

Update or Wait: How to Keep Your Data Fresh

Yin Sun[†], Elif Uysal-Biyikoglu[‡], Roy D. Yates^{*}, C. Emre Koksal[†], and Ness B. Shroff^{†§}

[†]Dept. of ECE, [§]Dept. of CSE, The Ohio State University, Columbus, OH

[‡]Dept. of EEE, Middle East Technical University, Ankara, Turkey

^{*}Dept. of ECE, Rutgers University, North Brunswick, NJ

Abstract—In this work, we study how to optimally manage the freshness of information updates sent from a source node to a destination via a channel. A proper metric for data freshness at the destination is the *age-of-information*, or simply *age*, which is defined as how old the freshest received update is since the moment that this update was generated at the source node (e.g., a sensor). A reasonable update policy is the *zero-wait* policy, i.e., the source node submits a fresh update once the previous update is delivered and the channel becomes free, which achieves the maximum throughput and the minimum delay. Surprisingly, this zero-wait policy does *not* always minimize the age. This counter-intuitive phenomenon motivates us to study how to optimally control information updates to keep the data fresh and to understand when the zero-wait policy is optimal. We introduce a general age penalty function to characterize the level of dissatisfaction on data staleness and formulate the average age penalty minimization problem as a constrained semi-Markov decision problem (SMDP) with an uncountable state space. We develop efficient algorithms to find the optimal update policy among all causal policies, and establish sufficient and necessary conditions for the optimality of the zero-wait policy. Our investigation shows that the zero-wait policy is far from the optimum if (i) the age penalty function grows quickly with respect to the age, (ii) the packet transmission times over the channel are positively correlated over time, or (iii) the packet transmission times are highly random (e.g., following a heavy-tail distribution).

I. INTRODUCTION

In recent years, the proliferation of mobile devices and applications has significantly boosted the need for real-time information updates, such as news, weather reports, traffic alerts, email notifications, stock quotes, social updates, mobile ads, etc. In addition, timely information updates are also critical in real-time monitoring and control systems, including sensor networks used in temperature and air pollution monitoring [2], surround monitoring in autonomous vehicles [3], phasor data updates in power grid stabilization systems [4], and so on.

A common need in these real-time applications is to optimize the freshness of the data. In light of this, a metric of data freshness called the *age-of-information*, or simply *age*, was defined in, e.g., [5], [6]. At any time t , if the freshest update delivered at the destination of information updates (e.g.,

a monitor or controller) was generated at time $U(t)$, then the age at the destination is

$$\Delta(t) = t - U(t). \quad (1)$$

Hence, the age is the amount of time elapsed since the moment that the freshest delivered update was generated.

Most existing research on the age-of-information focuses on an “enqueue-and-forward” model [6]–[21], where information update packets arrive stochastically at the source node. The source node enqueues these updates and then forwards them to a destination through a channel. It is worth noting that there is a significant difference between *age* and *delay*: A low update frequency results in a short queueing delay because the queue is almost empty, but the destination may end up having stale data due to infrequent updates. On the other hand, a high update frequency will increase the queueing delay, and the age is also high because the updates are becoming stale during their long waiting time in the queue. Hence, delay is an increasing function of the update frequency, but the age first decreases and then increases with respect to the update frequency [6]. In [6], [13], it was found that a good policy is to discard the old updates waiting in the queue when a new sample arrives, which can greatly reduce the negative impact of queueing delay. In [20], [21], a Last Generated First Served (LGFS) policy was shown to minimize the age process in a strong sense (in the sense of stochastic ordering) for multi-channel and multi-hop networks with an arbitrary (e.g., out-of-order) update arrival process.

In this paper, we study a “generate-at-will” model proposed in [22]: As depicted in Fig. 1, the source node (e.g., a sensor) has access to the channel’s idle/busy state through acknowledgements (ACKs), and in contrast to [6]–[21], is able to generate information updates at any time by its own will. Because of queueing, update packets may need to wait in the queue for their transmission opportunity, and become stale while waiting. Hence, it is better not to generate updates when the channel is busy, which completely eliminates the waiting time in the queue. In this case, a reasonable update policy is the *zero-wait* policy, also called *just-in-time updating* in [22] and the *work-conserving* policy in queueing theory [23], that submits a fresh update once the previous update is delivered and an ACK is received.¹ The zero-wait policy achieves the maximum throughput and the minimum delay. Surprisingly, this zero-wait policy does **not** always minimize

¹This policy was proposed in [6, Section VII] to provide a lower bound to the age in the “enqueue-and-forward” model.

This paper was presented in part at IEEE INFOCOM 2016 [1].

This work was supported in part by DTRA grant HDTRA1-14-1-0058, NSF grants CNS-1446582, CNS-1409336, CNS-1518829, CNS-1514260, CNS-1422988, and CNS-1054738, ARO grant W911NF-14-1-0368, ONR grants N00014-15-1-2166, and N00014-17-1-2417. E. Uysal-Biyikoglu was supported in part by TUBITAK and in part by a Science Academy BAGEP award.

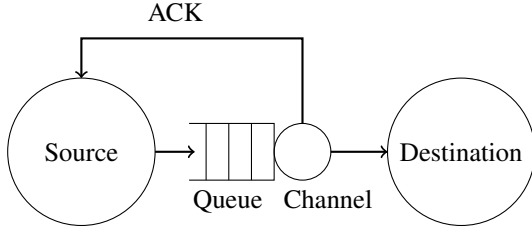


Fig. 1. System model.

the age-of-information. In particular, an optimal policy was obtained in [22] for minimizing the time-average age, from which one can deduce that the zero-wait policy is not age-optimal. The following example reveals the reason behind this counter-intuitive phenomenon:

Example: Suppose that the source node sends a sequence of information updates to a destination. The transmission times of these updates form a periodic sequence

$$0, 0, 2, 2, 0, 0, 2, 2, 0, 0, 2, 2, \dots$$

Suppose that Update 1 is generated and submitted at time 0 and delivered at time 0. Under the zero-wait policy, Update 2 is also generated at time 0 and delivered at time 0. However, despite its negligible transmission time, Update 2 has not brought any new information to the destination after Update 1 was delivered, because both updates were sampled at the same time. Therefore, the potential benefit of the zero transmission time of Update 2 is wasted! This issue occurs periodically over time: Whenever two consecutive updates have zero transmission time, the second update of the two is wasted. Therefore, a 1/4 of the updates in this sequence are wasted in the zero-wait policy!

For comparison, consider a ϵ -wait policy that waits for ϵ seconds after each update with a zero transmission time, and does not wait after each update with a transmission time of 2 seconds. Note that the control decisions in the ϵ -wait policy are made causally. The time-evolution of the age $\Delta(t)$ in the ϵ -wait policy is shown in Fig. 2. Update 1 is generated and delivered at time 0. Update 2 is generated and delivered at time ϵ . Update 3 is generated at time 2ϵ and is delivered at time $2 + 2\epsilon$. Because the transmission time of Update 3 is 2 seconds, the latest delivered update at time $2 + 2\epsilon$ is of the age 2 seconds. Hence, the age $\Delta(t)$ drops to 2 seconds at time $2 + 2\epsilon$. Update 4 is generated at time $2 + 2\epsilon$ and is delivered at time $4 + 2\epsilon$. At time $4 + 2\epsilon$, the age drops to zero because Update 5 is generated at this time and is delivered immediately.

According to Fig. 2, one can compute the time-average age of the ϵ -wait policy, which is given by

$$\begin{aligned} & (\epsilon^2/2 + \epsilon^2/2 + 2\epsilon + 4^2/2)/(4 + 2\epsilon) \\ & = (\epsilon^2 + 2\epsilon + 8)/(4 + 2\epsilon) \text{ seconds.} \end{aligned}$$

If the waiting time is $\epsilon = 0.5$, the time-average age of the ϵ -wait policy is 1.85 seconds. If the waiting time is $\epsilon = 0$, it reduces to the zero-wait policy, whose time-average age is 2 seconds. Hence, the zero-wait policy is not optimal! In Section V, we will see more numerical results showing that the zero-

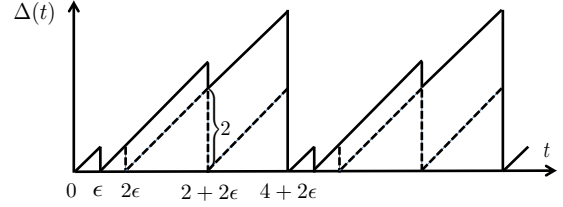


Fig. 2. Evolution of the age $\Delta(t)$ in the ϵ -wait policy in the example.

wait policy can be far from the optimum.

This example points out a key difference between data communication systems and information update systems: In communication systems, all packets are equally important; however, in information update systems, an update packet is useful only if it carries some fresh information to the destination. While the theory of data communications is quite mature, the optimal control of information updates remains open.

The aim of this paper is to answer the following questions: How to optimally submit update packets to minimize the age-of-information at the destination? When is the zero-wait policy optimal? To that end, the following are the key contributions of this paper:

- We generalize [22] by introducing two new features: (i) age penalty functions and (ii) correlated (non-*i.i.d.*) packet transmission times. We define a general age penalty function $g(\Delta)$ to characterize the level of dissatisfaction for data staleness, where $g(\cdot)$ can be any non-negative and non-decreasing function. This age penalty model is quite general, as it allows $g(\cdot)$ to be discontinuous and non-convex. In practice, one can choose the age penalty function based on the specific applications; a few examples are provided in Section III-B. In addition, the packet transmission times are modeled as a stationary ergodic Markov chain with an uncountable state space, which is more general than the *i.i.d.* transmission time processes assumed in related studies.
- We formulate the average age penalty minimization problem as a constrained semi-Markov decision problem (SMDP) with an uncountable state space. The set of feasible policies in this problem contains all causal policies, such that control decisions are made based on the history and current information of the system, which is more general than the feasible policy space considered in [22]. Despite the difficulty of this problem, we manage to solve it by a divide-and-conquer approach: We first prove that there exists a stationary randomized policy that is optimal for this problem (Theorem 1). Further, we prove that there exists a stationary deterministic policy that is optimal for this problem (Theorem 2). Finally, we develop a low-complexity algorithm to find the optimal stationary deterministic policy (Theorem 3).
- We further investigate when the zero-wait policy is optimal. For the special case of proportional penalty function and *i.i.d.* transmission times, we devise a simpler solution to minimize the average age (Theorem 4), and obtain a sufficient and necessary condition to characterize when

the zero-wait policy is optimal (Theorem 5). We find that the zero-wait policy is optimal if the transmission times are constant; and is not optimal for many commonly used transmission time distributions in communication and queueing theory, such as exponential distribution, geometric distribution, Erlang distribution, hyperexponential distribution, etc. (Corollary 1). In addition, several sufficient conditions for the optimality of the zero-wait policy are provided for general age penalty functions and correlated transmission time processes (Lemma 3).

- Our theoretical and simulation results demonstrate that, in many scenarios, the optimal information update policy is to wait for a certain amount of time before submitting a new update. In particular, the zero-wait policy is far from the optimum if (i) the age penalty function grows quickly with respect to the age, (ii) the packet transmission times over the channel are positively correlated over time, or (iii) the packet transmission times are highly random (e.g., following a heavy-tail distribution).

The rest of this paper is organized as follows. In Section II, we discuss some related work. In Section III, we describe the system model and the formulation of the average age penalty minimization problem. In Section IV, we develop the optimal update policy that minimizes the average age penalty among all causal policies. In Section V, we provide sufficient and necessary conditions for the optimality of the zero-wait policy. Finally, in Section VI, we conclude the paper.

II. RELATED WORK

The age-of-information was defined as a metric of data freshness as early as 1990s in the studies of real-time databases [5], [24]–[26]. In recent years, queueing theoretic techniques were used to evaluate the age-of-information in various system settings. The average age was analyzed for First-Come First-Served (FCFS) systems [6], Last-Come First-Served (LCFS) systems [7], multi-source networks [8], [9], and multi-channel networks [10]–[12]. A peak age metric was introduced in [13] and studied for a multi-class M/G/1 queueing system [16].

In [13]–[15], a packet management policy was shown to reduce the age, in which the old updates waiting in the queue are discarded when a new sample arrives. In [17], [27], it was shown that controlling buffer size and packet deadline can improve the age. The age performance in the presence of errors was analyzed in [18]. Gamma-distributed transmission times was considered in [19]. The age-of-information under energy-harvesting constraints was studied in [22], [28], [29]. Source coding techniques for reducing the age were evaluated in [30]. In [20], [21], it was proven that the LGFS policy achieves a smaller age process (in a stochastic ordering sense) than any other causal policy for multi-channel and multi-hop networks with an arbitrary (e.g., out-of-order) update arrival process. Scheduling of updates broadcasted to multiple users with unreliable channels was optimized in [31]. NP-hardness of wireless scheduling for minimizing the age in general networks was investigated in [32], [33]. Reducing the age of channel state information in wireless systems was studied in [34], [35]. Experimental evaluation of the age-of-information was conducted via emulation in [36].

The work that is most relevant to this paper is [22]. We have generalized the results in [22] by finding the optimal update policy for more general age penalty model, transmission time model, and in a larger feasible policy space.

III. MODEL AND FORMULATION

A. System Model

We consider an information update system depicted in Fig. 1, where a source node generates update packets and sends them to a destination through a channel. The source node generates and submits update packets at successive times S_0, S_1, \dots . The source node has access to the idle/busy state of the channel through ACK and is able to generate updates at any time by its own will. As we have mentioned before, the source node should not generate a new update when the channel is busy sending previous updates, because this will incur an unnecessary waiting time in the queue.

Suppose that Update i is submitted at time S_i , and its transmission time is $Y_i \geq 0$. Hence, Update i is delivered at time $D_i = S_i + Y_i$. We assume that Update 0 is submitted to an idle channel at time $S_0 = -Y_0$ and is delivered at $D_0 = 0$, as shown in Fig. 3. After Update i is delivered at time D_i , the source node may insert a *waiting time* $Z_i \in [0, M]$ before submitting Update $i + 1$ at time $S_{i+1} = D_i + Z_i$, where $M < \infty$ is the maximum waiting time allowed by the system. The source node can switch to a low-power sleep mode during the waiting period $[D_i, S_{i+1})$. We assume that the transmission time process (Y_0, Y_1, \dots) is a *stationary and ergodic Markov chain* with a possibly uncountable state space and a positive mean $0 < \mathbb{E}[Y_i] < \infty$.² This model generalizes the *i.i.d.* transmission time processes in related studies. This Markovian model is introduced to study the impact of temporal correlation on the optimality of the zero-wait policy. In Section V, we will see that the zero-wait policy is close to the optimum when the transmission times are negatively correlated; and is far from the optimum when the transmission times are positively correlated.

At any time t , the most recently received update packet is generated at time

$$U(t) = \max\{S_i : D_i \leq t\}. \quad (2)$$

The *age-of-information* $\Delta(t)$ is defined as [5], [6]

$$\Delta(t) = t - U(t), \quad (3)$$

which is also referred to as *age*. The age $\Delta(t)$ is a stochastic process that increases linearly with t between updates, with downward jumps occurring when updates are delivered. As shown in Fig. 3, Update i is sent at time $t = S_i$, and is delivered at time $D_i = S_i + Y_i$ with age $\Delta(D_i) = D_i - S_i = Y_i$. After that, the age increases linearly and reaches $\Delta(D_{i+1}^-) = Y_i + Z_i + Y_{i+1}$ just before Update $i+1$ is delivered. Then, at time D_{i+1} , the age drops to $\Delta(D_{i+1}) = Y_{i+1}$.

²The results in this paper can be readily extended to a more general Markovian transmission time process (Y_0, Y_1, \dots) with a longer memory, in which \mathbf{W}_i is defined as $\mathbf{W}_i = (Y_i, Y_{i+1}, \dots, Y_{i+k})$ for some finite k , and the sequence $(\mathbf{W}_0, \mathbf{W}_1, \dots)$ forms a Markov chain.

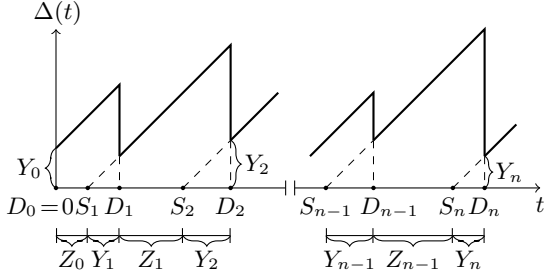


Fig. 3. Evolution of the age-of-information $\Delta(t)$.

B. Problem Formulation

We introduce an age penalty function $g(\Delta)$ to represent the level of dissatisfaction for data staleness or the need for new information update. The function $g: [0, \infty) \rightarrow [0, \infty)$ is assumed to be *measurable, non-negative, and non-decreasing*. This age penalty model is quite general, as it allows $g(\cdot)$ to be discontinuous and non-convex. In practice, one can specify the age penalty function based on the applications. A few examples are discussed in the following:

1. *Online Learning*: Fresh data is critical in online learning which is of great interest to practitioners due to the recent emergence of real-time applications such as advertisement placement and online web ranking [37], [38]. One can employ the age penalty functions that grow quickly with respect to the age, such as the power function $g(\Delta) = \Delta^a$ and the exponential function $g(\Delta) = e^{a\Delta}$ with $a \geq 0$, to characterize the desire for data refreshing in these applications.
2. *Periodic Inspection and Monitoring*: One can use the stair-step function $g(\Delta) = \lfloor a\Delta \rfloor$, where $\lfloor x \rfloor$ is the largest integer no greater than x , to characterize the dissatisfaction of data staleness if the information of interest is checked periodically [39], [40].

Two age penalty functions are depicted in Figure 4. To analyze the average age penalty, we decompose the area under the curve $g(\Delta(t))$ into a sum of disjoint components: Consider the time interval $[0, D_n]$, where $D_n = \sum_{i=0}^{n-1} (Z_i + Y_{i+1})$. In this interval, the area under $g(\Delta(t))$ can be seen as the concatenation of the areas Q_i , $0 \leq i \leq n-1$, such that

$$\int_0^{D_n} g(\Delta(t)) dt = \sum_{i=0}^{n-1} Q_i,$$

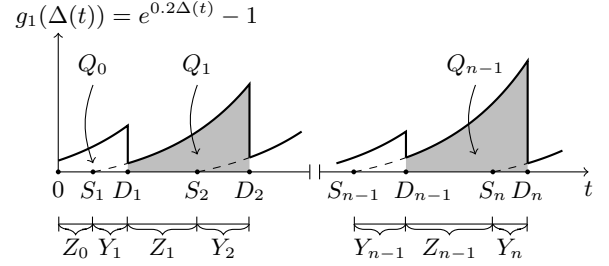
where

$$Q_i = \int_{D_i}^{D_i + Z_i + Y_{i+1}} g(\Delta(t)) dt = \int_{Y_i}^{Y_i + Z_i + Y_{i+1}} g(\tau) d\tau. \quad (4)$$

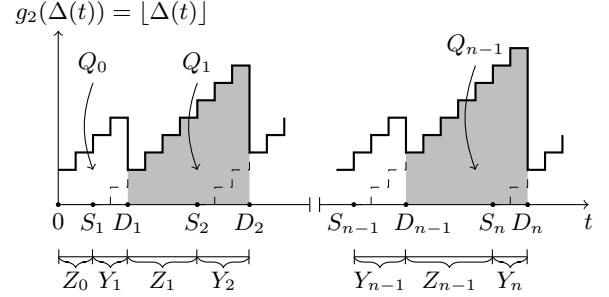
In the second equation of (4), we have used the fact that $\Delta(t) = t - S_i = t - (D_i - Y_i)$ for $t \in [D_i, D_i + Z_i + Y_{i+1}]$. Let us define

$$q(y, z, y') = \int_y^{y+z+y'} g(\tau) d\tau. \quad (5)$$

Then, Q_i can be expressed as $Q_i = q(Y_i, Z_i, Y_{i+1})$. Since $g(\Delta)$ is non-negative, the function $q(y, z, y')$ is non-decreasing



(a) An exponential age penalty function $g_1(\Delta) = e^{0.2\Delta} - 1$.



(b) A stair-step age penalty function $g_2(\Delta) = \lfloor \Delta \rfloor$.

Fig. 4. Evolutions of two age penalty functions.

in z . We assume

$$\mathbb{E}[q(Y_i, M, Y_{i+1})] < \infty, \quad (6)$$

which implies $\mathbb{E}[q(Y_i, Z_i, Y_{i+1})] < \infty$ for all $Z_i \in [0, M]$.

Our goal is to minimize the average age penalty by controlling the sequence of waiting times (Z_0, Z_1, \dots) . Let $\pi \triangleq (Z_0, Z_1, \dots)$ denote an information update policy. We consider the class of *causal* policies, in which control decisions are made based on the history and current information of the system, as well as the distribution of the transmission time process (Y_0, Y_1, \dots) . Specifically, Z_i is determined based on the past states (Y_0, Y_1, \dots, Y_i) and control actions $(Z_0, Z_1, \dots, Z_{i-1})$, without using the realizations of future transmission times $(Y_{i+1}, Y_{i+2}, \dots)$; but the conditional distribution of $(Y_{i+1}, Y_{i+2}, \dots)$ given (Y_0, Y_1, \dots, Y_i) is available. Let Π denote the set of all causal policies satisfying $Z_i \in [0, M]$ for all i .

The average age penalty per unit time is defined by³

$$\limsup_{n \rightarrow \infty} \frac{\mathbb{E} \left[\int_0^{D_n} g(\Delta(t)) dt \right]}{\mathbb{E}[D_n]}. \quad (7)$$

Because the transmission time process (Y_0, Y_1, \dots) is stationary and ergodic, we can obtain $\mathbb{E}[Y_i] = \mathbb{E}[Y_{i+1}]$ and hence

³There are two widely used definitions of average cost per unit time in infinite-horizon undiscounted semi-Markov decision problems (SMDP) [41]–[45]: In one definition, the average cost is the limit of expected total cost over a finite deterministic horizon divided by the length of the horizon, i.e., $\limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T g(\Delta(t)) dt \right]$; in the second definition, the average cost is the limit of the expected total cost over a finite number of stages divided by the expected cumulative time of these stages, as in (7). These two definitions are both reasonable [41], [44], [45]. They are equal under stationary randomized policies in which the generated semi-Markov chain has one ergodic class; see [41]–[43] for finite and countable state models. In general, however, these criteria are different. In our study, the second definition turns out to be analytically convenient.

$\mathbb{E}[D_n] = \mathbb{E}[\sum_{i=0}^{n-1} (Z_i + Y_{i+1})] = \mathbb{E}[\sum_{i=0}^{n-1} (Y_i + Z_i)]$. Using this, the optimal information update problem for minimizing the average age penalty can be formulated as

$$\bar{g}_{\text{opt}} = \min_{\pi \in \Pi} \limsup_{n \rightarrow \infty} \frac{\mathbb{E} \left[\sum_{i=0}^{n-1} q(Y_i, Z_i, Y_{i+1}) \right]}{\mathbb{E}[\sum_{i=0}^{n-1} (Y_i + Z_i)]} \quad (8)$$

$$\text{s.t. } \liminf_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \left[\sum_{i=0}^{n-1} (Y_i + Z_i) \right] \geq \frac{1}{f_{\text{max}}}, \quad (9)$$

where \bar{g}_{opt} is the optimum objective value of Problem (8), the expectation \mathbb{E} is taken over the stochastic process (Y_0, Y_1, \dots) for given policy π , and f_{max} is the maximum allowed average update frequency due to a long-term average resource constraint (i.e., on the power resource or CPU cycles) spent on generating information updates. We assume $M > 1/f_{\text{max}}$ such that Problem (8) is always feasible and $\bar{g}_{\text{opt}} < \infty$. In this paper, we will study Problem (8) both with and without the constraint (9). In Section V, sufficient and necessary conditions will be provided to characterize when the zero-wait policy is optimal for solving Problem (8) without the constraint (9).

In Problem (8), Y_i is the state of an embedded Markov chain, Z_i is the control action taken after observing Y_i , $Y_i + Z_i$ is the period of stage i , and $q(Y_i, Z_i, Y_{i+1})$ is the reward related to both stage i and $i+1$. Therefore, Problem (8) belongs to the class of constrained semi-Markov decision problems (SMDP) with an uncountable state space, which is generally known to be quite difficult. The class of SMDPs includes Markov decision problems (MDPs) [45], [46] and optimization problems of renewal processes [47] as special cases. Most existing studies on SMDPs deal with (i) unconstrained SMDPs, e.g., [41], [45], [48], [49], or (ii) constrained SMDPs with a countable state space, e.g., [44], [50]–[52]. However, the optimality equations (e.g., Bellman's equation) for solving unconstrained SMDPs are not applied to constrained SMDPs [53], and the studies on constrained SMDPs with a countable state space cannot be directly applied to Problem (8) which has an uncountable state space.

IV. OPTIMAL INFORMATION UPDATE POLICY

In this section, we develop a chain of new theoretical results to solve Problem (8) in a divide-and-conquer fashion: First, we prove that there exists a *stationary randomized* policy that is optimal for Problem (8). Further, we prove that there exists a *stationary deterministic* policy that is optimal for Problem (8). Finally, we develop a low-complexity algorithm to find the *optimal* stationary deterministic policy that solves Problem (8).

A. Optimality of Stationary Randomized Policies

A policy $\pi \in \Pi$ is said to be a *stationary randomized* policy, if it observes Y_i and then chooses a waiting time $Z_i \in [0, M]$ based only on the observed value of Y_i . In this case, Z_i is determined according to a conditional probability measure $p(y, A) \triangleq \Pr[Z_i \in A | Y_i = y]$ that is invariant for

all $i = 0, 1, \dots$. We use Π_{SR} ($\Pi_{\text{SR}} \subseteq \Pi$) to denote the set of stationary randomized policies such that

$\Pi_{\text{SR}} = \{\pi \in \Pi : \text{Given the observation } Y_i = y_i, Z_i \text{ is chosen according to the probability measure } p(y_i, A) \text{ for all } i\}$.

Note that (Y_i, Z_i, Y_{i+1}) is stationary and ergodic for all stationary randomized policies. In the sequel, when we refer to the stationary distribution of a stationary randomized policy $\pi \in \Pi_{\text{SR}}$, we will remove subscript i . In particular, the random variables (Y_i, Z_i, Y_{i+1}) are replaced by (Y, Z, Y') , where Z is chosen based on the conditional probability measure $\Pr[Z \in A | Y = y] = p(y, A)$ after observing $Y = y$, and (Y, Y') have the same joint distribution as (Y_0, Y_1) . The first key result of this paper is stated as follows:

Theorem 1: (Optimality of Stationary Randomized Policies) If $M < \infty$, $g : [0, \infty) \rightarrow [0, \infty)$ is measurable and non-negative, (Y_0, Y_1, \dots) is a stationary ergodic Markov chain with $Y_i \geq 0$ and $0 < \mathbb{E}[Y_i] < \infty$, and condition (6) is satisfied, then there exists a stationary randomized policy that is optimal for Problem (8).

Proof sketch of Theorem 1: For any policy $\pi \in \Pi$, define the finite time-horizon average occupation measures

$$a_{n,\pi} \triangleq \frac{1}{n} \mathbb{E} \left[\sum_{i=0}^{n-1} q(Y_i, Z_i, Y_{i+1}) \right] - \frac{\bar{g}_{\text{opt}}}{n} \mathbb{E} \left[\sum_{i=0}^{n-1} (Y_i + Z_i) \right], \quad (10)$$

$$b_{n,\pi} \triangleq \frac{1}{n} \mathbb{E} \left[\sum_{i=0}^{n-1} (Y_i + Z_i) \right]. \quad (11)$$

Let Γ_{SR} be the set of limit points of sequences $((a_{n,\pi}, b_{n,\pi}), n = 1, 2, \dots)$ associated with all stationary randomized policies in Π_{SR} . We first prove that Γ_{SR} is convex and compact. Then, we show that there exists an optimal policy π_{opt} of Problem (8), such that the sequence $((a_{n,\pi_{\text{opt}}}, b_{n,\pi_{\text{opt}}}), n = 1, 2, \dots)$ associated with policy π_{opt} has a limit point (a^*, b^*) satisfying $(a^*, b^*) \in \Gamma_{\text{SR}}$, $a^* \leq 0$, and $b^* \geq \frac{1}{f_{\text{max}}}$. Since $(a^*, b^*) \in \Gamma_{\text{SR}}$, there exists a stationary randomized policy π^* achieving this limit point (a^*, b^*) . Finally, we show that policy π^* is optimal for Problem (8), which completes the proof. The details are available in Appendix A. ■

The convexity and compactness properties of the set of average occupation measures are essential in the study of constrained MDPs [54, Sec. 1.5], which dates back to Derman's monograph in 1970 [55]. Recently, it was used in stochastic optimization for discrete-time queueing systems and renewal processes, e.g., [47], [56]. The techniques in these studies cannot directly handle constrained SMDPs with an uncountable state space, like Problem (8). One crucial novel idea in our proof is to introduce \bar{g}_{opt} in the definition of average occupation measures in (10), which turns out to be essential in later steps for showing the optimality of the stationary randomized policy π^* . In addition, we have also used one property of Problem (8) in the proof: the observation Y_{i+1} depends only on the immediately preceding state Y_i and not on earlier system states (Y_0, \dots, Y_{i-1}) and control actions (Z_0, \dots, Z_{i-1}) .

By Theorem 1, we only need to consider the class of

stationary randomized policies Π_{SR} . Using this, Problem (8) can be simplified to the following functional optimization problem, as shown in Appendix A:

$$\begin{aligned} \min_{p(y,A)} \quad & \frac{\mathbb{E}[q(Y, Z, Y')]}{\mathbb{E}[Y + Z]} \\ \text{s.t.} \quad & \mathbb{E}[Y + Z] \geq \frac{1}{f_{\max}} \\ & 0 \leq Z \leq M, \end{aligned} \quad (12)$$

where $p(y, A) = \Pr[Z \in A | Y = y]$ is the conditional probability measure of some stationary randomized policy, and (Y, Y') have the same distribution as (Y_0, Y_1) .

B. Optimality of Stationary Deterministic Policies

A policy $\pi \in \Pi_{\text{SR}}$ is said to be a *stationary deterministic policy* if $Z_i = z(Y_i)$ for all $i = 0, 1, \dots$, where $z : [0, \infty) \rightarrow [0, M]$ is a deterministic function. We use Π_{SD} ($\Pi_{\text{SD}} \subseteq \Pi_{\text{SR}}$) to denote the set of stationary deterministic policies such that

$$\Pi_{\text{SD}} = \{\pi \in \Pi_{\text{SR}} : Z_i = z(Y_i) \text{ for all } i\}.$$

Theorem 2: (Optimality of Stationary Deterministic Policies) If $g : [0, \infty) \rightarrow [0, \infty)$ is measurable and non-decreasing, then there exists a stationary deterministic policy that is optimal for Problem (12).

Proof sketch of Theorem 2: Since $g(\Delta)$ is non-decreasing, $q(y, z, y')$ is convex in z for any fixed y and y' . Using Jensen's inequality, we can show that for any feasible stationary randomized policy $\pi_1 \in \Pi_{\text{SR}}$, there is a feasible stationary deterministic policy that is no worse than policy π_1 . The details are provided in Appendix B. ■

Let μ_Y be the probability measure of Y_i , then any bounded measurable function $z : [0, \infty) \rightarrow [0, M]$ belongs to the Lebesgue space $L^2(\mu_Y)$ [57, Section 3], because

$$\int_0^\infty |z(y)|^2 d\mu_Y(y) \leq \int_0^\infty M^2 d\mu_Y(y) = M^2 < \infty.$$

By Theorems 1 and 2, we only need to consider the class of stationary deterministic policies Π_{SD} and Problem (8) is simplified as the following functional optimization problem:

$$\min_{z(\cdot) \in L^2(\mu_Y)} \frac{\mathbb{E}[q(Y, z(Y), Y')]}{\mathbb{E}[Y + z(Y)]} \quad (13)$$

$$\begin{aligned} \text{s.t.} \quad & \mathbb{E}[Y + z(Y)] \geq \frac{1}{f_{\max}} \\ & 0 \leq z(y) \leq M, \forall y \geq 0, \end{aligned} \quad (14)$$

where $z(\cdot)$ is the function associated with a stationary deterministic policy $\pi \in \Pi_{\text{SD}}$, and (Y, Y') have the same distribution as (Y_0, Y_1) . The optimum objective value of Problem (13) is equal to \bar{g}_{opt} .

C. A Low Complexity Solution to Problem (13)

Lemma 1: If $g : [0, \infty) \rightarrow [0, \infty)$ is measurable, non-negative, and non-decreasing, then the functional $h : L^2(\mu_Y) \rightarrow [0, \infty)$ defined by

$$h(z) = \frac{\mathbb{E}[q(Y, z(Y), Y')]}{\mathbb{E}[Y + z(Y)]}$$

Algorithm 1 Two-layer bisection method for Problem (13)

given $l = 0$, sufficiently large $u > \bar{g}_{\text{opt}}$, tolerance ϵ_1 .
repeat
 $c := (l + u)/2$.
given $\zeta_l = 0$, sufficiently large $\zeta_u > 0$, tolerance ϵ_2 .
 $\zeta := \zeta_l$, $\nu := \zeta + c$.
 Compute $z_\nu(\cdot)$ in (17).
if $\mathbb{E}[Y + z_\nu(Y)] < \frac{1}{f_{\max}}$ **then**
repeat
 $\zeta := (\zeta_l + \zeta_u)/2$, $\nu := \zeta + c$.
 Compute $z_\nu(\cdot)$ in (17).
if $\mathbb{E}[Y + z_\nu(Y)] \geq \frac{1}{f_{\max}}$, $\zeta_u := \zeta$; **else**, $\zeta_l := \zeta$.
until $\zeta_u - \zeta_l \leq \epsilon_2$.
end if
if $f(c) \leq 0$, $u := c$; **else**, $l := c$.
until $u - l \leq \epsilon_1$.
return $z(\cdot) := z_\nu(\cdot)$.

is quasi-convex.

Proof: See Appendix C. ■

Therefore, Problem (13) is a functional quasi-convex optimization problem. In order to solve Problem (13), we consider the following functional convex optimization problem with a parameter c :

$$f(c) = \min_{z(\cdot) \in L^2(\mu_Y)} \mathbb{E}[q(Y, z(Y), Y')] - c\mathbb{E}[Y + z(Y)] \quad (15)$$

$$\begin{aligned} \text{s.t.} \quad & \mathbb{E}[Y + z(Y)] \geq \frac{1}{f_{\max}} \\ & 0 \leq z(y) \leq M, \forall y \geq 0. \end{aligned} \quad (16)$$

It is easy to show that $\bar{g}_{\text{opt}} \leq c$ if and only if $f(c) \leq 0$ [58]. Therefore, we can solve Problem (13) by a two-layer nested algorithm, such as Algorithm 1. In the inner layer, we use *bisection* to solve Problem (15) for any given parameter c ; in the outer layer, we employ *bisection* again to search for a c^* such that $f(c^*) = 0$ and thus $\bar{g}_{\text{opt}} = c^*$. Algorithm 1 has low complexity. It requires at most $\lceil \log_2((u - l)/\epsilon_1) \rceil \times \lceil \log_2((\zeta_u - \zeta_l)/\epsilon_2) \rceil$ iterations to terminate and each iteration involves computing $\mathbb{E}[z_\nu(Y)]$ based on (17). The optimality of Algorithm 1 is guaranteed by the following theorem:

Theorem 3: If $g : [0, \infty) \rightarrow [0, \infty)$ is measurable, non-negative, and non-decreasing, then an optimal solution π_{opt} to Problem (13) is obtained by Algorithm 1, where the function $z_\nu(\cdot)$ is determined by

$$z_\nu(y) = \sup\{z \in [0, M] : \mathbb{E}[g(y + z + Y') | Y = y] \leq \nu\}, \quad (17)$$

and (Y, Y') follow the same distribution as (Y_0, Y_1) .

Proof Sketch of Theorem 3: We use Lagrangian duality theory to solve Problem (13). Different from traditional finite dimensional optimization problems [58], Problem (13) is an infinite dimensional functional optimization problem. Therefore, the Karush-Kuhn-Tucker (KKT) theorem for infinite dimensional space [59], [60] and the calculus of variations are required in the analysis. In particular, since the Lagrangian may not be strictly convex for some penalty functions, one-sided Gâteaux derivative (similar to sub-gradient in finite

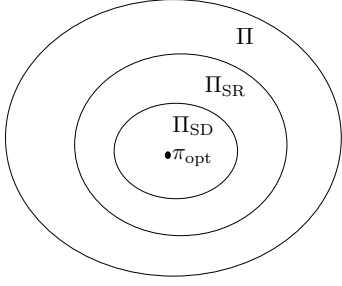


Fig. 5. Illustration of the set of causally feasible policies Π , the set of stationary randomized policies Π_{SR} , the set of stationary deterministic policies Π_{SD} , and the obtained optimal policy π_{opt} .

dimensional space) is used to solve the KKT conditions in Lebesgue space $L^2(\mu_Y)$. The proof details are provided in Appendix D. ■

The policy spaces Π , Π_{SR} , Π_{SD} , and the obtained optimal policy π_{opt} are depicted in Fig. 5.

V. WHEN IS IT BETTER TO WAIT THAN TO UPDATE?

When $f_{max} = \infty$, the constraint (16) is always satisfied. In this case, a logical policy is the *zero-wait* policy: the source node submits a fresh update once the prior update is delivered, i.e., $\pi_{zero-wait} = (0, 0, \dots)$. According to the example in the introduction, this zero-wait policy is not always optimal. In this section, we will study when it is optimal to submit updates with the minimum average waiting time and when it is not.

A. A Special Case of $g(\Delta) = \Delta$ with i.i.d. Transmission Times

Consider the case that $g(\Delta) = \Delta$ and the Y_i 's are i.i.d. with $0 < \mathbb{E}[Y] < \infty$. In this case, Problem (13) has a simpler solution than that provided by Algorithm 1. Interestingly, this solution explicitly characterizes whether the optimal control $z(\cdot)$ can have the minimum average waiting time such that $\mathbb{E}[Y + z(Y)] = 1/f_{max}$.

As shown in Fig. 3, $Q_n = \frac{1}{2} [(Y_n + Z_n + Y_{n+1})^2 - Y_n^2]$ is the area of a trapezoid. This corresponds to

$$q(y, z, y') = \frac{1}{2} [(y + z + y')^2 - y^2].$$

Because the Y_i 's are i.i.d., Y and Y' in Problem (13) are also i.i.d. Using this, we can obtain

$$\begin{aligned} & \mathbb{E}[q(Y, z(Y), Y')] \\ &= \mathbb{E} \left[\frac{1}{2} (Y + z(Y) + Y')^2 - \frac{1}{2} Y'^2 \right] \\ &= \frac{1}{2} \mathbb{E} [(Y + z(Y))^2] + \mathbb{E}[Y + z(Y)] \mathbb{E}[Y'], \end{aligned} \quad (18)$$

where in (18) we have used that $\mathbb{E}[Y^2] = \mathbb{E}[Y'^2]$. Hence, Problem (13) can be reformulated as

$$\min_{z \in L^2(\mu_Y)} \frac{\mathbb{E}[(Y + z(Y))^2]}{2\mathbb{E}[Y + z(Y)]} + \mathbb{E}[Y] \quad (19)$$

$$\begin{aligned} \text{s.t. } & \mathbb{E}[Y + z(Y)] \geq \frac{1}{f_{max}} \\ & 0 \leq z(y) \leq M, \quad \forall y \geq 0. \end{aligned} \quad (20)$$

Algorithm 2 Bisection method for solving Problem (19)

given $l = 0$, sufficiently large u , tolerance ϵ .
repeat
 $\beta := (l + u)/2$.
 $o := \mathbb{E}[(Y + z(Y))] - \max\left(\frac{1}{f_{max}}, \frac{\mathbb{E}[(Y + z(Y))^2]}{2\beta}\right)$,
where $z(\cdot)$ is given by (21).
if $o \geq 0$, $u := \beta$; **else**, $l := \beta$.
until $u - l \leq \epsilon$.
Compute $z(\cdot)$ by (21).
return $z(\cdot)$.

The following lemma tells us that Problem (19) is a functional convex optimization problem.

Lemma 2: The functional $h_1 : L^2(\mu_Y) \rightarrow \mathbb{R}$ defined by

$$h_1(z) = \frac{\mathbb{E}[(Y + z(Y))^2]}{\mathbb{E}[Y + z(Y)]}$$

is convex on the domain

$$\text{dom } h_1 = \{z \in L^2(\mu_Y) : z(y) \in [0, M], \forall y \geq 0\}.$$

Proof: See Appendix E. ■

Using the KKT theorem for infinite dimensional space and the calculus of variations, we can obtain

Theorem 4: If $\mathbb{E}[Y] > 0$, the optimal solution to Problem (19) is

$$z(y) = (\beta - y)_0^M, \quad (21)$$

where $(x)_0^M \triangleq \min\{\max\{x, 0\}, M\}$ and $\beta > 0$ satisfies

$$\mathbb{E}[Y + z(Y)] = \max\left(\frac{1}{f_{max}}, \frac{\mathbb{E}[(Y + z(Y))^2]}{2\beta}\right). \quad (22)$$

Proof: See Appendix F. ■

Equation (21) has the form of a *water-filling* solution, where the water-level β is given by the root of equation (22). One can observe that (17) reduces to (21) if $g(\Delta) = \Delta$, the Y_i 's are i.i.d., and ν is replaced by $\beta + \mathbb{E}[Y]$. The root β of equation (22) can be simply solved by the bisection search method in Algorithm 2. We note that Algorithm 2 has lower complexity than Algorithm 1 in the special case of $g(\Delta) = \Delta$ and i.i.d. transmission process, while Algorithm 1 can obtain the optimal policy in more general scenarios.

Theorem 4 provides a closed-form criterion on whether the optimal $z(\cdot)$ satisfies $\mathbb{E}[Y + z(Y)] = 1/f_{max}$. Specifically, (21) and (22) tell us that if $1/f_{max} \geq \frac{\mathbb{E}[(Y + z(Y))^2]}{2\beta}$, then the optimal control $z(\cdot)$ satisfies

$$\mathbb{E}[Y + z(Y)] = \frac{1}{f_{max}} \geq \frac{\mathbb{E}[(Y + z(Y))^2]}{2\beta}, \quad (23)$$

such that the optimal policy achieves the minimum possible average waiting time; otherwise, if $1/f_{max} < \frac{\mathbb{E}[(Y + z(Y))^2]}{2\beta}$, the optimal control $z(\cdot)$ satisfies

$$\mathbb{E}[Y + z(Y)] = \frac{\mathbb{E}[(Y + z(Y))^2]}{2\beta} > \frac{1}{f_{max}}, \quad (24)$$

such that the optimal policy can not achieve the minimum possible average waiting time. In [22], the author solved a

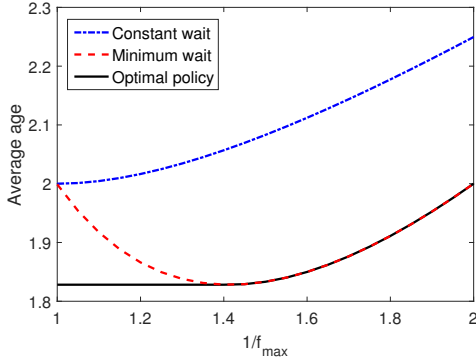


Fig. 6. Average age vs. $1/f_{\max}$ with *i.i.d.* discrete transmission times.

slightly different version of Problem (19): an equality constraint on the updating frequency was considered in [22], while an inequality constraint is adopted in Problem (19). It was observed in [22] that the optimal time-average age is not necessarily decreasing in the update frequency. The solution to Problem (19) in Theorem 4 further allows us to obtain the optimal update frequency.

Furthermore, if $\mathbb{E}[Y] \geq 1/f_{\max}$ (e.g., $f_{\max} = \infty$), the constraint (20) is always satisfied and can be removed. In this case, the optimality of the zero-wait policy is characterized by a sufficient and necessary condition in the following theorem.

Theorem 5: If $\mathbb{E}[Y] > 0$ and $f_{\max} = \infty$, then the zero-wait policy is optimal for Problem (19) if, and only if,

$$\mathbb{E}[Y^2] \leq 2y_{\inf}\mathbb{E}[Y], \quad (25)$$

where $y_{\inf} = \inf\{y \in [0, \infty) : \Pr\{Y \leq y\} > 0\}$.

Proof: See Appendix G. ■

Informally speaking, y_{\inf} is the smallest possible value of the random transmission time Y . From Theorem 5, it immediately follows that:

Corollary 1: If $\mathbb{E}[Y] > 0$ and the zero-wait policy is feasible, then the following assertions are true:

- If the transmission times are positive and constant (i.e., $Y = \text{const} > 0$), the zero-wait policy is optimal for Problem (19).
- If the transmission times satisfy $y_{\inf} = 0$, the zero-wait policy is *not* optimal for Problem (19).

Remark: As one can readily see from Corollary 1(b), the zero-wait policy is not optimal for many commonly used distributions in communication and queueing theory, such as exponential distribution, geometric distribution, Erlang distribution, hyperexponential distribution, etc.

1) *Numerical Results:* We use “optimal policy” to refer to the policy provided in Theorem 3 (or its special case in Theorem 4), and compare it with three reference policies:

- “Constant wait”: Each update is followed by a constant waiting time $Z = 1/f_{\max} - \mathbb{E}[Y]$ before submitting the next update.
- “Minimum wait”: The update waiting time is determined by $Z = z(Y)$, where $z(\cdot)$ is given by (21) and β in (22)

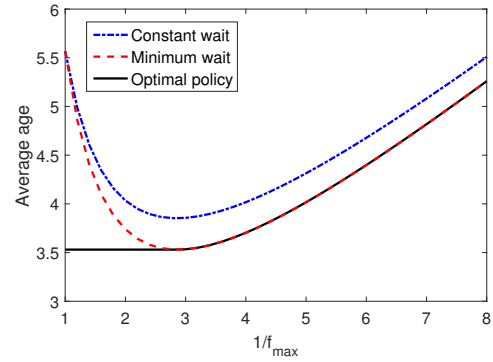


Fig. 7. Average age vs. $1/f_{\max}$ with *i.i.d.* log-normal distributed transmission times.

is chosen to satisfy $\mathbb{E}[z(Y)] = 1/f_{\max} - \mathbb{E}[Y]$.⁴

When $\mathbb{E}[Y] = 1/f_{\max}$, both the constant wait and minimum wait policies reduce to the zero-wait policy.

Two transmission time models are considered: The first is a discrete Markov chain with a probability mass function $\Pr\{Y_i = 0\} = \Pr\{Y_i = 2\} = 0.5$ and a transition matrix

$$P = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}.$$

Hence, the Y_i 's are *i.i.d.* when $p = 0.5$, and the correlation coefficient between Y_i and Y_{i+1} is $\rho = 2p - 1$. The second is a log-normal distributed Markov chain, where $Y_i = e^{\sigma X_i} / \mathbb{E}[e^{\sigma X_i}]$ and (X_0, X_1, \dots) is a Gaussian Markov process satisfying the first-order autoregressive (AR) equation

$$X_{i+1} = \eta X_i + \sqrt{1 - \eta^2} W_i,$$

where $\sigma > 0$ is the scale parameter of log-normal distribution, $\eta \in [-1, 1]$ is the parameter of the AR model, and the W_i 's are *i.i.d.* Gaussian random variables with zero mean and unit variance. The log-normal distributed Markov chain is normalized such that $\mathbb{E}[Y_i] = 1$. According to the properties of log-normal distribution, the correlation coefficient between Y_i and Y_{i+1} is $\rho = (e^\eta - 1)/(e - 1)$. Then, the Y_i 's are *i.i.d.* when $\eta = 0$. The value of M is set to be 10.

Figures 6 and 7 illustrate the average age vs. f_{\max} for *i.i.d.* discrete and log-normal distributed transmission times, respectively, where $\sigma = 1.5$. In both figures, one can observe that the constant wait policy always incurs a larger average age than the optimal policy. In addition, as expected from (23) and (24), as $1/f_{\max}$ exceeds a certain threshold, the optimal policy meets the constraint (20) with equality. For smaller values of $1/f_{\max}$, the constraint (20) is not active in the optimal solution. Consequently, the minimum wait policy deviates from the optimal policy for small values of $1/f_{\max}$, which is in accordance with Corollary 1(b).

B. General Age Penalties and Correlated Transmission Times

For general age penalties and correlated transmission time processes, it is essentially difficult to find closed-form characterization on whether the optimal control $z(\cdot)$ can have the

⁴This policy was called “ β -minimum” in [22].

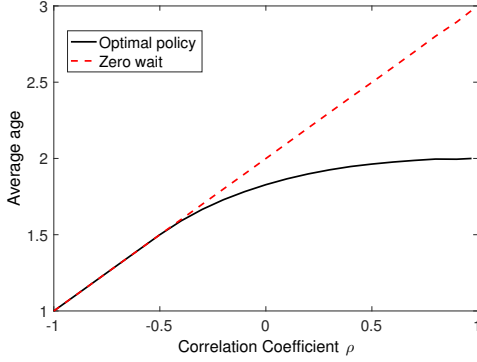


Fig. 8. Average age vs. the correlation coefficient ρ between Y_i and Y_{i+1} for discrete transmission times, where $\mathbb{E}[Y] \geq 1/f_{\max}$.

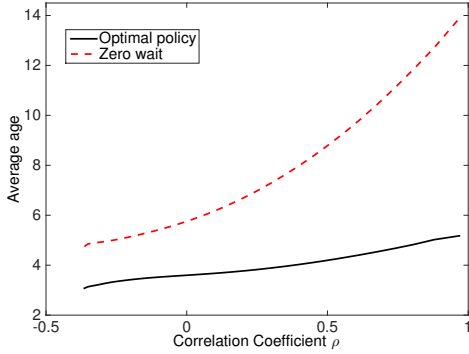


Fig. 9. Average age vs. the correlation coefficient ρ between Y_i and Y_{i+1} for log-normal distributed transmission times, where $\mathbb{E}[Y] \geq 1/f_{\max}$.

minimum average waiting time such that $\mathbb{E}[Y + z(Y)] = 1/f_{\max}$. Therefore, we focus on the case of $\mathbb{E}[Y] \geq 1/f_{\max}$ (this is equivalent to removing the update frequency constraint (16)) and study when the zero-wait policy minimizes the average age penalty. Sufficient conditions for the optimality of the zero-wait policy are provided as follows:

Lemma 3: Suppose that $\mathbb{E}[Y] \geq 1/f_{\max}$, $g(\cdot)$ is measurable, non-negative, and non-decreasing. The zero-wait policy is optimal for Problem (13) if one of the following conditions is satisfied:

- 1). The correlation coefficient between Y_i and Y_{i+1} is -1 ;
- 2). The Y_i 's are equal to a constant value;
- 3). $g(\cdot)$ is a constant function.

Proof: See Appendix H. ■

1) *Numerical Results:* We now provide some Numerical Results for general age penalties and/or correlated transmission time processes. Figures 8 and 9 depict the average age vs. the correlation coefficient ρ between Y_i and Y_{i+1} for discrete and log-normal distributed transmission times, respectively. In Fig. 8, the range of ρ is $[-1, 1)$. We observe that the zero-wait policy is optimal when $\rho \in [-1, -0.5]$, and the performance gap between the optimal policy and the zero-wait policy grows with ρ when $\rho \geq -0.5$. This is in accordance with the example in the introduction: As ρ grows, the occurrence of two consecutive zero transmission times (i.e., $(Y_i, Y_{i+1}) = (0, 0)$) increases. Therefore, more and more updates are wasted in the zero-wait policy, leading to a larger gap from the optimum.

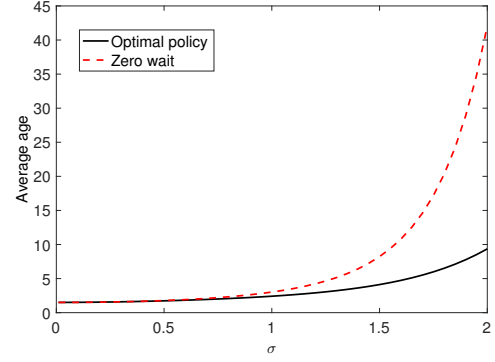


Fig. 10. Average age vs. the distribution parameter σ of *i.i.d.* log-normal distributed transmission times, where $\mathbb{E}[Y] \geq 1/f_{\max}$.

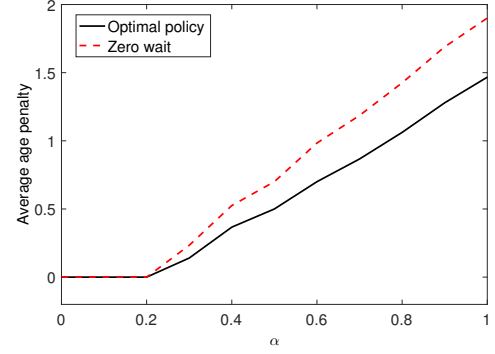


Fig. 11. Average age penalty vs. the parameter α of stair-step penalty functions with discrete transmission times, where $\mathbb{E}[Y] \geq 1/f_{\max}$, $g(\Delta) = \lfloor \alpha \Delta \rfloor$, and $\rho = 0.4$.

In Fig. 9, the range of ρ is $[(e^{-1} - 1)/(e - 1), 1)$ and the sub-optimality gap of the zero-wait policy also increases with ρ . The point $\rho = 1$ is not plotted in these figures because the corresponding Markov chains are not ergodic.

Figure 10 considers the average age vs. the parameter σ of log-normal distributed transmission times, where $\rho = (e^{0.5} - 1)/(e - 1)$. We observe that the zero-wait policy is optimal for small σ and is not optimal for large σ . When $\sigma = 0$, the transmission times are constant, i.e., $Y_i = 1$ for all i , and hence by Lemma 3, the zero-wait policy is optimal. For large σ , the time-average age of the zero-wait policy is significantly larger than the optimum. This implies that the sub-optimality gap of the zero-wait policy can be quite large for heavy-tail transmission time distributions.

Figures 11-16 show the average age penalty vs. the parameter α of three types of age penalty functions, where the stair-step function $g(\Delta) = \lfloor \alpha \Delta \rfloor$ is considered in Fig. 11 and 12, the power function $g(\Delta) = \Delta^\alpha$ is considered in Fig. 13 and 14, and the exponential function $g(\Delta) = e^{\alpha \Delta} - 1$ is considered in Fig. 15 and 16. The correlation coefficient is $\rho = 0.4$ for discrete transmission times, and is $\rho = (e^{0.5} - 1)/(e - 1)$ for log-normal distributed transmission times. We find that the zero-wait policy is optimal if $\alpha = 0$, in which case $g(\Delta)$ is a constant function. When $\alpha > 0$, the zero-wait policy may not be optimal.

These numerical results suggest that the conditions in Lemma 3 are sufficient but not necessary.

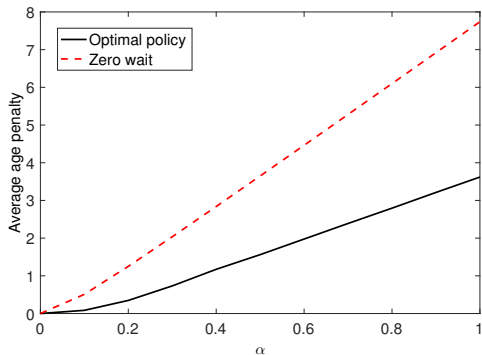


Fig. 12. Average age penalty vs. the parameter α of stair-step penalty functions with log-normal distributed transmission times, where $\mathbb{E}[Y] \geq 1/f_{\max}$, $g(\Delta) = \lfloor \alpha \Delta \rfloor$, and $\rho = (e^{0.5} - 1)/(e - 1)$.

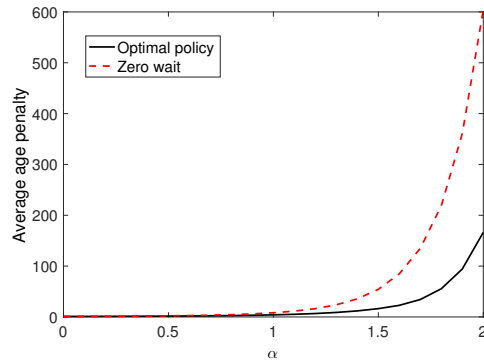


Fig. 14. Average age penalty vs. the parameter α of power penalty functions with log-normal distributed service times, where $\frac{1}{f_{\max}} \leq \mathbb{E}[Y]$, $g(\Delta) = \Delta^\alpha$, and $\rho = (e^{0.5} - 1)/(e - 1)$.

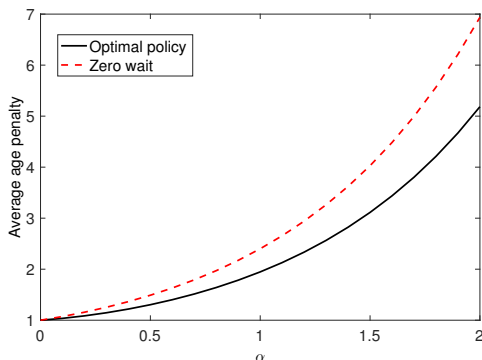


Fig. 13. Average age penalty vs. the parameter α of power penalty functions with discrete transmission times, where $\mathbb{E}[Y] \geq 1/f_{\max}$, $g(\Delta) = \Delta^\alpha$, and $\rho = 0.4$.

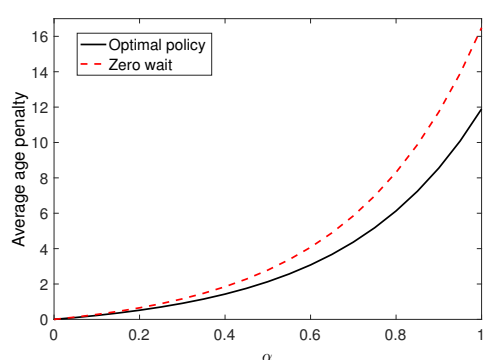


Fig. 15. Average age penalty vs. the parameter α of exponential penalty functions with discrete service times, where $\frac{1}{f_{\max}} \leq \mathbb{E}[Y]$, $g(\Delta) = e^{\alpha \Delta} - 1$, and $\rho = 0.4$.

VI. CONCLUSIONS

We studied the optimal control of information updates sent from a source node to a remote destination via a communication server. We generalized prior study on the age-of-information by considering general age penalty functions and non-*i.i.d.* transmission time processes. We developed efficient algorithms to find the optimal update policy for minimizing the average age penalty among all causal update policies. We showed that, surprisingly, the optimal policy is to wait for a certain amount of time before submitting a new update in many scenarios. Sufficient and necessary conditions were established to characterize the optimality of the zero-wait policy. In particular, the zero-wait policy is far from the optimum if (i) the age penalty function grows quickly with respect to the age, (ii) the packet transmission times over the channel are positively correlated over time, or (iii) the packet transmission times are highly random (e.g., following a heavy-tail distribution). In our future work, we will further investigate how to improve the freshness of real-time signals transmitted over a channel. Some interesting initial result has been obtained in [61].

APPENDIX A PROOF OF THEOREM 1

A. An Upper Bound of \bar{g}_{opt}

By restricting Π in Problem (8) to Π_{SR} , we obtain the following problem:

$$\begin{aligned} \bar{g}_{SR} = \min_{\pi \in \Pi_{SR}} \limsup_{n \rightarrow \infty} & \frac{\mathbb{E} \left[\sum_{i=0}^{n-1} q(Y_i, Z_i, Y_{i+1}) \right]}{\mathbb{E} \left[\sum_{i=0}^{n-1} (Y_i + Z_i) \right]} \\ \text{s.t. } \liminf_{n \rightarrow \infty} & \frac{1}{n} \mathbb{E} \left[\sum_{i=0}^{n-1} (Y_i + Z_i) \right] \geq \frac{1}{f_{\max}}, \end{aligned} \quad (26)$$

where \bar{g}_{SR} is the optimum objective value of Problem (26). Since $\Pi_{SR} \subseteq \Pi$, we can obtain

$$\bar{g}_{SR} \geq \bar{g}_{opt}. \quad (27)$$

It is easy to show that the (Y_i, Z_i, Y_{i+1}) 's are stationary and ergodic for all stationary randomized policies. This, together with the condition that $g(\cdot)$ is measurable, tells us that $q(Y_i, Z_i, Y_{i+1})$ is stationary and ergodic [62, Theorems 7.1.1 and 7.1.3]. For any stationary randomized policy $\pi =$

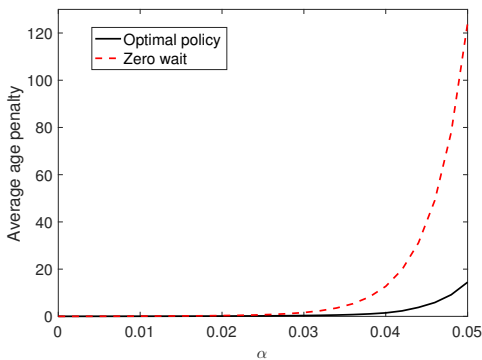


Fig. 16. Average age penalty vs. the parameter α of exponential penalty functions with log-normal distributed service times, where $\frac{1}{f_{\max}} \leq \mathbb{E}[Y]$, $g(\Delta) = e^{\alpha\Delta} - 1$, and $\rho = (e^{0.5} - 1)/(e - 1)$.

$(Z_0, Z_1, \dots) \in \Pi_{\text{SR}}$, we obtain

$$\frac{1}{n} \mathbb{E} \left[\sum_{i=0}^{n-1} q(Y_i, Z_i, Y_{i+1}) \right] = \mathbb{E}[q(Y_0, Z_0, Y_1)], \quad (28)$$

$$\frac{1}{n} \mathbb{E} \left[\sum_{i=0}^{n-1} (Y_i + Z_i) \right] = \mathbb{E}[Y_0 + Z_0]. \quad (29)$$

Hence, Problem (26) can be reformulated as Problem (12).

B. The Upper Bound of \bar{g}_{opt} is Tight, i.e., $\bar{g}_{\text{SR}} = \bar{g}_{\text{opt}}$

We will show $\bar{g}_{\text{SR}} = \bar{g}_{\text{opt}}$ in 4 steps. The following definitions are needed: Since \bar{g}_{opt} is finite, for each causally feasible policy $\pi = (Z_0, Z_1, \dots) \in \Pi$ we can define $a_{n,\pi}$ and $b_{n,\pi}$ as in (10) and (11), respectively.

Further, define Γ_{SR} as the set of limit points of sequences $((a_{n,\pi}, b_{n,\pi}), n = 1, 2, \dots)$ associated with all stationary randomized policies $\pi \in \Pi_{\text{SR}}$. Because the reward $q(Y_i, Z_i, Y_{i+1})$ and interval $Y_i + Z_i$ are stationary and ergodic for all stationary randomized policies $\pi \in \Pi_{\text{SR}}$, the sequence $(a_{n,\pi}, b_{n,\pi})$ has a unique limit point in the form of

$$(\mathbb{E}[q(Y, Z, Y')] - \bar{g}_{\text{opt}} \mathbb{E}[Y + Z], \mathbb{E}[Y + Z]). \quad (30)$$

Hence, Γ_{SR} is the set of all points $(\mathbb{E}[q(Y, Z, Y')] - \bar{g}_{\text{opt}} \mathbb{E}[Y + Z], \mathbb{E}[Y + Z])$, where each point is associated with a conditional probability measure $p(y, A) = \Pr[Z \in A | Y = y]$, and the measure of (Y, Y') is the same as that of (Y_0, Y_1) . Note that $\mathbb{E}[Y] = \mathbb{E}[Y']$.

Step 1: We will show that Γ_{SR} is a convex and compact set.

It is easy to show that Γ_{SR} is convex by considering a stationary randomized policy that is a mixture of two stationary randomized policies.

For compactness, let $((d_j, e_j), j = 1, 2, \dots)$ be any sequence of points in Γ_{SR} , we need to show that there is a convergent subsequence (d_{j_k}, e_{j_k}) whose limit is also in Γ_{SR} . Since $(d_j, e_j) \in \Gamma_{\text{SR}}$, there must exist $(Y, Z_{(j)}, Y')$ with conditional probability $p_j(y, A) = \Pr[Z_{(j)} \in A | Y = y]$, such that $d_j = \mathbb{E}[q(Y, Z_{(j)}, Y')] - \bar{g}_{\text{opt}} \mathbb{E}[Y + Z_{(j)}]$, $e_j = \mathbb{E}[Y + Z_{(j)}]$. Let μ_j be the joint probability measure of $(Y, Z_{(j)}, Y')$, then

(d_j, e_j) is uniquely determined by μ_j . For any L satisfying $L \geq M$, we can obtain

$$\begin{aligned} & \mu_j(Y \leq L, Z_{(j)} \leq L, Y' \leq L) \\ &= \Pr(Y \leq L, Y' \leq L) \\ &\geq \Pr(Y + Y' \leq L) \\ &\geq 1 - \frac{\mathbb{E}[Y + Y']}{L}, \quad \forall j, \end{aligned}$$

where the equality is due to the fact that $Z_{(j)} \leq M \leq L$ and the last inequality is due to Markov's inequality. Therefore, for any ϵ , there is an L such that

$$\liminf_{j \rightarrow \infty} \mu_j(|Y| \leq L, |Z_{(j)}| \leq L, |Y'| \leq L) \geq 1 - \epsilon.$$

Hence, the sequence of measures μ_j is tight. By Helly's selection theorem [62, Theorem 3.9.2], there is a subsequence of measures μ_{j_k} that converges weakly to a limit measure μ_∞ .

Let $(Y, Z_{(\infty)}, Y')$ and $p_\infty(y, A) = \Pr[Z_\infty \in A | Y = y]$ denote the random vector and conditional probability corresponding to the limit measure μ_∞ , respectively. We can define $d_\infty = \mathbb{E}[q(Y, Z_{(\infty)}, Y')] - \bar{g}_{\text{opt}} \mathbb{E}[Y + Z_{(\infty)}]$, $e_\infty = \mathbb{E}[Y + Z_{(\infty)}]$. Since the function $q(y, z, y')$ is in the form of an integral, it is continuous and thus measurable. Using the continuous mapping theorem [62, Theorem 3.2.4], we can obtain that $q(Y, Z_{(j_k)}, Y')$ converges weakly to $q(Y, Z_{(\infty)}, Y')$. Then, using the condition (6), together with the dominated convergence theorem (Theorem 1.6.7 of [62]) and Theorem 3.2.2 of [62], we can obtain $\lim_{k \rightarrow \infty} (d_{j_k}, e_{j_k}) = (d_\infty, e_\infty)$. Hence, $((d_j, e_j), j = 1, 2, \dots)$ has a convergent subsequence. Further, we can generate a stationary randomized policy $\pi_{\infty, \text{SR}}$ by using the conditional probability $p_\infty(y, A)$ corresponding to μ_∞ . Then, (d_∞, e_∞) is the limit point generated by the stationary randomized policy $\pi_{\infty, \text{SR}}$, which implies $(d_\infty, e_\infty) \in \Gamma_{\text{SR}}$. In summary, any sequence (d_j, e_j) in Γ_{SR} has a convergent subsequence (d_{j_k}, e_{j_k}) whose limit (d_∞, e_∞) is also in Γ_{SR} . Therefore, Γ_{SR} is a compact set.

Step 2: We will show that there exists an optimal policy $\pi_{\text{opt}} \in \Pi$ of Problem (8) such that the sequence $(a_{n,\pi_{\text{opt}}}, b_{n,\pi_{\text{opt}}})$ associated with policy π_{opt} has at least one limit point in Γ_{SR} .

Since the sequence (Y_0, Y_1, \dots) is a Markov chain, the observation Y_{i+1} depends only on the immediately preceding state Y_i and not on the history state and control $Y_0, \dots, Y_{i-1}, Z_0, \dots, Z_{i-1}$. Therefore, Y_i is the *sufficient statistic* [45, p. 252] for solving Problem (8). This tells us that there exists an optimal policy $\pi_{\text{opt}} = (Z_0, Z_1, \dots) \in \Pi$ of Problem (8) in which the control action Z_i is determined based on only Y_i , but not the history state and control $Y_0, \dots, Y_{i-1}, Z_0, \dots, Z_{i-1}$ [45]. We will show that the sequence $(a_{n,\pi_{\text{opt}}}, b_{n,\pi_{\text{opt}}})$ associated with this policy π_{opt} has at least one limit point in Γ_{SR} .

It is known that Z_i takes values in the standard Borel space $(\mathbb{R}, \mathcal{R})$, where \mathcal{R} is the Borel σ -field. According to [62, Theorem 5.1.9], for each i there exists a conditional probability measure $p'_i(y, A)$ such that $p'_i(y, A) = \Pr[Z_i \in A | Y_i = y]$ for almost all y . That is, the control action Z_i is determined based on Y_i and the conditional probability measure $p'_i(y, A) = \Pr[Z_i \in A | Y_i = y]$. One can use this conditional probability $p'_i(y, A)$ to generate a stationary randomized policy $\pi'_{i, \text{SR}} \in \Pi_{\text{SR}}$. Then, the one-stage expectation

$(\mathbb{E}[q(Y_i, Z_i, Y_{i+1})] - \bar{g}_{\text{opt}}\mathbb{E}[Y_i + Z_i], \mathbb{E}[Y_i + Z_i])$ is exactly the limit point generated by the stationary randomized policy $\pi'_{i,\text{SR}}$. Thus, $(\mathbb{E}[q(Y_i, Z_i, Y_{i+1})] - \bar{g}_{\text{opt}}\mathbb{E}[Y_i + Z_i], \mathbb{E}[Y_i + Z_i]) \in \Gamma_{\text{SR}}$ for all $i = 0, 1, 2, \dots$. Using (10), (11), and the fact that Γ_{SR} is convex, we can obtain $(a_n, \pi_{\text{opt}}, b_n, \pi_{\text{opt}}) \in \Gamma_{\text{SR}}$ for all $n = 1, 2, 3, \dots$. In other words, the sequence $(a_n, \pi_{\text{opt}}, b_n, \pi_{\text{opt}})$ is within Γ_{SR} . Since Γ_{SR} is a compact set, the sequence $(a_n, \pi_{\text{opt}}, b_n, \pi_{\text{opt}})$ must have a convergent subsequence, whose limit is in Γ_{SR} .

Step 3: Let $(a^, b^*) \in \Gamma_{\text{SR}}$ be one limit point of the sequence $(a_n, \pi_{\text{opt}}, b_n, \pi_{\text{opt}})$ associated with policy π_{opt} . We will show that $a^* \leq 0$ and $b^* \geq \frac{1}{f_{\text{max}}}$.*

Policy π_{opt} is feasible for Problem (8) and meanwhile achieves the optimum objective value \bar{g}_{opt} . Hence,

$$\limsup_{n \rightarrow \infty} \frac{c_n, \pi_{\text{opt}}}{b_n, \pi_{\text{opt}}} = \bar{g}_{\text{opt}}, \quad (31)$$

$$\liminf_{n \rightarrow \infty} b_n, \pi_{\text{opt}} \geq \frac{1}{f_{\text{max}}}, \quad (32)$$

where

$$c_n, \pi_{\text{opt}} \triangleq \frac{1}{n} \mathbb{E} \left[\sum_{i=0}^{n-1} q(Y_i, Z_i, Y_{i+1}) \right].$$

By (11), b_n, π_{opt} is upper bounded by

$$b_n, \pi_{\text{opt}} \leq M + \mathbb{E}[Y] < \infty.$$

Hence, by (10), we have

$$\begin{aligned} a_n, \pi_{\text{opt}} &= c_n, \pi_{\text{opt}} - \bar{g}_{\text{opt}} b_n, \pi_{\text{opt}} \\ &\leq \max\{c_n, \pi_{\text{opt}} - \bar{g}_{\text{opt}} b_n, \pi_{\text{opt}}, 0\} \\ &= \max\left\{\frac{c_n, \pi_{\text{opt}}}{b_n, \pi_{\text{opt}}} - \bar{g}_{\text{opt}}, 0\right\} b_n, \pi_{\text{opt}} \\ &\leq \max\left\{\frac{c_n, \pi_{\text{opt}}}{b_n, \pi_{\text{opt}}} - \bar{g}_{\text{opt}}, 0\right\} (M + \mathbb{E}[Y]). \end{aligned}$$

Taking the lim sup in this inequality and using (31), yields

$$\limsup_{n \rightarrow \infty} a_n, \pi_{\text{opt}} \leq 0. \quad (33)$$

Because (a^*, b^*) is one limit point of $(a_n, \pi_{\text{opt}}, b_n, \pi_{\text{opt}})$, we have

$$a^* \leq \limsup_{n \rightarrow \infty} a_n, \pi_{\text{opt}}, \quad b^* \geq \liminf_{n \rightarrow \infty} b_n, \pi_{\text{opt}}. \quad (34)$$

By (32)-(34), we have $a^* \leq 0$ and $b^* \geq \frac{1}{f_{\text{max}}}$.

Step 4: We will show that there exists a stationary randomized policy that is optimal for Problems (8) and (12), and thus $\bar{g}_{\text{SR}} = \bar{g}_{\text{opt}}$. By the definition of Γ_{SR} , $(a^*, b^*) \in \Gamma_{\text{SR}}$ must be the limit point generated by a stationary randomized policy $\pi^* \in \Pi_{\text{SR}}$. Let (Y, Z^*, Y') be a random vector with the stationary distribution of policy π^* . Then, (30) implies

$$(a^*, b^*) = (\mathbb{E}[q(Y, Z^*, Y')] - \bar{g}_{\text{opt}}\mathbb{E}[Y + Z^*], \mathbb{E}[Y + Z^*]).$$

Using $a^* \leq 0$ and $b^* \geq \frac{1}{f_{\text{max}}}$, we can obtain

$$\mathbb{E}[q(Y, Z^*, Y')] - \mathbb{E}[Y + Z^*]\bar{g}_{\text{opt}} \leq 0, \quad (35)$$

$$\mathbb{E}[Y + Z^*] \geq \frac{1}{f_{\text{max}}}. \quad (36)$$

By (35) and $\mathbb{E}[Y + Z^*] > 0$, we have

$$\frac{\mathbb{E}[q(Y, Z^*, Y')]}{\mathbb{E}[Y + Z^*]} \leq \bar{g}_{\text{opt}}.$$

Further, the inequality (36) suggests that the stationary randomized policy π^* is feasible for Problem (12). Hence,

$$\frac{\mathbb{E}[q(Y, Z^*, Y')]}{\mathbb{E}[Y + Z^*]} \geq \bar{g}_{\text{SR}}.$$

Therefore, $\bar{g}_{\text{SR}} \leq \bar{g}_{\text{opt}}$. This and (27) suggest that

$$\frac{\mathbb{E}[q(Y, Z^*, Y')]}{\mathbb{E}[Y + Z^*]} = \bar{g}_{\text{SR}} = \bar{g}_{\text{opt}}.$$

This completes the proof.

APPENDIX B PROOF OF THEOREM 2

Consider an arbitrarily chosen stationary randomized policy $\pi_1 \in \Pi_{\text{SR}}$ that is feasible for Problem (12). We will show that there exists a feasible stationary deterministic policy that is no worse than policy π_1 .

For any y , we can use the conditional probability $p(y, A)$ associated with policy π_1 to compute the conditional expectation $\mathbb{E}[Z|Y = y]$ by

$$\mathbb{E}[Z|Y = y] = \int_0^M zp(y, dz).$$

Since the conditional expectation $\mathbb{E}[Z|Y]$ is unique w.p.1 [62, Section 5.1], there is a deterministic function $z(\cdot)$ such that $z(y) = \mathbb{E}[Z|Y = y]$ w.p.1. Consider the set $\Lambda \subset \Pi_{\text{SR}}$ of all stationary randomized policies that satisfy $\mathbb{E}[Z|Y = y] = z(y)$ w.p.1. Then, the stationary randomized policy π_1 is in Λ . It is also easy to show that the stationary deterministic policy $(Z_i = z(Y_i), i = 1, 2, \dots)$ is also in Λ .

Using the iterated expectation, for any policy in Λ

$$\mathbb{E}[Y + Z] = \mathbb{E}[Y + \mathbb{E}[Z|Y]] = \mathbb{E}[Y + z(Y)]. \quad (37)$$

Because $\pi_1 \in \Lambda$ is feasible for Problem (12), any policy in Λ is feasible for Problem (12).

Since $q(y, z, y')$ is the integral of a non-decreasing function g , it is easy to show that the function $q(y, \cdot, y')$ is convex. For any policy $\pi \in \Lambda$, Jensen's inequality tells us that

$$\begin{aligned} &\mathbb{E}[q(Y, Z, Y')|Y, Y'] \\ &\geq q(Y, \mathbb{E}[Z|Y, Y'], Y') \\ &= q(Y, \mathbb{E}[Z|Y], Y') \\ &= q(Y, z(Y), Y'), \quad (\text{w.p.1}), \end{aligned} \quad (38)$$

where (38) is due to the fact that Z is determined based on Y , but not Y' . Taking the expectation over (Y, Y') , yields

$$\mathbb{E}[q(Y, z(Y), Y')] \leq \mathbb{E}[q(Y, Z, Y')]$$

for any policy $\pi \in \Lambda$, where equality holds if $Z = z(Y)$. This and (37) suggest that the stationary deterministic policy $(Z_i = z(Y_i), i = 1, 2, \dots)$ achieves the smallest objective value for Problem (12) among all policies in Λ . In conclusion, for any feasible stationary randomized policy $\pi_1 \in \Pi_{\text{SR}}$, we can find

a feasible stationary deterministic policy that is no worse than policy π_1 . This completes the proof.

APPENDIX C PROOF OF LEMMA 1

We need the following lemma:

Lemma 4: If $l : \mathbb{R} \rightarrow \mathbb{R}$ is a convex function, then the functional $w : L^2(\mu_Y) \rightarrow \mathbb{R}$ defined by

$$w(z) = \int_0^\infty l(z(y)) d\mu_Y(y) \quad (39)$$

is also convex.

Proof: For any $\lambda \in [0, 1]$ and $z_1, z_2 \in L^2(\mu_Y)$, we have

$$\begin{aligned} & w(\lambda z_1 + (1 - \lambda)z_2) \\ &= \int_0^\infty l(\lambda z_1(y) + (1 - \lambda)z_2(y)) d\mu_Y(y) \\ &\leq \int_0^\infty [\lambda l(z_1(y)) + (1 - \lambda)l(z_2(y))] d\mu_Y(y) \\ &= \lambda w(z_1) + (1 - \lambda)w(z_2). \end{aligned} \quad (40)$$

By this, $w(z)$ is convex. ■

We now prove Lemma 1. Since $q(y, z, y')$ is the integral of a non-negative and non-decreasing function g , it is easy to show that the function $q(y, \cdot, y')$ is non-negative and convex. Hence, the conditional expectation $\mathbb{E}[q(y, \cdot, Y')|Y = y]$ is also convex. We can obtain

$$\begin{aligned} & \mathbb{E}[q(Y, z(Y), Y')] \\ &= \int_0^\infty \mathbb{E}[q(y, z(y), Y')|Y = y] d\mu_Y(y). \end{aligned} \quad (41)$$

According to Lemma 4, $\mathbb{E}[q(Y, z(Y), Y')]$ is a convex functional of z and $\mathbb{E}[Y + z(Y)]$ is an affine functional of z . It is known that the ratio of a non-negative convex functional and positive affine functional is quasi-convex [58, p. 103]. Hence, $h(z)$ is quasi-convex, which completes the proof.

APPENDIX D PROOF OF THEOREM 3

We use the Lagrangian duality approach to solve Problem (15). The Lagrangian of Problem (15) is

$$\begin{aligned} & L(z, \zeta, \gamma, \tau) \\ &= \int_0^\infty \mathbb{E}[q(y, z(y), Y')|Y = y] d\mu_Y(y) \\ &\quad - c \int_0^\infty [y + z(y)] d\mu_Y(y) \\ &\quad + \zeta \left[\frac{1}{f_{\max}} - \int_0^\infty [y + z(y)] d\mu_Y(y) \right] \\ &\quad - \int_0^\infty \gamma(y) z(y) d\mu_Y(y) + \int_0^\infty \tau(y) (z(y) - M) d\mu_Y(y) \\ &= \int_0^\infty \left\{ \mathbb{E}[q(y, z(y), Y')|Y = y] - (c + \zeta)[y + z(y)] \right. \\ &\quad \left. - \gamma(y) z(y) + \tau(y) [z(y) - M] \right\} d\mu_Y(y) + \zeta \frac{1}{f_{\max}}. \end{aligned} \quad (42)$$

Since Problem (15) is feasible, all constraints are affine, the refined Slater's condition [58, Sec. 5.2.3] is satisfied. According to [59, Proposition 3.3.2] and [60, pp. 70-72], the Karush-Kuhn-Tucker (KKT) theorem remains valid for the Lebesgue space $L^2(\mu_Y)$. Hence, if a vector (z, ζ, γ, τ) satisfies the KKT conditions (43)-(49), it is an optimal solution to (15). The KKT conditions are given by

$$z = \min_{x \in L^2(\mu_Y)} L(x, \zeta, \gamma, \tau), \quad (43)$$

$$\zeta \geq 0, \int_0^\infty [y + z(y)] d\mu_Y(y) \geq \frac{1}{f_{\max}}, \quad (44)$$

$$\gamma(y) \geq 0, z(y) \geq 0, \forall y \geq 0, \quad (45)$$

$$\tau(y) \geq 0, z(y) \leq M, \forall y \geq 0, \quad (46)$$

$$\zeta \left[\frac{1}{f_{\max}} - \int_0^\infty [y + z(y)] d\mu_Y(y) \right] = 0, \quad (47)$$

$$\gamma(y) z(y) = 0, \forall y \geq 0, \quad (48)$$

$$\tau(y) (z(y) - M) = 0, \forall y \geq 0. \quad (49)$$

We now solve the KKT conditions (43)-(49) by using the calculus of variations. The one-sided Gâteaux derivative (similar to sub-gradient in finite dimensional space) of a functional h in the direction of $w \in L^2(\mu_Y)$ at $z \in L^2(\mu_Y)$ is defined as

$$\delta h(z; w) \triangleq \lim_{\epsilon \rightarrow 0^+} \frac{h(z + \epsilon w) - h(z)}{\epsilon}. \quad (50)$$

If h is a function on \mathbb{R} , then (50) becomes the common one-sided derivative. Let $l(z, y, \zeta, \gamma, \tau)$ denote the integrand in (42), and $r(z, y) = \mathbb{E}[q(y, z(y), Y')|Y = y]$. According to Lemma 4, the function $q(y, z, y')$ and functionals $r(z, y)$, $l(z, y, \zeta, \gamma, \tau)$, and $L(z, \zeta, \gamma, \tau)$ are all convex in z . Therefore, their one-sided Gâteaux derivatives with respect to z exist [63, p. 709]. Since $g(x)$ is right-continuous, for any given (y, y') , the one-sided derivative $\delta q(y, z; w, y')$ of function $q(y, z, y')$ with respect to z is given by

$$\begin{aligned} & \delta q(y, z; w, y') \\ &= \lim_{\epsilon \rightarrow 0^+} \frac{1}{\epsilon} \{q(y, z + \epsilon w, y') - q(y, z, y')\} \\ &= \begin{cases} \lim_{x \rightarrow z^+} g(y + x + y') w, & \text{if } w \geq 0; \\ \lim_{x \rightarrow z^-} g(y + x + y') w, & \text{if } w < 0. \end{cases} \end{aligned}$$

Next, consider the one-sided Gâteaux derivative $\delta r(z; w, y)$ of functional $r(z, y)$. Since the function $g : [0, \infty) \rightarrow [0, \infty)$ is non-decreasing, $z \rightarrow q(y, z, y')$ is convex and finite for all $z \in [0, M]$. Hence, the function $\epsilon \rightarrow [q(y, z + \epsilon w, y') - q(y, z, y')]/\epsilon$ is non-decreasing and bounded from above on $(0, a]$ for some $a > 0$ [59, Proposition 1.1.2(i)]. By using the monotone convergence theorem [62, Theorem 1.5.6], we can interchange the limit and integral operators in $\delta r(z; w, y)$ such

that

$$\begin{aligned}
& \delta r(z; w, y) \\
&= \lim_{\epsilon \rightarrow 0^+} \frac{1}{\epsilon} \mathbb{E} [q(y, z(y) + \epsilon w(y), Y') - q(y, z(y), Y') | Y = y] \\
&= \mathbb{E} \left[\lim_{\epsilon \rightarrow 0^+} \frac{1}{\epsilon} \{q(y, z(y) + \epsilon w(y), Y') - q(y, z(y), Y')\} \middle| Y = y \right] \\
&= \mathbb{E} \left[\lim_{x \rightarrow z(y)^+} g(y + x + Y') w(y) 1_{\{w(y) > 0\}} \right. \\
&\quad \left. + \lim_{x \rightarrow z(y)^-} g(y + x + Y') w(y) 1_{\{w(y) < 0\}} \middle| Y = y \right] \\
&= \lim_{x \rightarrow z(y)^+} \mathbb{E} \left[g(y + x + Y') w(y) 1_{\{w(y) > 0\}} \middle| Y = y \right] \\
&\quad + \lim_{x \rightarrow z(y)^-} \mathbb{E} \left[g(y + x + Y') w(y) 1_{\{w(y) < 0\}} \middle| Y = y \right], \quad (51)
\end{aligned}$$

where 1_E is the indicator function of event E . By using the monotone convergence theorem again, we have

$$\begin{aligned}
& \delta L(z; w, \zeta, \gamma, \tau) \\
&= \int_0^\infty \delta l(z; w, y, \zeta, \gamma, \tau) d\mu_Y(y) \\
&= \int_0^\infty \delta r(z; w, y) d\mu_Y(y) \\
&\quad + \int_0^\infty [-(c + \zeta) - \gamma(y) + \tau(y)] w(y) d\mu_Y(y). \quad (52)
\end{aligned}$$

According to [63, p. 710], z is an optimal solution to (43) if and only if

$$\delta L(z; w, \zeta, \gamma, \tau) \geq 0, \quad \forall w \in L^2(\mu_Y). \quad (53)$$

Since $w(\cdot)$ is an arbitrary function in $L^2(\mu_Y)$, considering positive functions $w(y) > 0$, we can obtain from (51)-(53) that for each $y \geq 0$, $z(y)$ must satisfy

$$\lim_{x \rightarrow z(y)^+} \mathbb{E} [g(y + x + Y') | Y = y] - (c + \zeta) - \gamma(y) + \tau(y) \geq 0. \quad (54)$$

Similarly, considering negative functions $w(y) < 0$, we can obtain that for each $y \geq 0$, $z(y)$ must satisfy

$$\lim_{x \rightarrow z(y)^-} \mathbb{E} [g(y + x + Y') | Y = y] - (c + \zeta) - \gamma(y) + \tau(y) \leq 0. \quad (55)$$

Because $g(\cdot)$ is non-decreasing, we can obtain from (54) and (55) that for each $y \geq 0$, $z(y)$ needs to satisfy

$$\mathbb{E} [g(y + x + Y') | Y = y] - (c + \zeta) - \gamma(y) + \tau(y) \geq 0 \quad (56)$$

for all $x > z(y)$, and

$$\mathbb{E} [g(y + x + Y') | Y = y] - (c + \zeta) - \gamma(y) + \tau(y) \leq 0 \quad (57)$$

for all $x < z(y)$.

We solve the optimal primal solution $z(\cdot)$ by considering the following three cases:

Case 1: $\gamma(y) = \tau(y) = 0$. The solutions to (56) and (57) may not be unique, if function g is not strictly increasing. In particular, there may exist an interval $[a(y), b(y)]$ such that each $z(y) \in [a(y), b(y)]$ satisfies (56) and (57) for some y .

In this case, we choose the largest possible solution of $z(y)$ to make sure that the constraint (14) is satisfied. The largest solution satisfying (56) and (57) is given by

$$z(y) = \sup\{x \in [0, M] : \mathbb{E} [g(y + x + Y') | Y = y] \leq c + \zeta\}, \quad \forall y \geq 0,$$

which is exactly (17).

Case 2: $\gamma(y) > 0$. By (48), we have $z(y) = 0$.

Case 3: $\tau(y) > 0$. By (49), we have $z(y) = M$.

In summary, the optimal primal solution $z(\cdot)$ is given by (17).

Next, we find the optimal dual variable ζ . By (44) and (47), the optimal ζ satisfies

$$\zeta = 0, \mathbb{E} [Y + z(Y)] \geq \frac{1}{f_{\max}} \quad (58)$$

or

$$\zeta > 0, \mathbb{E} [Y + z(Y)] = \frac{1}{f_{\max}}, \quad (59)$$

where $\mathbb{E} [Y + z(Y)]$ is determined by the optimal primal solution (17). Since $\mathbb{E} [Y + z(Y)]$ is non-decreasing in ζ , we can use bisection to search for the optimal ζ . By this, an optimal solution to (15) is obtained for any given c . Finally, according to Sections 4.2.5 and 11.4 of [58], the optimal c is solved by an outer-layer bisection search. Therefore, an optimal solution to Problem (13) is given by Algorithm 1. This completes the proof.

APPENDIX E PROOF OF LEMMA 2

Let us rewrite the functional h_1 as

$$h_1(z) = \frac{\int_0^\infty [y + z(y)]^2 d\mu_Y(y)}{\int_0^\infty [y + z(y)] d\mu_Y(y)}.$$

We need to prove that the functional h_1 is convex when restricted to any line that intersects its domain. For any $w \in L^2(\mu_Y)$, consider the function $u : \mathbb{R} \rightarrow \mathbb{R}$ defined as

$$u(\epsilon) = \frac{\int_0^\infty [z(y) + \epsilon w(y) + y]^2 d\mu_Y(y)}{\int_0^\infty [z(y) + \epsilon w(y) + y] d\mu_Y(y)}$$

with domain

$$\text{dom } u = \{\epsilon : z(y) + \epsilon w(y) \in [0, M], \forall y \geq 0, \epsilon \in \mathbb{R}\}.$$

Since the function $\epsilon \rightarrow [z(y) + \epsilon w(y) + y]^2$ is convex, the function $x \rightarrow \{[z(y) + (\epsilon + x)w(y) + y]^2 - [z(y) + \epsilon w(y) + y]^2\}/x$ is non-decreasing and bounded from above on $(0, a]$ for some $a > 0$. By using monotone convergence theorem [62, Theorem 1.5.6], we can interchange the limit and integral operators such that

$$\begin{aligned}
& \frac{d}{d\epsilon} \int_0^\infty [z(y) + \epsilon w(y) + y]^2 d\mu_Y(y) \\
&= \int_0^\infty 2[z(y) + \epsilon w(y) + y] w(y) d\mu_Y(y).
\end{aligned}$$

Similarly,

$$\frac{d}{d\epsilon} \int_0^\infty [z(y) + \epsilon w(y) + y] d\mu_Y(y) = \int_0^\infty w(y) d\mu_Y(y).$$

By this, we have

$$\begin{aligned} \frac{du}{d\epsilon} &= \frac{\int_0^\infty 2[z(y) + \epsilon w(y) + y] w(y) d\mu_Y(y)}{\int_0^\infty [z(y) + \epsilon w(y) + y] d\mu_Y(y)} \\ &\quad - \frac{\int_0^\infty [z(y) + \epsilon w(y) + y]^2 d\mu_Y(y) \int_0^\infty w(y) d\mu_Y(y)}{[\int_0^\infty [z(y) + \epsilon w(y) + y] d\mu_Y(y)]^2}. \end{aligned} \quad (60)$$

After some additional manipulations, we can obtain

$$\begin{aligned} \frac{d^2 u}{d^2 \epsilon} &= \frac{2 [\int_0^\infty [y + z(y)] d\mu_Y(y) \int_0^\infty w(y) d\mu_Y(y)]^2}{[\int_0^\infty [z(y) + \epsilon w(y) + y] d\mu_Y(y)]^3} \\ &\quad \times \int_0^\infty \left[\frac{y + z(y)}{\int_0^\infty [y + z(y)] d\mu_Y(y)} - \frac{w(y)}{\int_0^\infty w(y) d\mu_Y(y)} \right]^2 d\mu_Y(y). \end{aligned}$$

Since $z(y) + \epsilon w(y) \geq 0$ for all y on **dom** u , we have $\frac{d^2 u}{d^2 \epsilon} \geq 0$. Hence, the function u is convex for all $w \in L^2(\mu_Y)$. By this, the functional h_1 is convex, which completes the proof.

APPENDIX F PROOF OF THEOREM 4

The Lagrangian of Problem (19) is determined as

$$\begin{aligned} L_1(z, \zeta, \gamma, \tau) &= \frac{\int_0^\infty [y + z(y)]^2 d\mu_Y(y)}{2 \int_0^\infty [y + z(y)] d\mu_Y(y)} + \zeta \left[\frac{1}{f_{\max}} - \int_0^\infty [y + z(y)] d\mu_Y(y) \right] \\ &\quad - \int_0^\infty \gamma(y) z(y) d\mu_Y(y) + \int_0^\infty \tau(y) (z(y) - M) d\mu_Y(y), \end{aligned}$$

where $\zeta \in \mathbb{R}$, $\gamma, \tau \in L^2(\mu_Y)$ are dual variables. According to [59, Proposition 3.3.2] and [60, pp. 70-72], the KKT theorem remains valid for the Lebesgue space $L^2(\mu_Y)$. Hence, if a vector (z, ζ, γ, τ) satisfies the KKT conditions (61)-(67), it is an optimal solution to (19). The KKT conditions are given by:

$$z = \min_{x \in L^2(\mu_Y)} L_1(x, \zeta, \gamma, \tau), \quad (61)$$

$$\zeta \geq 0, \int_0^\infty [y + z(y)] d\mu_Y(y) \geq \frac{1}{f_{\max}}, \quad (62)$$

$$\gamma(y) \geq 0, z(y) \geq 0, \forall y \geq 0, \quad (63)$$

$$\tau(y) \geq 0, z(y) \leq M, \forall y \geq 0, \quad (64)$$

$$\zeta \left[\frac{1}{f_{\max}} - \int_0^\infty [y + z(y)] d\mu_Y(y) \right] = 0, \quad (65)$$

$$\gamma(y) z(y) = 0, \forall y \geq 0, \quad (66)$$

$$\tau(y) (z(y) - M) = 0, \forall y \geq 0. \quad (67)$$

We now solve the KKT conditions by using the calculus of variations. For any fixed (ζ, γ, τ) , the Gâteaux derivative of the Lagrange L_1 in the direction of $w \in L^2(\mu_Y)$ at $z \in L^2(\mu_Y)$ is defined as

$$\delta L_1(z; w, \zeta, \gamma, \tau) \triangleq \lim_{\epsilon \rightarrow 0} \frac{L_1(z + \epsilon w, \zeta, \gamma, \tau) - L_1(z, \zeta, \gamma, \tau)}{\epsilon}.$$

Similar to the derivations of (60), we can obtain

$$\begin{aligned} &\delta L_1(z; w, \zeta, \gamma, \tau) \\ &= \int_0^\infty \left[\frac{y + z(y)}{\int_0^\infty [y + z(y)] d\mu_Y(y)} - \frac{\int_0^\infty [y + z(y)]^2 d\mu_Y(y)}{2 [\int_0^\infty [y + z(y)] d\mu_Y(y)]^2} \right. \\ &\quad \left. - \zeta - \gamma(y) + \tau(y) \right] w(y) d\mu_Y(y), \quad \forall w \in L^2(\mu_Y). \end{aligned}$$

Then, $z(\cdot)$ is an optimal solution to (61) if and only if [63, p. 710]

$$\delta L_1(z; w, \zeta, \gamma, \tau) \geq 0, \quad \forall w \in L^2(\mu_Y).$$

By $\delta L_1(z; w, \zeta, \gamma, \tau) = -\delta L_1(z; -w, \zeta, \gamma, \tau)$, we deduce

$$\delta L_1(z; w, \zeta, \gamma, \tau) = 0, \quad \forall w \in L^2(\mu_Y).$$

Since $w(\cdot)$ is arbitrary, we have

$$\begin{aligned} &\frac{y + z(y)}{\int_0^\infty [y + z(y)] d\mu_Y(y)} - \frac{\int_0^\infty [y + z(y)]^2 d\mu_Y(y)}{2 [\int_0^\infty [y + z(y)] d\mu_Y(y)]^2} \\ &\quad - \zeta - \gamma(y) + \tau(y) = 0, \quad \forall y \geq 0. \end{aligned} \quad (68)$$

For notational simplicity, let us define

$$\beta \triangleq \zeta \int_0^\infty [y + z(y)] d\mu_Y(y) + \frac{\int_0^\infty [y + z(y)]^2 d\mu_Y(y)}{2 \int_0^\infty [y + z(y)] d\mu_Y(y)}. \quad (69)$$

Since $\mathbb{E}[Y] > 0$, we have $\beta > 0$. The optimal primal solution $z(\cdot)$ is obtained by considering the following three cases:

Case 1: If $\gamma(y) = \tau(y) = 0$, then by (68) and (69), we obtain $z(y) = \beta - y$. In this case, we require $\beta - y \in [0, M]$ by (63) and (64).

Case 2: If $\gamma(y) > 0$, then by (66), $z(y) = 0$.

Case 3: If $\tau(y) > 0$, then by (67), $z(y) = M$.

In summary, the optimal primal solution $z(\cdot)$ is given by (21).

The optimal dual variable β is obtained by considering two cases:

Case 1: $\zeta > 0$. Then, (65) and (69) imply that

$$\int_0^\infty [y + z(y)] d\mu_Y(y) = \frac{1}{f_{\max}}, \beta \geq \frac{\int_0^\infty [y + z(y)]^2 d\mu_Y(y)}{2 \int_0^\infty [y + z(y)] d\mu_Y(y)}. \quad (70)$$

Case 2: $\zeta = 0$. Then, (62) and (69) imply that

$$\int_0^\infty [y + z(y)] d\mu_Y(y) \geq \frac{1}{f_{\max}}, \beta = \frac{\int_0^\infty [y + z(y)]^2 d\mu_Y(y)}{2 \int_0^\infty [y + z(y)] d\mu_Y(y)}. \quad (71)$$

Combining (70) and (71), yields

$$\int_0^\infty [y + z(y)] d\mu_Y(y) = \max \left(\frac{1}{f_{\max}}, \frac{\int_0^\infty [z(y) + y]^2 d\mu_Y(y)}{2\beta} \right).$$

Then, (22) is obtained. This completes the proof.

APPENDIX G PROOF OF THEOREM 5

Since $\mathbb{E}[Y] \geq \frac{1}{f_{\max}}$, the constraint (20) is always satisfied and can be removed. If the zero-wait policy is optimal,

we solve the KKT conditions of Problem (19) without the constraint (20). By this, we can obtain that the optimal primal solution is given by (21) almost everywhere, and the optimal dual solution β must satisfy

$$\beta \leq y_{\text{inf}}, \quad \mathbb{E}[Y^2] = 2\beta\mathbb{E}[Y], \quad (72)$$

from which (25) follows.

Next, we prove the reverse direction. If (25) holds, by $\mathbb{E}[Y] > 0$, we can get $\mathbb{E}[Y^2] > 0$ and $y_{\text{inf}} > 0$. By (25) and choosing $\beta = \frac{\mathbb{E}[Y^2]}{2\mathbb{E}[Y]} > 0$, we obtain (72). Substituting (72) and $\mathbb{E}[Y] \geq \frac{1}{f_{\text{max}}}$ into Theorem 4, yields that the zero-wait policy is optimal. This completes the proof.

APPENDIX H PROOF OF LEMMA 3

1). When the correlation coefficient between Y_i and Y_{i+1} is -1 , $Y + Y'$ is equal a constant value with probability one. Choosing $z(y) = 0$, $\zeta = 0$, $c = g(Y + Y')$, $\gamma(y) = \tau(y) = 0$, one can show that the KKT conditions (43)-(49) are satisfied.

2). If the Y_i 's are equal to a constant value, $Y + Y'$ is equal a constant value. The remaining proof follows from part 1).

3). When $g(\cdot)$ is a constant function, all policies are optimal. This completes the proof.

REFERENCES

- [1] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," in *Proc. IEEE INFOCOM*, 2016.
- [2] A. R. Al-Ali, I. Zuakernan, and F. Aloul, "A mobile GPRS-sensors array for air pollution monitoring," *IEEE Sensors Journal*, vol. 10, no. 10, pp. 1666–1671, Oct 2010.
- [3] T. Gandhi and M. M. Trivedi, "Pedestrian protection systems: Issues, survey, and challenges," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 3, pp. 413–430, Sept 2007.
- [4] V. Terzija, G. Valverde, D. Cai, P. Regulski, V. Madani, J. Fitch, S. Skok, M. M. Begovic, and A. Phadke, "Wide-area monitoring, protection, and control of future electric power networks," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 80–93, Jan 2011.
- [5] X. Song and J. W. S. Liu, "Performance of multiversion concurrency control algorithms in maintaining temporal consistency," in *Fourteenth Annual International Computer Software and Applications Conference*, Oct 1990, pp. 132–139.
- [6] S. Kaul, R. D. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. IEEE INFOCOM*, Apr. 2012, pp. 2731–2735.
- [7] —, "Status updates through queues," in *Conf. on Info. Sciences and Systems*, Mar. 2012.
- [8] R. D. Yates and S. Kaul, "Real-time status updating: Multiple sources," in *IEEE ISIT*, Jul. 2012.
- [9] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," CoRR, abs/1608.08622, submitted to *IEEE Trans. Inf. Theory*, 2016.
- [10] C. Kam, S. Kompella, and A. Ephremides, "Age of information under random updates," in *IEEE ISIT*, 2013.
- [11] —, "Effect of message transmission diversity on status age," in *IEEE ISIT*, June 2014, pp. 2411–2415.
- [12] C. Kam, S. Kompella, G. D. Nguyen, and A. Ephremides, "Effect of message transmission path diversity on status age," *IEEE Trans. Inf. Theory*, vol. 62, no. 3, pp. 1360–1374, March 2016.
- [13] M. Costa, M. Codreanu, and A. Ephremides, "Age of information with packet management," in *IEEE ISIT*, June 2014, pp. 1583–1587.
- [14] —, "On the age of information in status update systems with packet management," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1897–1910, April 2016.
- [15] N. Pappas, J. Gunnarsson, L. Kratz, M. Kountouris, and V. Angelakis, "Age of information of multiple sources with queue management," in *2015 IEEE ICC*, 2015.
- [16] L. Huang and E. Modiano, "Optimizing age-of-information in a multi-class queueing system," in *IEEE ISIT*, 2015.
- [17] C. Kam, S. Kompella, G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, "Age of information with a packet deadline," in *IEEE ISIT*, July 2016, pp. 2564–2568.
- [18] K. Chen and L. Huang, "Age-of-information in the presence of error," in *IEEE ISIT*, July 2016, pp. 2579–2583.
- [19] E. Najm and R. Nasser, "Age of information: The gamma awakening," in *IEEE ISIT*, July 2016, pp. 2574–2578.
- [20] A. M. Bedewy, Y. Sun, and N. B. Shroff, "Optimizing data freshness, throughput, and delay in multi-server information-update systems," in *IEEE ISIT*, July 2016, pp. 2569–2573.
- [21] —, "Age-optimal information updates in multihop networks," in *IEEE ISIT*, 2017.
- [22] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *IEEE ISIT*, 2015.
- [23] L. Kleinrock, *Queueing Systems*. New York: Wiley-Interscience, 1975.
- [24] A. Segev and W. Fang, "Optimal update policies for distributed materialized views," *Manage. Sci.*, vol. 37, no. 7, pp. 851–870, Jul. 1991.
- [25] B. Adelberg, H. Garcia-Molina, and B. Kao, "Applying update streams in a soft real-time database system," in *Proc. ACM SIGMOD*, 1995, pp. 245–256.
- [26] J. Cho and H. Garcia-Molina, "Synchronizing a database to improve freshness," in *Proc. ACM SIGMOD*, 2000, pp. 117–128.
- [27] C. Kam, S. Kompella, G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, "Controlling the age of information: Buffer size, deadline, and packet replacement," in *IEEE MILCOM*, Nov 2016, pp. 301–306.
- [28] B. T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu, "Age of information under energy replenishment constraints," in *ITA*, 2015.
- [29] B. T. Bacinoglu and E. Uysal-Biyikoglu, "Scheduling status updates to minimize age of information with an energy harvesting sensor," in *IEEE ISIT*, 2017.
- [30] J. Zhong and R. D. Yates, "Timeliness in lossless block coding," in *Data Compression Conference (DCC)*, March 2016.
- [31] I. Kadota, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Minimizing the age of information in broadcast wireless networks," in *Allerton Conference*, Sept 2016, pp. 844–851.
- [32] Q. He, D. Yuan, and A. Ephremides, "Optimizing freshness of information: On minimum age link scheduling in wireless systems," in *WiOpt*, May 2016, pp. 1–8.
- [33] —, "On optimal link scheduling with min-max peak age of information in wireless systems," in *IEEE ICC*, May 2016, pp. 1–7.
- [34] M. Costa, S. Valentin, and A. Ephremides, "On the age of channel state information for non-reciprocal wireless links," in *IEEE ISIT*, June 2015, pp. 2356–2360.
- [35] —, "On the age of channel information for a finite-state Markov model," in *IEEE ICC*, June 2015, pp. 4101–4106.
- [36] C. Kam, S. Kompella, and A. Ephremides, "Experimental evaluation of the age of information via emulation," in *IEEE MILCOM*, Oct 2015, pp. 1070–1075.
- [37] S. Shalev-Shwartz, "Online learning and online convex optimization," *Foundations and Trends in Machine Learning*, vol. 4, no. 2, pp. 107–194, 2012.
- [38] X. He, J. Pan, O. Jin, T. Xu, B. Liu, T. Xu, Y. Shi, A. Atallah, R. Herbrich, S. Bowers, and J. Q. Candela, "Practical lessons from predicting clicks on Ads at Facebook," in *Proceedings of the Eighth International Workshop on Data Mining for Online Advertising*, 2014.
- [39] J. Chiang and J. Yuan, "Optimal maintenance policy for a Markovian system under periodic inspection," *Reliability Engineering & System Safety*, vol. 71, no. 2, pp. 165 – 172, 2001.
- [40] V. A. Kottapalli, A. S. Kiremidjian, J. P. Lynch, E. Carryer, T. W. Kenny, K. H. Law, and Y. Lei, "Two-tiered wireless sensor network architecture for structural health monitoring," in *Proc. SPIE*, 2003.
- [41] S. M. Ross, *Applied Probability Models with Optimization Applications*. San Francisco, CA: Holden-Day, 1970.
- [42] H. Mine and S. Osaki, *Markovian Decision Processes*. New York: Elsevier, 1970.
- [43] D. Hayman and M. Sobel, *Stochastic models in Operations Research, Volume II: Stochastic Optimizations*. New York: McGraw-Hill, 1984.
- [44] E. A. Feinberg, "Constrained semi-Markov decision processes with average rewards," *Zeitschrift für Operations Research*, vol. 39, no. 3, pp. 257–288, 1994.
- [45] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Belmont, MA: Athena Scientific, 2005, vol. 1.
- [46] R. Bellman, *Dynamic Programming*. Princeton University Press, 1957.

- [47] M. J. Neely, "Dynamic optimization and learning for renewal systems," *IEEE Trans. Auto. Control*, vol. 58, no. 1, pp. 32–46, Jan 2013.
- [48] A. Federgruen, P. J. Schweitzer, and H. C. Tijms, "Denumerable undiscounted semi-Markov decision processes with unbounded rewards," *Mathematics of Operations Research*, vol. 8, no. 2, pp. 298–313, 1983.
- [49] D. Klabjan and D. Adelman, "Existence of optimal policies for semi-Markov decision processes using duality for infinite linear programming," *SIAM J. Control Optim.*, vol. 44, no. 6, pp. 2104–2122, Jan. 2006.
- [50] F. J. Beutler and K. W. Ross, "Optimal policies for controlled Markov chains with a constraint," *Journal of Mathematical Analysis and Applications*, vol. 112, no. 1, pp. 236 – 252, 1985.
- [51] —, "Time-average optimal constrained semi-Markov decision processes," *Adv. in Appl. Probab.*, vol. 18, no. 2, pp. 341–359, 1986.
- [52] M. Baykal-Gürsoy and K. Gürsoy, "Semi-Markov decision processes: Nonstandard criteria," *Probability in the Engineering and Information Sciences*, vol. 21, pp. 635–657, Oct. 2007.
- [53] E. A. Feinberg, "Constrained discounted semi-Markov decision processes," in *Markov Processes and Controlled Markov Chains*, Z. Hou, J. A. Filar, and A. Chen, Eds. Kluwer Academic Publishers, 2002.
- [54] E. Altman, *Constrained Markov decision processes*. Boca Raton, FL, USA: Chapman & Hall/CRC, 1999.
- [55] C. Derman, *Finite State Markovian Decision Processes*. Orlando, FL, USA: Academic Press, Inc., 1970.
- [56] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*. Morgan & Claypool, 2010.
- [57] W. Rudin, *Real and Complex Analysis*, 3rd ed. New York: McGraw-Hill Book Co., 1987.
- [58] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.
- [59] D. Butnariu and A. N. Iusem, *Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization*. Norwell, MA: Kluwer Academic Publisher, 2000.
- [60] G. G. Magaril-Il'yayev and V. M. Tikhomirov, *Convex Analysis: Theory and Applications*. American Mathematical Society, 2003.
- [61] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the Wiener process over a channel with random delay," in *IEEE ISIT*, 2017.
- [62] R. Durrett, *Probability: Theory and Examples*, 4th ed. Cambridge University Press, 2010.
- [63] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.