_j > Q - D_j)
\]

(18)

where (a) follows from (8). \( Q - D_{j+1} = Q - D_j - Z_j - Y_{j+1} < KB_L \) as \( Y_{j+1} \geq B_L \) and \( Q - D_j < (K + 1)B_L \). This means that we can exploit the induction assumption in the RHS of (18) to claim that \((Y_i)_{i=0}^{j+1} \) and \((Z_i)_{i=0}^{j+1} \) do not affect the value of the term with expectation given \( Q - D_{j+1} = Q - D_j - Y_{j+1} - Z_j \) and \( \Delta(D_{j+1}) = Y_{j+1} \). This is because \( \pi \in \Pi_K \). In the term with penalty function, only \( Q - D_j \) and \( \Delta(D_j) \) appear. This means that the optimal control problem of choosing an optimal \( Z_j \) at the delivery point \( D_j \) does not depend on \((Y_i)_{i=0}^{j+1} \) and \((Z_i)_{i=0}^{j+1} \). This completes the induction. Once the single query problem can be minimized in the set of \( \bigcup_{k=1}^{\infty} \Pi_k \), it is easy to show that the calculation of the functions \( G^\pi_D \) and \( G^\pi_R \) can be performed by only knowing \( Q - D_j \) and \( \Delta(D_j) \) for every \( \pi \in \bigcup_{k=1}^{\infty} \Pi_k \). The proof can be performed with a similar induction. From now on, we can omit \((Y_i)_{i=0}^{j+1} \) and \((Z_i)_{i=0}^{j+1} \) from \( G^\pi_R \) and \( G^\pi_D \). For the part related to the existence of a deterministic optimal policy, we construct a deterministic optimal policy by performing another induction on \( Q - D_j \). Before move on to the induction, we state some simple observation.

**Lemma 2:** Let us assume that there exists a deterministic optimal policy \( \pi^o_p \).

1. Let \( h: \mathbb{R} \rightarrow \mathbb{R} \) such that \( h(\epsilon) \triangleq \max_{x \in [0, M + B_L]} g(x + \epsilon) - g(x) \). Then, we can obtain the following for every \( t_1, t_2 \in \mathbb{R} \)

\[
0 \leq G^\pi_R(t_1, t_2 + \epsilon) - G^\pi_R(t_1, t_2) \leq h(\epsilon)
\]

(19)

2. If \( f(x) \triangleq G^\pi_R(Q - D_j - x, \Delta(D_j)) \) is a lower semi-continuous function for a given \( Q - D_j \) and \( \Delta(D_j) \), then \( f(x) \triangleq G^\pi_R(Q - D_j - x, \Delta(D_j)) + x \) is a lower semi-continuous function as well.

3. If \( f(x) \triangleq G^\pi_R(Q - D_j - x, \Delta(D_j)) \) is a lower semi-continuous function for every \( Q - D_j \) and \( \Delta(D_j) \) satisfying \( Q - D_j < C \), where \( C \) is an arbitrary real number, then \( f''(x) \triangleq G^\pi_R(Q - D_j - Y_{j+1} - x, Y_{j+1}) \) is a...
lower semi-continuous function as well for every $Q - D_i$ and $\Delta(D_i)$ satisfying $Q - D_i < C$.

(iv) For every $\epsilon > 0$, there exists $\delta > 0$ such that

$$\Pr(t_1 < Y_j < t_1 + \delta) < \epsilon$$ (20)

where $t_1$ is a given real number satisfying $t_1 \in [B_L, B_U]$. 

For every $Q - D_j$, $\Delta(D_j)$, and $\epsilon > 0$, there exists $\delta > 0$ such that

$$G_{D_j}^{\text{opt}}(Q - D_j - x, \Delta(D_j)) < \epsilon$$ (21)

Proof:

(i) It follows from (8) and the facts that the penalty function $g$ is continuous and non-decreasing.

(ii) It follows from the definition of lower semi-continuity and Lemma 2(i).

(iii) It follows from Lemma 2(ii) and the fact that $\pi_1^{\text{opt}}$ is a deterministic optimal policy.

(iv) The transmission delay is measurable on Borel algebra on the real line.

(v) It follows from (9) and the fact that the penalty function $g$ is continuous and non-decreasing.

The idea which will be proven by the induction is that $f(x) \triangleq G_R^{\text{opt}}(Q - D_j - x, \Delta(D_j))$ is a lower semi-continuous function. From Lemma 2(ii), $f(x) \triangleq G_R^{\text{opt}}(Q - D_j - x, \Delta(D_j) + x)$ is a lower semi-continuous function as well. Therefore, it attains its infimum for every $Q - D_j$ and $\Delta(D_j)$ due to the extension of Extreme Value Theorem to semi-continuity. Then, this infimum point can be determined as the waiting time at the delivery point $D_j$. This policy is a deterministic optimal policy that decides the waiting periods solely based on $Q - D_j$ and $\Delta(D_j)$. Now, let us move on to the induction. When $Q - D_j < B_L$, all waiting periods result in the same age penalty. This means that there exists a deterministic optimal policy $\pi_1^{\text{opt}}$ for $n = 1$. Additionally, $f(x) \triangleq G_R^{\text{opt}}(Q - D_j - x, \Delta(D_j))$ is lower semi-continuous for every $D_j$ satisfying $Q - D_j < B_L$. Let us assume for $n = 2$ that $f(x) \triangleq G_R^{\text{opt}}(Q - D_j - x, \Delta(D_j))$ is lower semi-continuous for every $D_j$ satisfying $B_L \leq Q - D_j < 2B_L$. Note that the superscript $\pi_1^{\text{opt}}$ refers in the definition of the function $f$ that the deterministic optimal policy $\pi_1^{\text{opt}}$ is performed starting with $(j + 2)^\text{th}$ request because $(j + 1)^\text{th}$ request has already determined as $D_j + x$. The delivery point $D_{j+1}$ must satisfy $Q - D_{j+1} < (2 - 1)B_L$ in which there exists a deterministic optimal policy. As $f(x) \triangleq G_R^{\text{opt}}(Q - D_j - x, \Delta(D_j))$ is lower semi-continuous for $n = 2$, $f(x) \triangleq G_R^{\text{opt}}(Q - D_j - x, \Delta(D_j) + x)$ is a lower semi-continuous function as well by Lemma 2(ii). Hence, the function $f(x)$ attains its minimum for every $B_L \leq Q - D_j < 2B_L$ and $\Delta(D_j)$. Therefore, there exists a deterministic optimal policy for $n = 2$ as well. Similar to the transition from $n = 1$ to $n = 2$, let us assume by one that the function $f(x) \triangleq G_R^{\text{opt}}(Q - D_j - x, \Delta(D_j))$ is lower semi-continuous and there exists a deterministic optimal policy for $n = 2, 3, \ldots, K$ where $K$ is an arbitrary natural number. Let us prove that $f(x) \triangleq G_R^{\text{opt}}(Q - D_j - x, \Delta(D_j))$ is a lower semi-continuous function for $n = K + 1$. To reach contradiction, suppose that the claim is false. Then, there exists $Q - D_j$, $\Delta(D_j)$, and $x_0$ satisfying $KB_L \leq Q - D_j < (K + 1)B_L$ such that $f(x) \triangleq G_R^{\text{opt}}(Q - D_j - x, \Delta(D_j))$ is not lower semi-continuous at $x_0$. Hence, there exist either an increasing or a decreasing sequence $(x_n)$ and $C > 0$ such that $\lim_{n \to \infty} x_n = x_0$ and $f(x_n) - f(x_0) < -C$ for every $n \in \mathbb{N}$. If $(x_n)$ is a increasing sequence, then we obtain the following by (8):

$$f(x_n) - f(x_0) = A \times \Pr(Y_{j+1} \leq Q - D_j - x_0) + B \times \Pr(Q - D_j - x_n < Y_{j+1} \leq Q - D_j - x_n) + C \times \Pr(Q - D_j - x_n < Y_{j+1} < Q - D_j - x_n)$$

where $A, B, C$ are the following:

$$A = E\left[G_R^{\text{opt}}(Q - D_j - x_n - Y_{j+1}, Y_{j+1}) \left| Y_{j+1} \leq Q - D_j - x_n \right. \right] - E\left[G_R^{\text{opt}}(Q - D_j - x_n, Y_{j+1}) \left| Y_{j+1} < Q - D_j - x_n \right. \right]$$

$$B = g(Q - D_j + \Delta(D_j)) - E\left[G_R^{\text{opt}}(Q - D_j - x_n, \Delta(D_j)) - g(Q - D_j - x_n + \Delta(D_j)) \right]$$

From the induction assumption and Lemma 2(iii), $A$ can be arbitrarily small. $B$ is upper bounded by $g(M + B_U)$ and the multipliers of $B$ in (22) can be arbitrarily small by Lemma 2(iv). $C$ can be arbitrarily small due to the continuity of the penalty function $g$. Therefore, there exists $x_0$ such that $f(x_n) - f(x) \geq -C$, which is a contradiction. If $(x_n)$ is a decreasing sequence, then an equation similar to (22) can be written. The terms that are similar to $A$ and $C$ can be analyzed similarly. The term that is similar $B$ can be analyzed with the help of Lemma 2(v). After the analysis, a similar contradiction can be achieved. As a result, the function $f(x)$ is lower semi-continuous. From Lemma 2(ii), the function $f'(x)$ is lower semi-continuous for every $Q - D_j$ and $\Delta(D_j)$ satisfying $KB_L \leq Q - D_j < (K + 1)B_L$. Thus, the function attains its infimum, and the infimum point can be determined as a deterministic optimal waiting period $Z_j$, which completes the induction.

APPENDIX B

PROOF OF PROPOSITION 3

We start this proof with a lemma:

**Lemma 3:** For any delivery point $D_j \in [0, Q - 2B_U]$ and its AoI $\Delta(D_j)$, an optimal request point $R_{j+1}$ must be until $Q - B_U$ i.e. $R_{j+1} \leq Q - B_U$.

**Proof:** Let us assume that this lemma is not true: There exist a delivery point $D_j \in [0, Q - 2B_U]$ and its AoI at the delivery $\Delta(D_j)$ such that an optimal request point is $R_{j+1} > Q - B_U$. Let this policy follows $\pi^{\text{opt}}$ and let $R^* \triangleq R_{j+1}$. We will show that there exists $\pi^{\text{modified}}$ such that $G_{D_j}^{\text{opt}}(Q - D_j, \Delta(D_j)) \leq G_{D_j}^{\text{modified}}(Q - D_j, \Delta(D_j))$. Let $\pi^{\text{modified}}$ determine $R_{j+1}^{\text{modified}} = D_j$ and $R_{j+2}^{\text{modified}} = R^*$. As the time duration between $D_j$ and $R^*$ is greater than $B_U$, $\pi^{\text{modified}}$ can determine $R_{j+2}^{\text{modified}}$ as $R^*$.
Next, we prove that there is no point $D_j$ is the optimal policy. Hence, there is no such point $R_j$ satisfying $K_{BL} \leq Q - 3B_U - R_j < (K + 1)B_L$. As a result of Lemma 3, $R_j$ cannot be in the interval $[0, B_L]$. This completes the proof.

As a result of Lemma 3, $R^*_1 \leq Q - B_U$. From Proposition 2, AoI at $R^*_1$ does not affect the age penalty at the query. Next, we prove that there is no $R_j \in [0, Q - B_U]$ such that $G_R^\text{opt}(Q - R^*_1, \Delta(D_j)) > G_R^\text{opt}(Q - R_j, \Delta(D_j))$. If there existed such a point $R_j$ such that $R_j \in [Q - 3B_U, Q - B_U]$, then the destination node would determine the optimal request point for the delivery point $D_j$ as $R^*_1$. Hence, we can state the following for every delivery point $D_j$ and its transmission delay $Y_j$ satisfying $D_j \in [Q - 3B_U, Q - 2B_U]$ and $Y_j \in [B_L, B_U]$:

$$G_R^\text{opt}(Q - R^*_1) \leq G_D^\text{opt}(Q - D_j, Y_j)$$

On the other hand, such $R^*_1$ cannot be in the interval $[0, Q - 3B_U]$ as well. This statement is proved by induction. Similar to the proof of Proposition 1, $R^*_1$ is mapped to a natural number $n$ if it satisfies $(n - 1)B_L \leq Q - 3B_U - R^*_1 < nB_L$. It is true for $n = 1$ i.e. such $R^*_1$ cannot be in the interval $0 \leq Q - 3B_U - R^*_1 < B_L$ because of the following:

$$G_R^\text{opt}(Q - R^*_1) \leq G_D^\text{opt}(Q - R^*_1, Y_j)$$

where (a) follows from (8), and (b) follows from (28). Let us assume that the induction statement is true for $n = 2, 3, \ldots, K$ where $K$ is an arbitrary natural number. This statement assumes the following for every request point $R^*_1$ satisfying $0 \leq Q - 3B_U - R^*_1 < KB_L$:

$$G_R^\text{opt}(Q - R^*_1) \leq G_D^\text{opt}(Q - R^*_1)$$

Let us prove the induction statement for $n = K + 1$. For every $R_j$ satisfying $K_{BL} \leq Q - 3B_U - R_j < (K + 1)B_L$, we have the following:

$$G_R^\text{opt}(Q - R_j) \leq G_D^\text{opt}(Q - R_{j+1} - Y_j, Y_{j+1})$$

where (a) follows from (8), and (b) follows from (30). This implies that the induction is completed. As a result, for every $R_{j+1} \in [0, Q - B_U]$, we have the following:

$$G_R^\text{opt}(Q - R_{j+1}) \leq G_R^\text{opt}(Q - R^*_1)$$

It means that $R^*_1$ attains its infimum value on the interval $[0, Q - B_U]$. This completes the proof.

APPENDIX C

PROOF OF PROPOSITION 5

We perform a similar induction included in the proof of Proposition 1. Let us map each $Q - D_j$ to a natural number $n$ that satisfies $(n - 1)B_L \leq Q - D_j < nB_L$. If $n = 1$, the request point does not affect the expected age penalty at the query. Thus, requesting at the query is an optimal request point that proves the proposition statement for $n = 1$. Let us assume that the optimal request point is in the set $\{0, N\frac{Q}{B_L}, 2N\frac{Q}{B_L}, \ldots, Q\}$ when a delivery occurs at time $D_j$ satisfying $(n - 1)B_L \leq Q - D_j < nB_L$ for $n = 1, 2, \ldots, K$ where $K$ is an arbitrary natural number. Let us prove that the optimal request point is in the set $\{0, N\frac{Q}{B_L}, 2N\frac{Q}{B_L}, \ldots, Q\}$ when a delivery occurs at time $D_j$ satisfying $K_{BL} \leq R - D_j < (K + 1)B_L$. Let us assume the inverse. There exists a delivery point $D_j$ such that $K_{BL} \leq Q - D_j < (K + 1)B_L$ and the is no optimal request point in the set $\{0, N\frac{Q}{B_L}, 2N\frac{Q}{B_L}, \ldots, Q\}$. As there exists an optimal policy from Proposition 1, there exists an optimal request point $R_{j+1} \notin \{0, N\frac{Q}{B_L}, 2N\frac{Q}{B_L}, \ldots, Q\}$. For every quantized transmission delay $Y_{j+1}$, the next delivery point satisfies $Q - D_{j+1} < KB_L$. If $Q - D_{j+1} > 0$, the optimal next request point should be in the set $\{0, N\frac{Q}{B_L}, 2N\frac{Q}{B_L}, \ldots, Q\}$ due to the induction assumption. Instead of requesting at $R^*_1$, if the request was performed at $R^*_1 \triangleq \frac{\lceil \frac{Q}{B_L} \rceil N}{Q}$, there would be two cases based on the transmission delay $Y_{j+1}$. For every $m \in \mathbb{N}$; if $D^\text{opt}_{j+1} > m\frac{Q}{N}$, then $D_{j+1} > m\frac{Q}{N}$, if $D^\text{opt}_{j+1} < m\frac{Q}{N}$, then $D_{j+1} > m\frac{Q}{N}$ because of the quantized transmission delay process, where $m\frac{Q}{N}$ represents the possible next request point or the query. Therefore, requesting an update packet at $R^*_1$ is optimal given that $R^*_1$ is an optimal request point. This conclusion contradicts with the assumption. Hence, there exists an optimal request point in the set $\{0, N\frac{Q}{B_L}, 2N\frac{Q}{B_L}, \ldots, Q\}$ for every delivery point, which completes the proof.

APPENDIX D

PROOF OF PROPOSITION 7

Let $\delta = \frac{1}{2} \max_{x \in [0,M+2]} (y(x) - g(x))$. Let $N_1$ be a natural number satisfying $N_1 > 3\frac{Q}{4B_L}$. There exists a deterministic optimal policy $\pi_1^\text{opt}$ whose first request point is the border...
point for lower quantization of the transmission delay with $N \geq N_1$ by Corollary 2. We construct an update policy $\pi_{1}^{mod}$ for upper quantization of the transmission delay by $N$ by utilizing $\pi_{1}^{opt}$. Let the border point corresponding to the lower quantized transmission delay and $\pi_{1}^{opt}$ be $QBP$. Let $\pi_{1}^{mod}$ pull its first request at $R_{1}^{mod} \triangleq QBP - \frac{QBP - Q_{BL}}{Q}$. Note that $R_{1}^{mod} - QBP < \delta$ as $QBP \geq Q - 3B_{U} > \frac{Q}{4}$ and $N \geq N_1$. Let $(Y_{j})_{j=1}^{\pi_{1}^{opt}}$ be an arbitrary transmission delay sequence from the unquantized transmission delay process where $\sum_{j=1}^{\pi_{1}^{opt}} Y_{j} > Q$. Let $(Y_{j})_{j=1}^{\pi_{1}^{opt}}$ and $(Y_{j})_{j=1}^{\pi_{1}^{opt}}$ be the sequences that correspond to upper and lower quantized of $(Y_{j})_{j=1}^{\pi_{1}^{opt}}$, respectively. Let $(Z_{j})_{j=1}^{\pi_{1}^{opt}}$ be the waiting time sequences that is causally determined by $\pi_{1}^{opt}$. This completes the proof.

ACKNOWLEDGMENT

The authors thank Semanur Avsar for her assistance with simulations.

REFERENCES

[1] G. Mois, S. Folea, and T. Sanislav, “Analysis of three IoT-based wireless sensors for environmental monitoring,” IEEE Trans. Instrum. Meas., vol. 66, no. 8, pp. 2056–2064, Aug. 2017.

[2] M. Hassanali and C. E. Koksal, “Health monitoring and management using internet-of-things (IoT) sensing with cloud-based processing: Opportunities and challenges,” in Proc. IEEE Int. Conf. Services Comput., 2015, pp. 285–292.

[3] H. Li, Y. Liu, Z. Qin, H. Rong, and Q. Liu, “A large-scale urban vehicular network framework for IoT in smart cities,” IEEE Access, vol. 7, pp. 74437–74449, 2019.

[4] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, “Internet of Things for smart cities,” IEEE Internet Things J., vol. 1, no. 1, pp. 22–32, Feb. 2014.

[5] S. Kaul, R. Yates, and M. Gruteser, “Real-time status: How often should one update?” in Proc. IEEE INFOCOM, 2012, pp. 2731–2735.

[6] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, “Age of information: An introduction and survey,” IEEE J. Sel. Areas Commun., vol. 39, no. 5, pp. 1183–1210, May 2021.

[7] Y. Sun, I. Kadota, R. Talak, and E. Modiano, Age of Information: A New Metric for Information Freshness. San Rafael, CA, USA: Morgan Clouprint, 2019.

[8] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, “Scheduling policies for minimizing age of information in broadcast wireless networks,” IEEE/ACM Trans. Netw., vol. 26, no. 6, pp. 2637–2650, Dec. 2018.

[9] B. T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu, “Age of information under energy replenishment constraints,” in Proc. Inf. Theory Appl. Workshop (ITA), 2015, pp. 25–31.

[10] B. T. Bacinoglu and E. Uysal-Biyikoglu, “Finite-horizon online transmission scheduling on an energy harvesting communication link with a discrete set of rates,” J. Commun. Netw., vol. 16, no. 3, pp. 300–309, Jun. 2014.

[11] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, “Update or wait: How to keep your data fresh,” IEEE Trans. Inf. Theory, vol. 63, no. 11, pp. 7492–7508, Nov. 2017.

[12] R. D. Yates, “Lazy is timely: Status updates by an energy harvesting sensor,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), 2015, pp. 3008–3012.
[38] J. Bao et al., “Towards a theory of semantic communication,” in Proc. IEEE Netw. Sci. Workshop, 2011, pp. 110–117.
[39] B. Güler, A. Yener, and A. Swami, “The semantic communication game,” IEEE Trans. Cogn. Commun. Netw., vol. 4, no. 4, pp. 787–802, Dec. 2018.
[40] P. Popovski, O. Simeone, F. Boccardi, D. Gündüz, and O. Sahin, “Semantic-effectiveness filtering and control for post-5G wireless connectivity,” J. Indian Inst. Sci., vol. 100, no. 2, pp. 435–443, Apr. 2020. [Online]. Available: https://doi.org/10.1007/s41745-020-00165-6
[41] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides, “The age of incorrect information: A new performance metric for status updates,” IEEE/ACM Trans. Netw., vol. 28, no. 5, pp. 2215–2228, Oct. 2020.
[42] Y. Chen and A. Ephremides, “Minimizing the age of incorrect information for unreliable channel with power constraint,” in Proc. IEEE Global Commun. Conf. (GLOBECOM), 2021, pp. 1–6.
[43] S. Kriouile and M. Assaad, “Minimizing the age of incorrect information for real-time tracking of Markov remote sources,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), 2021, pp. 2978–2983.
[44] C. Kam, S. Kompella, and A. Ephremides, “Age of incorrect information for remote estimation of a binary Markov source,” in Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS), 2020, pp. 1–6.
[45] X. Zheng, S. Zhou, and Z. Niu, “Urgency of information for context-aware timely status updates in remote control systems,” IEEE Trans. Wireless Commun., vol. 19, no. 11, pp. 7237–7250, Nov. 2020.
[46] W. Lin, X. Wang, C. Xu, X. Sun, and X. Chen, “Average age of changed information in the Internet of Things,” in Proc. IEEE Wireless Commun. Netw. Conf. (WCNC), 2020, pp. 1–6.
[47] Y. Sang, B. Li, and B. Ji, “The power of waiting for more than one response in minimizing the age-of-information,” in Proc. IEEE Global Commun. Conf., 2017, pp. 1–6.
[48] F. Li, Y. Sang, Z. Liu, B. Li, H. Wu, and B. Ji, “Waiting but not aging: Optimizing information freshness under the pull model,” IEEE/ACM Trans. Netw., vol. 29, no. 1, pp. 465–478, Feb. 2021.
[49] Y. Dong, Z. Chen, S. Liu, and P. Fan, “Age of information upon decisions,” in Proc. IEEE 39th Sarnoff Symp., 2018, pp. 1–5.
[50] Y. Dong and P. Fan, “Age upon decisions with general arrivals,” in Proc. 9th IEEE Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON), 2018, pp. 825–829.
[51] B. Yin et al., “Only those requested count: Proactive scheduling policies for minimizing effective age-of-information,” in Proc. IEEE Conf. Comput. Commun., 2019, pp. 109–117.
[52] M. E. Ildiz, O. T. Yavascan, E. Uysal, and O. T. Kartal, “Query age of information: Optimizing AoI at the right time,” in Proc. IEEE Int. Symp. Inform. Theory (ISIT), 2022, pp. 144–149.
[53] M. E. Ildiz, S. Aykar, and E. Uysal, “An inequality for query age of information and age of information,” in Proc. 30th Signal Process. Commun. Appl. Conf. (SIU), 2022, pp. 1–4.
[54] D. P. Bertsekas and J. N. Tsitsiklis, “An analysis of stochastic shortest path problems,” Math. Oper. Res., vol. 16, no. 3, pp. 580–595, 1991. [Online]. Available: http://www.jstor.org/stable/2032412
[55] D. P. Bertsekas, “Proper policies in infinite-state stochastic shortest path problems,” IEEE Trans. Automat. Control, vol. 63, no. 11, pp. 3787–3792, Nov. 2018.
[56] S. R. Pliska, “On the transient case for Markov decision chains with general state spaces,” in Proc. Dyn. Program. Appl., 1978, pp. 355–349. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780125681506500246
[57] H. W. James and E. J. Collins, “An analysis of transient Markov decision processes,” J. Appl. Probab., vol. 43, no. 3, pp. 603–621, 2006. [Online]. Available: http://www.jstor.org/stable/27595759
[58] D. P. Bertsekas, Dynamic Programming and Optimal Control, 3rd ed, vol. 2, Belmon, MA, USA: Athena Sci., 2005.
[59] H. Robbins, “On the Equidistribution of sums of independent random variables,” Proc. Amer. Math. Soc., vol. 4, no. 5, pp. 786–799, 1953. [Online]. Available: http://www.jstor.org/stable/2032412