Optimal Sampling and Scheduling for Timely Status Updates in Multi-source Networks

Ahmed M. Bedewy, Yin Sun, *Member, IEEE*, Sastry Kompella, *Senior Member, IEEE*, and Ness B. Shroff, *Fellow, IEEE*

Abstract

We consider a joint sampling and scheduling problem for optimizing data freshness in multi-source systems. Data freshness is measured by a non-decreasing penalty function of the *age of information*, where all sources have the same age-penalty function. Sources take turns to generate update packets, and forward them to their destinations one-by-one through a shared channel with random delay. There is a scheduler, that chooses the update order of the sources, and a sampler, that determines when a source should generate a new packet in its turn. We aim to find the optimal scheduler-sampler pairs that minimize the total-average peak age-penalty (Ta-PAP) and the total-average age-penalty (Ta-AP). We prove that the Maximum Age First (MAF) scheduler and the zero-wait sampler are jointly optimal for minimizing the Ta-PAP. Meanwhile, the MAF scheduler and a relative value iteration with reduced complexity (RVI-RC) sampler are jointly optimal for minimizing the Ta-AP. The RVI-RC sampler is based on a relative value iteration algorithm whose complexity is reduced by exploiting a threshold property in the optimal sampler. Finally, a low-complexity threshold-type sampler is devised via an approximate analysis of Bellman's equation. This threshold-type sampler reduces to a simple water-filling sampler for a linear age-penalty function.
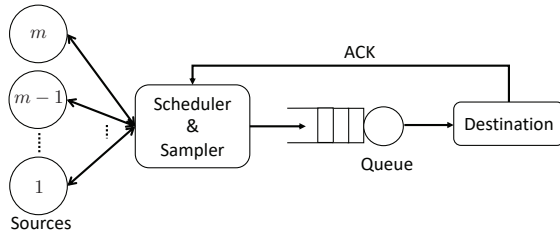
Figure 1: System model

I. INTRODUCTION

In recent years, significant attention has been paid to the *age of information* as a metric for data freshness. This is because there are a growing number of applications that require timely status updates in various networked monitoring and control systems. Examples include sensor and environment monitoring networks, surrounding monitoring autonomous vehicles, smart grid systems, etc. The age of information, or simply age, was introduced in [2]–[5], which is the time elapsed since the most recently received update was generated. Unlike traditional packet-based metrics, such as throughput and delay, age is a destination-based metric that captures the information lag at the destination, and is hence more apt for achieving the goal of timely updates.

There have been two major lines of research on age in single source networks: One direction is on systems with a stochastic arrival process. There are results on both queueing-based age analysis [5]–[8] and sample-path based age optimization [9]–[12]. The second direction is for the case that the packet arrival process is designable [13]–[16], where our study extends the findings in these studies to multi-source networks.

We consider random, yet discrete, transmission times such that a packet has to be processed for a random period before delivered to the destination. In practice, such random transmission times occur in many applications, such as autonomous vehicles. In particular, there are many electronic control units (ECUs) in a vehicle, that are connected to one or more sensors and actuators via a controller area network (CAN) bus [17], [18]. These ECUs are given different priority, based on their criticality level (e.g., ECUs in the powertrain have a higher priority compared to those connected to infotainment systems). Since high priority packets usually have hard deadlines, the transmissions of low priority packets are interrupted whenever the higher priority ones are transmitted. Therefore, information packets with lower priority see a time-varying bandwidth, and hence encounter a random transmission time.

When the transmission time is highly random, one can observe an interesting phenomenon: it is not necessarily optimal to generate a new packet as soon as the channel becomes available. This phenomenon was revealed in [13] and further explored in [15] and [16]. In this paper, our goal is to investigate timely status updates in multi-source systems with random transmission times, as depicted in Fig. 1. Sources take turns to generate update packets, and forward the packets to their destinations one-by-one through a shared channel with random delay. This results in a joint design problem of scheduling and sampling, where the scheduler chooses the update order of the sources, and the sampler determines when a source should generate a new packet in its turn. We find that it is optimal to first serve the source with the

highest age, and, similar to the single-user case, it is not always optimal to generate packets as soon as the channel becomes available. To that end, the main contributions of this paper are outlined as follows:

- We formulate the optimal scheduling and sampling problem to optimize data freshness in single-hop, multi-source networks. We use a non-decreasing age-penalty function to represent the level of dissatisfaction of data staleness, where all sources have the same age-penalty function. We focus on minimizing the total-average peak age-penalty (Ta-PAP) and the total-average age-penalty (Ta-AP), where Ta-AP is more challenging to minimize. We show that our optimization problem has an important *separation principle*: For any given sampler, we show that the optimal scheduling policy is the Maximum Age First (MAF) scheduler (Proposition 1). Hence, the optimal scheduler-sampler pair can be obtained by fixing the scheduling policy to the MAF scheduler, and then optimize the sampler design separately.

- We show that the MAF scheduler and zero-wait sampler, in which a new packet is generated once the channel becomes idle, are jointly optimal for minimizing the Ta-PAP (Theorem 2). We show this result by proving the optimality of the zero-wait sampler for minimizing the Ta-PAP, when the scheduling policy is fixed to the MAF scheduler.

- Interestingly, we find that zero-wait sampler does not always minimize the Ta-AP, when the MAF scheduler is employed. We show that the MAF scheduler and the relative value iteration with reduced complexity (RVI-RC) sampler are jointly optimal for minimizing the Ta-AP (Theorem 6). We take several steps to prove the optimality of the RVI-RC sampler: When the scheduling policy is fixed to the MAF scheduler, we reformulate the optimal sampling problem for minimizing the Ta-AP as an equivalent semi-Markov decision problem. We use Dynamic Programming (DP) to obtain the optimal sampler. In particular, we show that there exists a stationary deterministic sampler that can achieve optimality (Proposition 4). We also show that the optimal sampler has a threshold property (Proposition 5), that helps in reducing the complexity of the relative value iteration (RVI) algorithm (by reducing the computations required for some system states). This results in the RVI-RC sampler in Algorithm 1.

- Finally, in Section IV, we devise a low-complexity threshold-type sampler via an approximate analysis of Bellman's equation whose solution is the RVI-RC sampler. In addition, for the special case of a linear age-penalty function, this threshold sampler is further simplified to the water-filling solution. The numerical results in Figs. 5-10 indicate that, when the scheduler is fixed to the MAF, the performance of these approximated samplers is almost the same as that of the RVI-RC sampler.

A. Related Works

Early studies have characterized the age in many interesting variants of queueing models, such as First-Come, First-Served (FCFS) [5], [8], [19], [20], Last-Come, First-Served (LCFS) with and without preemption [6], [21], and the queueing model with packet management [7], [22]. The update packets in these studies arrive at the queue randomly according to a Poisson process. The work in [9]–[12] showed that Last-Generated, First-Served (LGFS)-type policies are optimal or near-optimal for minimizing a large class of age metrics in single flow multi-server and multi-hop networks.

Another line of research has considered the "generate-at-will" model [13]–[16], in which the generation times (sampling times) of the update packets are controllable. The work in [15], [16] motivated the usage of nonlinear age functions from various real-time applications and designed sampling policies for optimizing nonlinear age functions in single source systems. Our study here extends the work in [15], [16] to a multi-source system. In this system, only one packet can be sent through the channel at a time. Therefore, a decision maker does not only consist of a sampler, but also a scheduler, which makes the problem even more challenging.

The scheduling problem for multi-source networks with different scenarios was considered in [23]–[36]. In [24], the authors found that the scheduling problem for minimizing the age in wireless networks under physical interference constraints is NP-hard. Optimal scheduling for age minimization in a broadcast network was studied in [25]–[29], where a single source can be scheduled at a time. In addition, it was found that a maximum age first (MAF) service discipline is useful for reducing the age in various multi-source systems with different service time distributions in [25]–[27], [30], [31]. In contrast to our study, the generation of the update packets in [24]–[31] is uncontrollable and they arrive randomly at the transmitter. Age analysis of the status updates over a multiaccess channel was considered in [32]. The studies in [33]–[36] considered the age optimization problem in a wireless network with general interference constraints and channel uncertainty. Our results in Theorem 2 and Corollary 9 suggest that if (i) the packet transmission time is fixed as in time-slotted systems [24]–[30], [32] or (ii) the objective function is the peak age [33]–[36], then it is optimal to sample as soon as the channel becomes available. However, this is not necessarily true otherwise.

II. MODEL AND FORMULATION

A. Notation and Definitions

We use $\mathbb{N}^+$ to represent the set of non-negative integers, $\mathbb{R}^+$ is the set of non-negative real numbers, $\mathbb{R}$ is the set of real numbers, and $\mathbb{R}^n$ is the set of n-dimensional real Euclidean space. We use t^- to denote the time instant just before t. Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ be two vectors in $\mathbb{R}^n$, then we denote $\mathbf{x} \leq \mathbf{y}$ if $x_i \leq y_i$ for $i = 1, 2, \ldots, n$. Also, we use $x_{[i]}$ to denote the i-th largest component of vector $\mathbf{x}$. A set $U \subseteq \mathbb{R}^n$ is called upper if $\mathbf{y} \in U$ whenever $\mathbf{y} \geq \mathbf{x}$ and $\mathbf{x} \in U$. We will need the following definitions:

Definition 1. *Univariate Stochastic Ordering: [37] Let X and Y be two random variables. Then, X is said to be stochastically smaller than Y (denoted as $X \leq_{st} Y$), if*

$$\mathbb{P}\{X > x\} \leq \mathbb{P}\{Y > x\}, \quad \forall x \in \mathbb{R}.$$

Definition 2. *Multivariate Stochastic Ordering: [37] Let $\mathbf{X}$ and $\mathbf{Y}$ be two random vectors. Then, $\mathbf{X}$ is said to be stochastically smaller than $\mathbf{Y}$ (denoted as $\mathbf{X} \leq_{st} \mathbf{Y}$), if*

$$\mathbb{P}\{\mathbf{X} \in U\} \leq \mathbb{P}\{\mathbf{Y} \in U\}, \quad \text{for all upper sets} \quad U \subseteq \mathbb{R}^n.$$

Definition 3. *Stochastic Ordering of Stochastic Processes: [37] Let $\{X(t), t \in [0, \infty)\}$ and $\{Y(t), t \in [0, \infty)\}$ be two stochastic processes. Then, $\{X(t), t \in [0, \infty)\}$ is said to be stochastically smaller than*

$\{Y(t), t \in [0, \infty)\}$ *(denoted by* $\{X(t), t \in [0, \infty)\} \leq_{st} \{Y(t), t \in [0, \infty)\}$*), if, for all choices of an integer n and $t_1 < t_2 < \ldots < t_n$ in $[0, \infty)$, it holds that*

$$(X(t_1), X(t_2), \ldots, X(t_n)) \leq_{st} (Y(t_1), Y(t_2), \ldots, Y(t_n)), \tag{1}$$

where the multivariate stochastic ordering in (1) *was defined in Definition 2.*

B. System Model

We consider a status update system with m sources as shown in Fig. 1, where each source observes a time-varying process. An update packet is generated from a source and is then sent over an error-free delay channel to the destination, where only one packet can be sent at a time. A decision maker controls the transmission order of the sources and the generation times of the update packets for each source. This is known as the "generate-at-will" model [13]–[15] (i.e., the update packets can be generated at any time).

We use S_i to denote the generation time of the i-th generated packet from all sources, called packet i. Moreover, we use r_i to represent the source index from which packet i is generated. The channel is modeled as FCFS queue with random *i.i.d.* service time Y_i, where Y_i represents the service time of packet i, $Y_i \in \mathcal{Y}$, and $\mathcal{Y} \subset \mathbb{R}^+$ is a finite and bounded set. We also assume that $0 < \mathbb{E}[Y_i] < \infty$ for all i. We suppose that the decision maker knows the idle/busy state of the server through acknowledgments (ACKs) from the destination with zero delay. If an update packet is generated while the server is busy, this packet needs to wait in the queue until its transmission opportunity, and become stale while waiting. Hence, there is no loss of optimality to avoid generating an update packet during the busy periods. As a result, a packet is served immediately once it is generated. Let D_i denote the delivery time of packet i, where $D_i = S_i + Y_i$. After the delivery of packet i at time D_i, the decision maker may insert a waiting time Z_i before generating a new packet (hence, $S_{i+1} = D_i + Z_i$)[1], where $Z_i \in \mathcal{Z}$, and $\mathcal{Z} \subset \mathbb{R}^+$ is a finite and bounded set[2].

At any time t, the most recently delivered packet from source l is generated at time

$$U_l(t) = \max\{S_i : r_i = l, D_i \leq t\}. \tag{2}$$

The *age of information*, or simply the *age*, for source l is defined as [2]–[5]

$$\Delta_l(t) = t - U_l(t). \tag{3}$$

As shown in Fig. 2, the age increases linearly with t but is reset to a smaller value with the delivery of a fresher packet. We suppose that the age $\Delta_l(t)$ is right-continuous. The age process for source l is given by $\{\Delta_l(t), t \geq 0\}$. We suppose that the initial age value $\Delta_l(0^-)$ for all l are known to the system. For each source l, we consider an age-penalty function $g(\Delta_l(t))$ of the age $\Delta_l(t)$. The function $g : [0, \infty) \to \mathbb{R}$ is non-decreasing and is not necessarily convex or continuous. We suppose that $\mathbb{E}[|\int_a^{a+x} g(\tau)d\tau|] < \infty$ whenever $x < \infty$. It was recently shown in [16] that, under certain conditions, information freshness

---

[1] We suppose that $D_0 = 0$. Thus, we have $S_1 = Z_0$.

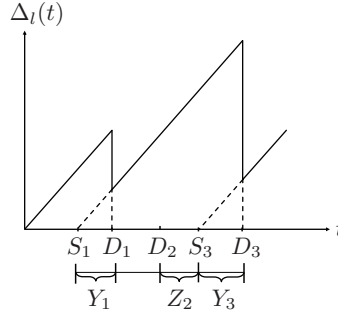[2] We suppose that we always have $0 \in \mathcal{Z}$.

Figure 2: The age $\Delta_l(t)$ of source l, where we suppose that the first and third packets are generated from source l, i.e., $r_1 = r_3 = l$.

metrics expressed in terms of auto-correlation functions, the estimation error of signal values, and mutual information, are monotonic functions of the age. Moreover, the age-penalty function $g(\cdot)$ can be used to represent the level of dissatisfaction of data staleness in different applications based on their demands. For instance, a stair-shape function $g(x) = \lfloor x \rfloor$ can be used to characterize the dissatisfaction for data staleness when the information of interest is checked periodically, an exponential function $g(x) = e^x$ can be utilized in online learning and control applications in which the demand for updating data increases quickly with age, and an indicator function $g(x) = \mathbb{1}(x > q)$ can be used to indicate the dissatisfaction of the violation of an age limit q.

C. Decision Policies

A decision policy, denoted by d, controls the following: i) the scheduler, denoted by π, that determines the source to be served at each transmission opportunity $\pi \triangleq (r_1, r_2, \ldots)$, ii) the sampler, denoted by f, that determines the packet generation times $f \triangleq (S_1, S_2, \ldots)$, or equivalently, the sequence of waiting times $f \triangleq (Z_0, Z_1, \ldots)$. Hence, $d = (\pi, f)$ implies that a decision policy d employs the scheduler π and the sampler f. Let $\mathcal{D}$ denote the set of causal decision policies in which decisions are made based on the history and current information of the system. Observe that $\mathcal{D}$ consists of Π and $\mathcal{F}$, where Π and $\mathcal{F}$ are the sets of causal schedulers and samplers, respectively.

After each delivery, the decision maker chooses the source to be served, and imposes a waiting time before the generation of the new packet. Next, we present our optimization problems.

D. Optimization Problem

We define two metrics to assess the long term age performance over our status update system in (4) and (5). Consider the time interval $[0, D_n]$. For any decision policy $d = (\pi, f)$, we define the total-average peak age-penalty (Ta-PAP) as

$$\Delta_{\text{peak}}(\pi, f) = \limsup_{n \to \infty} \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^{n} g\left(\Delta_{r_i}(D_i^-) \right) \right], \tag{4}$$

and the total-average age-penalty per unit time (Ta-AP) as

$$\Delta_{\text{avg}}(\pi, f) = \limsup_{n \to \infty} \frac{\mathbb{E}\left[\sum_{l=1}^{m} \int_{0}^{D_n} g\left(\Delta_l(t)\right) dt\right]}{\mathbb{E}\left[D_n\right]}. \tag{5}$$

In this paper, we aim to minimize both the Ta-PAP and the Ta-AP separately. In other words, we seek a decision policy $d = (\pi, f)$ that solves the following optimization problems:

$$\bar{\Delta}_{\text{peak-opt}} \triangleq \min_{\pi \in \Pi, f \in \mathcal{F}} \Delta_{\text{peak}}(\pi, f), \tag{6}$$

and

$$\bar{\Delta}_{\text{avg-opt}} \triangleq \min_{\pi \in \Pi, f \in \mathcal{F}} \Delta_{\text{avg}}(\pi, f), \tag{7}$$

where $\bar{\Delta}_{\text{peak-opt}}$ and $\bar{\Delta}_{\text{avg-opt}}$ are the optimum objective values of Problems (6) and (7), respectively. Due to the large decision policy space, the optimization problem is quite challenging. In particular, we need to seek the optimal decision policy that controls both the scheduler and sampler to minimize the Ta-PAP and the Ta-AP. In the next section, we discuss our approach to tackle these optimization problems.

III. OPTIMAL DECISION POLICY

We first show that our optimization problems in (6) and (7) have an important separation principle: Given the generation times of the update packets, the Maximum Age First (MAF) scheduler provides the best age performance compared to any other scheduler. What remains to be addressed is the question of finding the best sampler that solves Problems (6) and (7), given that the scheduler is fixed to the MAF. Next, we present our approach to solve our optimization problems in detail.

A. Optimal Scheduler

We start by defining the MAF scheduler as follows:

Definition 4 ([25]–[27], [30], [31])**.** *Maximum Age First scheduler: In this scheduler, the source with the maximum age is served first among all sources. Ties are broken arbitrarily.*

For simplicity, let π_{MAF} represent the MAF scheduler. The age performance of π_{MAF} scheduler is characterized in the following proposition:

Proposition 1. *For all $f \in \mathcal{F}$*

$$\Delta_{\text{peak}}(\pi_{MAF}, f) = \min_{\pi \in \Pi} \Delta_{\text{peak}}(\pi, f), \tag{8}$$

$$\Delta_{\text{avg}}(\pi_{MAF}, f) = \min_{\pi \in \Pi} \Delta_{\text{avg}}(\pi, f). \tag{9}$$

That is, the MAF scheduler minimizes both the Ta-PAP and the Ta-AP in (4) *and* (5) *among all schedulers in* Π.

Proof. One of the key ideas of the proof is as follows: Given any sampler, that controls the generation times of the update packets, we only control from which source a packet is generated. We couple the policies such that the packet delivery times are fixed under all decision policies. In the MAF scheduler, a

(a) The age evolution of source 1.
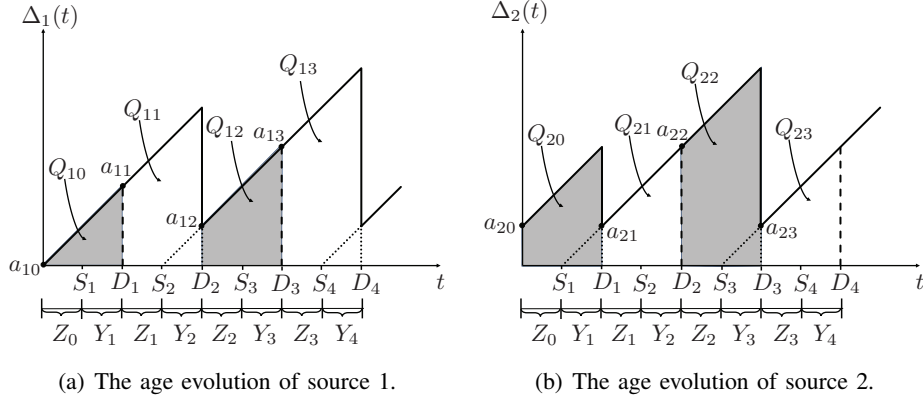
(b) The age evolution of source 2.

Figure 3: The age processes evolution of the MAF scheduler in a two-sources information update system. Source 2 has a higher initial age than Source 1. Thus, Source 2 starts service and packet 1 is generated from Source 2, which is delivered at time D_1. Then, Source 1 is served and packet 2 is generated from Source 1, which is delivered at time D_2. The same operation is repeated over time.

source with maximum age becomes the source with minimum age among the m sources after each delivery. Under any arbitrary scheduler, a packet can be generated from any source, which is not necessarily the one with maximum age, and the chosen source becomes the one with minimum age among the m sources after the delivery. Since the age-penalty function, $g(\cdot)$, is non-decreasing, the MAF scheduler provides a better age performance compared to any other scheduler. For details, see Appendix A. □

Proposition 1 is proven by using a sample-path proof technique that was recently developed in [31]. The difference is that the authors in [31] proved the results for symmetrical packet generation and arrival processes, while we consider here that the packet generation times are controllable. We found that the same proof technique applies to both cases.

Proposition 1 helps us conclude the separation principle that the optimal sampler can be optimized separately, given that the scheduling policy is fixed to the MAF scheduler. Hence, the optimization problems (6) and (7) reduce to the following:

$$\bar{\Delta}_{\text{peak-opt}} \triangleq \min_{f \in \mathcal{F}} \Delta_{\text{peak}}(\pi_{\text{MAF}}, f), \tag{10}$$

$$\bar{\Delta}_{\text{avg-opt}} \triangleq \min_{f \in \mathcal{F}} \Delta_{\text{avg}}(\pi_{\text{MAF}}, f). \tag{11}$$

By fixing the scheduling policy to the MAF scheduler, the evolution of the age processes of the sources is as follows: The sampler may impose a waiting time Z_i before generating packet $i+1$ at time $S_{i+1} = D_i + Z_i$ from the source with the maximum age at time $t = D_i$. Packet $i+1$ is delivered at time $D_{i+1} = S_{i+1} + Y_{i+1}$ and the age of the source with maximum age drops to the minimum age with the value of Y_{i+1}, while the age processes of other sources increase linearly with time without change. This operation is repeated with time and the age processes evolve accordingly. An example of age processes evolution is shown in Fig. 3. Next, we seek the optimal sampler for Problems (10) and (11).

B. Optimal Sampler for Problem (10)

Now, we show that the MAF scheduler and the zero-wait sampler are jointly optimal for minimizing the Ta-PAP as follows:

Theorem 2. *The MAF scheduler and the zero-wait sampler form an optimal solution for Problem* (6).

Proof. We prove Theorem 2 by proving that the zero-wait sampler is optimal for Problem (10). In particular, we show that the Ta-PAP is an increasing function of the packets waiting times Z_i's. For details, see Appendix B. $\qquad\square$

Remark 1. *The results in Proposition 1 and Theorem 2 hold even if $\mathcal{Y}$ and $\mathcal{Z}$ are unbounded and uncountable sets. Indeed, the finiteness assumption of $\mathcal{Y}$ and $\mathcal{Z}$ is only needed for the utilization of the DP technique in the next subsection.*

C. Optimal Sampler for Problem (11)

Although the zero-wait sampler is the optimal sampler for minimizing the Ta-PAP, it is not clear whether it also minimizes the Ta-AP. This is because the latter metric may not be a non-decreasing function of the waiting times as we will see later, which makes Problem (11) more challenging. Next, we derive the Ta-AP when the MAF scheduler is employed and reformulate Problem (11) as a semi-Markov decision problem.

1) Reformulation of Problem (11)*:* We start by analyzing the Ta-AP when the scheduling policy is fixed to the MAF scheduler. We decompose the area under each curve $g(\Delta_l(t))$ into a sum of disjoint geometric parts. Observing Fig. 3 [3], this area in the time interval $[0, D_n]$, where $D_n = \sum_{i=0}^{n-1} Z_i + Y_{i+1}$, can be seen as the concatenation of the areas Q_{li}, $0 \leq i \leq n-1$. Thus, we have

$$\int_0^{D_n} g(\Delta_l(t))dt = \sum_{i=0}^{n-1} Q_{li}, \tag{12}$$

where

$$Q_{li} = \int_{D_i}^{D_{i+1}} g(\Delta_l(t))dt = \int_{D_i}^{D_i + Z_i + Y_{i+1}} g(\Delta_l(t))dt. \tag{13}$$

Let a_{li} denote the age value of source l at time D_i, i.e., $a_{li} = \Delta_l(D_i)$[4]. Hence, for $t \in [D_i, D_{i+1})$, we have

$$\Delta_l(t) = t - U_l(t) = t - (D_i - a_{li}), \tag{14}$$

[3]Observe that a special age-penalty function is depicted in Fig. 3, where we choose $g(x) = x$ to simplify the illustration.

[4]Since the age process is right-continuous, if packet i is delivered from source l, then $\Delta_l(D_i)$ is the age value of source l just after the delivery time D_i.

where $(D_i - a_{li})$ represents the generation time of the last delivered packet from source l before time D_{i+1}. By performing a change of variable in (13), we get

$$Q_{li} = \int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau. \tag{15}$$

Hence, the Ta-AP can be rewritten as

$$\limsup_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}\left[\sum_{l=1}^{m} \int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau\right]}{\sum_{i=0}^{n-1} \mathbb{E}\left[Z_i + Y_{i+1}\right]}. \tag{16}$$

Using this, the optimal sampling problem for minimizing the Ta-AP, given that the scheduling policy is fixed to the MAF scheduler, can be cast as

$$\bar{\Delta}_{\text{avg-opt}} \triangleq \min_{f\in\mathcal{F}} \limsup_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}\left[\sum_{l=1}^{m} \int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau\right]}{\sum_{i=0}^{n-1} \mathbb{E}\left[Z_i + Y_{i+1}\right]}. \tag{17}$$

Since $|\int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau| < \infty$ for all $Z_i \in \mathcal{Z}$ and $Y_i \in \mathcal{Y}$, and $\mathbb{E}[Y_i] > 0$ for all i, $\bar{\Delta}_{\text{avg-opt}}$ is bounded. Note that Problem (17) is hard to solve in the current form. Therefore, we reformulate it. We consider the following optimization problem with a parameter $\beta \geq 0$:

$$p(\beta) \triangleq \min_{f\in\mathcal{F}} \limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}\left[\sum_{l=1}^{m} \int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau - \beta(Z_i + Y_{i+1})\right], \tag{18}$$

where $p(\beta)$ is the optimal value of (18).

Lemma 3. *The following assertions are true:*
(i). $\bar{\Delta}_{avg\text{-}opt} \lesseqgtr \beta$ *if and only if* $p(\beta) \lesseqgtr 0$.
(ii). *If* $p(\beta) = 0$, *then the optimal sampling policies that solve* (17) *and* (18) *are identical.*

Proof. The proof of Lemma 3 is similar to the proof of [38, Lemma 2]. The difference is that a regenerative assumption of the inter-sampling times is used to prove the result in [38]; instead, we use the boundedness of the inter-sampling times to prove the result. For the sake of completeness, we modify the proof accordingly and provide it in Appendix C. □

As a result of Lemma 3, the solution to (17) can be obtained by solving (18) and seeking a $\beta = \bar{\Delta}_{\text{avg-opt}} \geq 0$ such that $p(\bar{\Delta}_{\text{avg-opt}}) = 0$. Lemma 3 helps us to utilize the DP technique to obtain the optimal sampler. Note that without Lemma 3, it would be quite difficult to use the DP technique to solve (17) optimally. Next, we illustrate our solution approach to Problem (18) in detail.

2) The solution of Problem (18)*:* Following the methodology proposed in [39], when $\beta = \bar{\Delta}_{\text{avg-opt}}$, Problem (18) is equivalent to an average cost per stage problem. According to [39], we describe the components of this problem in detail below.

- **States:** At stage[5] i, the system state is specified by

$$\mathbf{s}(i) = (a_{[1]i}, \ldots, a_{[m]i}), \tag{19}$$

[5]From henceforward, we assume that the duration of stage i is $[D_i, D_{i+1})$.

where $a_{[l]i}$ is the l-th largest age of the sources at stage i, i.e., it is the l-th largest component of the vector $(a_{1i}, \ldots, a_{mi})$. We use $\mathcal{S}$ to denote the state-space including all possible states. Notice that $\mathcal{S}$ is finite and bounded because $\mathcal{Z}$ and $\mathcal{Y}$ are finite and bounded.

- **Control action:** At stage i, the action that is taken by the sampler is $Z_i \in \mathcal{Z}$.
- **Random disturbance:** In our model, the random disturbance occurring at stage i is Y_{i+1}, which is independent of the system state and the control action.
- **Transition probabilities:** If the control $Z_i = z$ is applied at stage i and the service time of packet $i+1$ is $Y_{i+1} = y$, then the evolution of the system state from $\mathbf{s}(i)$ to $\mathbf{s}(i+1)$ is as follows:

$$
\begin{aligned}
a_{[m]i+1} &= y, \\
a_{[l]i+1} &= a_{[l+1]i} + z + y, \ \ l = 1, \ldots, m-1.
\end{aligned}
\tag{20}
$$

We let $\mathbb{P}_{\mathbf{ss}'}(z)$ denote the transition probabilities

$$
\mathbb{P}_{\mathbf{ss}'}(z) = \mathbb{P}(\mathbf{s}(i+1) = \mathbf{s}' | \mathbf{s}(i) = \mathbf{s}, Z_i = z), \ \ \mathbf{s}, \mathbf{s}' \in \mathcal{S}.
\tag{21}
$$

When $\mathbf{s} = (a_{[1]}, \ldots, a_{[m]})$ and $\mathbf{s}' = (a'_{[1]}, \ldots, a'_{[m]})$, the law of the transition probability is given by

$$
\mathbb{P}_{\mathbf{ss}'}(z) =
\begin{cases}
\mathbb{P}(Y_{i+1} = y) & \text{if } a'_{[m]} = y \text{ and} \\
& \quad a'_{[l]} = a_{[l+1]} + z + y \text{ for } l \neq m; \\
0 & \text{else.}
\end{cases}
\tag{22}
$$

- **Cost Function:** Each time the system is in stage i and control Z_i is applied, we incur a cost

$$
C(\mathbf{s}(i), Z_i, Y_{i+1}) = \sum_{l=1}^{m} \int_{a_{[l]i}}^{a_{[l]i} + Z_i + Y_{i+1}} g(\tau) d\tau - \bar{\Delta}_{\text{avg-opt}}(Z_i + Y_{i+1}).
\tag{23}
$$

To simplify notation, we use the expected cost $C(\mathbf{s}(i), Z_i)$ as the cost per stage, i.e.,

$$
C(\mathbf{s}(i), Z_i) = \mathbb{E}_{Y_{i+1}}\left[C(\mathbf{s}(i), Z_i, Y_{i+1})\right],
\tag{24}
$$

where $\mathbb{E}_{Y_{i+1}}$ is the expectation with respect to Y_{i+1}, which is independent of $\mathbf{s}(i)$ and Z_i. It is important to note that there exists $c \in \mathbb{R}^+$ such that $|C(\mathbf{s}(i), Z_i)| \leq c$ for all $\mathbf{s}(i) \in \mathcal{S}$ and $Z_i \in \mathcal{Z}$. This is because $\mathcal{Z}$, $\mathcal{Y}$, $\mathcal{S}$, and $\bar{\Delta}_{\text{avg-opt}}$ are bounded.

In general, the average cost per stage under a sampling policy $f \in \mathcal{F}$ is given by

$$
\limsup_{n \to \infty} \frac{1}{n} \mathbb{E}\left[\sum_{i=0}^{n-1} C(\mathbf{s}(i), Z_i)\right].
\tag{25}
$$

We say that a sampling policy $f \in \mathcal{F}$ is *average-optimal* if it minimizes the average cost per stage in (25). Our objective is to find the average-optimal sampling policy. A policy f is called a stationary randomized policy if it assigns a probability distribution $q_{\mathcal{Z}}(\mathbf{s}(i))$ over the control set based on the state $\mathbf{s}(i)$ such that it chooses the control Z_i randomly according to this distribution; while a stationary deterministic policy chooses an action with certainty such that $Z_i = Z_j$ whenever $\mathbf{s}(i) = \mathbf{s}(j)$ for any i, j. According to [39], there may not exist a stationary deterministic policy that is average-optimal. In the next proposition, we show that there is a stationary deterministic policy that is average-optimal.

Proposition 4. *There exist a scalar λ and a function h that satisfy the following Bellman's equation:*

$$\lambda + h(\mathbf{s}) = \min_{z \in \mathcal{Z}} \left(C(\mathbf{s}, z) + \sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{s}\mathbf{s}'}(z) h(\mathbf{s}') \right), \tag{26}$$

where λ is the optimal average cost per stage that is independent of the initial state $\mathbf{s}(0)$ and satisfies

$$\lambda = \lim_{\alpha \to 1} (1 - \alpha) J_\alpha(\mathbf{s}), \forall \mathbf{s} \in \mathcal{S}, \tag{27}$$

and $h(\mathbf{s})$ is the relative cost function that, for any state $\mathbf{o}$, satisfies

$$h(\mathbf{s}) = \lim_{\alpha \to 1} (J_\alpha(\mathbf{s}) - J_\alpha(\mathbf{o})), \forall \mathbf{s} \in \mathcal{S}, \tag{28}$$

where $J_\alpha(\mathbf{s})$ is the optimal total expected α-discounted cost function, which is defined by

$$J_\alpha(\mathbf{s}) = \min_{f \in \mathcal{F}} \limsup_{n \to \infty} \mathbb{E} \left[\sum_{i=0}^{n-1} \alpha^i C(\mathbf{s}(i), Z_i) \right], \mathbf{s}(0) = \mathbf{s} \in \mathcal{S}, \tag{29}$$

where $0 < \alpha < 1$ is the discount factor. Furthermore, there exists a stationary deterministic policy that attains the minimum in (26) for each $\mathbf{s} \in \mathcal{S}$ and is average-optimal.

Proof. According to [39, Proposition 4.2.1 and Proposition 4.2.6], it is enough to show that for every two states $\mathbf{s}$ and $\mathbf{s}'$, there exists a stationary deterministic policy f such that for some k, we have $\mathbb{P}[\mathbf{s}(k) = \mathbf{s}'|\mathbf{s}(0) = \mathbf{s}, f] > 0$, i.e., we have a communicating Markov decision process (MDP). Observe that the proof idea of this proposition is different from those used in literature such as [27], [29], where they have used the discounted cost problem to show their results and then connect it to the average cost problem. For details, see Appendix D. $\qquad\square$

We can deduce from Proposition 4 that the optimal waiting time is a fixed function of the state $\mathbf{s}$. Next, we use the RVI algorithm to obtain the optimal sampler for minimizing the Ta-AP, and then exploit the structure of our problem to reduce its complexity.

Optimal Sampler Structure: The RVI algorithm [40, Section 9.5.3], [41, Page 171] can be used to solve Bellman's equation (26). Starting with an arbitrary state $\mathbf{o}$, a single iteration for the RVI algorithm is given as follows:

$$Q_{n+1}(\mathbf{s}, z) = C(\mathbf{s}, z) + \sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{s}\mathbf{s}'}(z) h_n(\mathbf{s}'),$$

$$J_{n+1}(\mathbf{s}) = \min_{z \in \mathcal{Z}} (Q_{n+1}(\mathbf{s}, z)), \tag{30}$$

$$h_{n+1}(\mathbf{s}) = J_{n+1}(\mathbf{s}) - J_{n+1}(\mathbf{o}),$$

where $Q_{n+1}(\mathbf{s}, z)$, $J_n(\mathbf{s})$, and $h_n(\mathbf{s})$ denote the state action value function, value function, and relative value function for iteration n, respectively. In the beginning, we set $J_0(\mathbf{s}) = 0$ for all $\mathbf{s} \in \mathcal{S}$, and then

---

Algorithm 1: RVI algorithm with reduced complexity.

1 given $l = 0$, sufficiently large u, tolerance $\epsilon_1 > 0$, tolerance $\epsilon_2 > 0$;

2 while $u - l > \epsilon_1$ **do**

3 $\beta = \frac{l+u}{2}$;

4 $J(\mathbf{s}) = 0$, $h(\mathbf{s}) = 0$, $h_{\text{last}}(\mathbf{s}) = 0$ for all states $\mathbf{s} \in \mathcal{S}$;

5   **while** $\max_{\mathbf{s} \in \mathcal{S}} |h(\mathbf{s}) - h_{last}(\mathbf{s})| > \epsilon_2$ **do**

6   **for** *each* $\mathbf{s} \in \mathcal{S}$ **do**

7   **if** $\mathbb{E}_Y\left[\sum_{l=1}^m g(a_{[l]} + Y)\right] \geq \beta$ **then**

8 $z_s^\star = 0$;

9   **else**

10 $z_s^\star = \text{argmin}_{z \in \mathcal{Z}} C(\mathbf{s}, z) + \sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z)h(\mathbf{s}')$;

11   **end**

12 $J(\mathbf{s}) = C(\mathbf{s}, z_s^\star) + \sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z_s^\star)h(\mathbf{s}')$;

13   **end**

14 $h_{\text{last}}(\mathbf{s}) = h(\mathbf{s})$;

15 $h(\mathbf{s}) = J(\mathbf{s}) - J(\mathbf{o})$;

16   **end**

17   **if** $J(\mathbf{o}) \geq 0$ **then**

18 $u = \beta$;

19   **else**

20 $l = \beta$;

21   **end**

22 end

---

we repeat the iteration of the RVI algorithm as described before[6].

The complexity of the RVI algorithm is high due to many sources (i.e., the curse of dimensionatlity [42]). Thus, we need to simplify the RVI algorithm. To that end, we show that the optimal sampler has a threshold property that can reduce the complexity of the RVI algorithm. Define $z_s^\star$ as the optimal waiting time for state $\mathbf{s}$, and Y as a random variable that has the same distribution as Y_i. The threshold property in the optimal sampler is manifested in the following proposition:

Proposition 5. *If the state* $\mathbf{s} = (a_{[1]}, \ldots, a_{[m]})$ *satisfies* $\mathbb{E}_Y\left[\sum_{l=1}^m g(a_{[l]} + Y)\right] \geq \bar{\Delta}_{avg\text{-}opt}$, *then we have* $z_s^\star = 0$.

Proof. See Appendix F.   □

We can exploit the threshold test in Proposition 5 to reduce the complexity of the RVI algorithm as follows: The optimal waiting time for any state $\mathbf{s}$ that satisfies $\mathbb{E}_Y\left[\sum_{l=1}^m g(a_{[l]} + Y)\right] \geq \bar{\Delta}_{avg\text{-}opt}$ is zero. Thus, we need to solve (30) only for the states that fail this threshold test. As a result, we reduce the number of computations required along the system state space, which reduces the complexity of the RVI

---

[6]According to [40], [41], a sufficient condition for the convergence of the RVI algorithm is the aperiodicity of the transition matrices of stationary deterministic optimal policies. In our case, these transition matrices depend on the service times. This condition can always be achieved by applying the aperiodicity transformation as explained in [40, Section 8.5.4], which is a simple transformation. However, This is not always necessary to be done.

algorithm. Note that $\bar{\Delta}_{\text{avg-opt}}$ can be obtained using the bisection method or any other one-dimensional search method. Combining this with the result of Proposition 5 and the RVI algorithm, we propose the "RVI with reduced complexity (RVI-RC) sampler" in Algorithm 1. In the outer layer of Algorithm 1, bisection is employed to obtain $\bar{\Delta}_{\text{avg-opt}}$, where β converges to $\bar{\Delta}_{\text{avg-opt}}$.

Note that, according to [40], [41], $J(\mathbf{o})$ in Algorithm 1 converges to the optimal average cost per stage. Moreover, the value of u in Algorithm 1 can be initialized to the value of the Ta-AP of the zero-wait sampler (as the Ta-AP of the zero-wait sampler provides an upper bound on the optimal Ta-AP), which can be easily calculated.

The RVI algorithm and Whittle's methodology have been used in literature to obtain the optimal age scheduler in time-slotted multi-source networks (e.g., [27], [29]). Since they considered a time-slotted system, their model belongs to the class of Markov decision problems. In contrast, we consider random discrete transmission times that can be more than one time slot. Thus, our model belongs to the class of semi-Markov decision problems, and hence is different from those in [27], [29].

In conclusion, an optimal solution for Problem (7) is manifested in the following theorem:

Theorem 6. *The MAF scheduler and the RVI-RC sampler form an optimal solution for Problem* (7).

Proof. The theorem follows directly from Proposition 1, Proposition 4, and Proposition 5 ☐

3) Special Case of $g(x) = x$: Now we consider the case of $g(x) = x$ and obtain some useful insights. Define $A_s = \sum_{l=1}^{m} a_{[l]}$ as the sum of the age values of state $\mathbf{s}$. The threshold test in Proposition 5 is simplified as follows:

Proposition 7. *If the state $\mathbf{s} = (a_{[1]}, \ldots, a_{[m]})$ satisfies $A_s \geq (\bar{\Delta}_{avg\text{-}opt} - m\mathbb{E}[Y])$, then we have $z_s^\star = 0$.*

Proof. The proposition follows directly by substituting $g(x) = x$ into the threshold test in Proposition 5. ☐

Hence, the only change in Algorithm 1 is to replace the threshold test in Step 7 by $A_s \geq (\bar{\Delta}_{\text{avg-opt}} - m\mathbb{E}[Y])$. Let $y_{\text{inf}} = \inf\{y \in \mathcal{Y} : \mathbb{P}[Y = y] > 0\}$, i.e., y_{inf} is the smallest possible transmission time in $\mathcal{Y}$. As a result of Proposition 7, we obtain the following sufficient condition for the optimality of the zero-wait sampler for minimizing the Ta-AP when $g(x) = x$:

Theorem 8. *If*

$$y_{inf} \geq \frac{(m^2 - m)\mathbb{E}[Y]^2 + m\mathbb{E}[Y^2]}{m(m + 1)\mathbb{E}[Y]}, \tag{31}$$

then the zero-wait sampler is optimal for Problem (18).

Proof. See Appendix F ☐

From Theorem 8, it immediately follows that:

Corollary 9. *If the transmission times are positive and constant (i.e., $Y_i = const > 0$ for all i), then the zero-wait sampler is optimal for Problem* (18).

Proof. The corollary follows directly from Theorem 8 by showing that (31) always holds in this case. □

Corollary 9 suggests that the designed schedulers in [24]–[29], [32]–[36] are indeed optimal in time-slotted systems. However, if there is a variation in the transmission times, these schedulers alone may not be optimal anymore, and we need to optimize the sampling times as well.

IV. LOW-COMPLEXITY SAMPLER DESIGN VIA BELLMAN'S EQUATION APPROXIMATION

In this section, we try to obtain low-complexity samplers via an approximate analysis for Bellman equation in (26). The obtained low-complexity samplers in this section will be shown to have near optimal age performance in our numerical results in Section V. For a given state $\mathbf{s}$, we denote the next state given z and y by $\mathbf{s}'(z, y)$. We can observe that the transition probability in (22) depends only on the distribution of the packet service time which is independent of the system state and the control action. Hence, the second term in Bellman's equation in (26) can be rewritten as

$$\sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{s}\mathbf{s}'}(z)h(\mathbf{s}'(z, y)) = \sum_{y \in \mathcal{Y}} \mathbb{P}(Y = y)h(\mathbf{s}'(z, y)). \tag{32}$$

As a result, Bellman's equation in (26) can be rewritten as

$$\lambda = \min_z \left(C(\mathbf{s}, z) + \sum_{y \in \mathcal{Y}} \mathbb{P}(Y = y)(h(\mathbf{s}'(z, y)) - h(\mathbf{s})) \right). \tag{33}$$

Although $h(\mathbf{s})$ is discrete, we can interpolate the value of $h(\mathbf{s})$ between the discrete values so that it is differentiable by following the same approach in [43] and [44]. Let $\mathbf{s} = (a_{[1]}, \ldots, a_{[m]})$, then using the first order Taylor approximation around a state $\mathbf{v} = (a_{[1]}^v, \ldots, a_{[m]}^v)$ (some fixed state), we get

$$h(\mathbf{s}) \approx h(\mathbf{v}) + \sum_{l=1}^m (a_{[l]} - a_{[l]}^v)\frac{\partial h(\mathbf{v})}{\partial a_{[l]}}. \tag{34}$$

Again, we use the first order Taylor approximation around the state $\mathbf{v}$, together with the state evolution in (20), to get

$$h(\mathbf{s}'(z, y)) \approx h(\mathbf{v}) + (y - a_{[m]}^v)\frac{\partial h(\mathbf{v})}{\partial a_{[m]}} + \sum_{l=1}^{m-1} (a_{[l+1]} - a_{[l]}^v + z + y)\frac{\partial h(\mathbf{v})}{\partial a_{[l]}}. \tag{35}$$

From (34) and (35), we get

$$h(\mathbf{s}'(z, y)) - h(\mathbf{s}) \approx (y - a_{[m]})\frac{\partial h(\mathbf{v})}{\partial a_{[m]}} + \sum_{l=1}^{m-1} (a_{[l+1]} - a_{[l]} + z + y)\frac{\partial h(\mathbf{v})}{\partial a_{[l]}}. \tag{36}$$

This implies that

$$\sum_{y \in \mathcal{Y}} \mathbb{P}(Y = y)(h(\mathbf{s}'(z, y)) - h(\mathbf{s})) \approx (\mathbb{E}[Y] - a_{[m]})\frac{\partial h(\mathbf{v})}{\partial a_{[m]}} + \sum_{l=1}^{m-1} (a_{[l+1]} - a_{[l]} + z + \mathbb{E}[Y])\frac{\partial h(\mathbf{v})}{\partial a_{[l]}}. \tag{37}$$

Using (33) with (37), we can get the following approximated Bellman's equation:

$$\lambda \approx \min_z \left(C(\mathbf{s}, z) + (\mathbb{E}[Y] - a_{[m]}) \frac{\partial h(\mathbf{v})}{\partial a_{[m]}} + \sum_{l=1}^{m-1} (a_{[l+1]} - a_{[l]} + z + \mathbb{E}[Y]) \frac{\partial h(\mathbf{v})}{\partial a_{[l]}} \right). \tag{38}$$

By following the same steps as in Appendix F to get the optimal z that minimizes the objective function in (38), we get the following condition: The optimal z, for a given state $\mathbf{s}$, must satisfy

$$\mathbb{E}_Y \left[\sum_{l=1}^{m} g(a_{[l]} + t + Y) \right] - \bar{\Delta}_{\text{avg-opt}} + \sum_{l=1}^{m-1} \frac{\partial h(\mathbf{v})}{\partial a_{[l]}} \geq 0 \tag{39}$$

for all $t > z$, and

$$\mathbb{E}_Y \left[\sum_{l=1}^{m} g(a_{[l]} + t + Y) \right] - \bar{\Delta}_{\text{avg-opt}} + \sum_{l=1}^{m-1} \frac{\partial h(\mathbf{v})}{\partial a_{[l]}} \leq 0 \tag{40}$$

for all $t < z$. The smallest z that satisfies (39)-(40) is

$$\hat{z}_s^\star = \inf \left\{ t \geq 0 : \mathbb{E}_Y \left[\sum_{l=1}^{m} g(a_{[l]} + t + Y) \right] \geq \bar{\Delta}_{\text{avg-opt}} - \sum_{l=1}^{m-1} \frac{\partial h(\mathbf{v})}{\partial a_{[l]}} \right\}, \tag{41}$$

where $\hat{z}_s^*$ is the optimal solution of the approximated Bellman's equation for state $\mathbf{s}$. Note that the term $\sum_{i=1}^{m-1} \frac{\partial h(\mathbf{v})}{\partial a_{[i]}}$ is constant. Hence, (41) can be rewritten as

$$\hat{z}_s^\star = \inf \left\{ t \geq 0 : \mathbb{E}_Y \left[\sum_{l=1}^{m} g(a_{[l]} + t + Y) \right] \geq T \right\}. \tag{42}$$

This simple threshold sampler can approximate the optimal sampler for the original Bellman's equation in (26). The optimal threshold (T) in (42) can be obtained using a golden-section method [45]. Moreover, for a given state $\mathbf{s}$ and the threshold (T), (42) can be solved using the bisection method or any other one-dimensional search method.

A. Low-complexity Water-filling Sampler

Consider the case that $g(x) = x$, the solution in (42) can be further simplified. Substituting $g(x) = x$ into (41), where the equality holds in this case, we get the following condition: The optimal z in this case, for a given state $\mathbf{s}$, must satisfy

$$A_s - \bar{\Delta}_{\text{avg-opt}} + mz + m\mathbb{E}[Y] + \sum_{l=1}^{m-1} \frac{\partial h(\mathbf{v})}{\partial a_{[l]}} = 0, \tag{43}$$

where A_s is the sum of the age values of state $\mathbf{s}$. Rearranging (43), we get

$$\hat{z}_s^* = \left[\frac{\bar{\Delta}_{\text{avg-opt}} - m\mathbb{E}[Y] - \sum_{l=1}^{m-1} \frac{\partial h(\mathbf{v})}{\partial a_{[l]}}}{m} - \frac{A_s}{m} \right]^+. \tag{44}$$
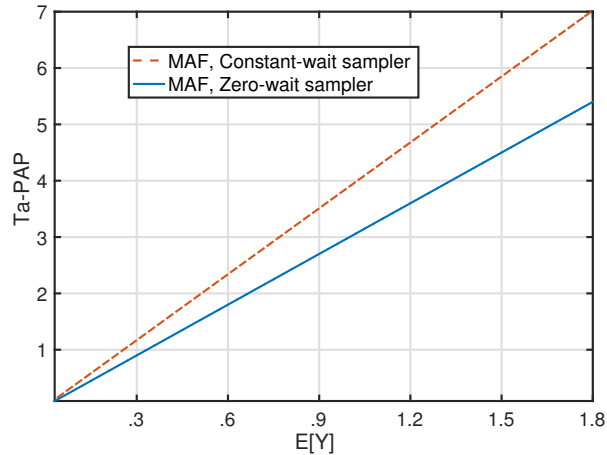
Figure 4: Ta-PAP versus the mean of the packet transmission time $\mathbb{E}[Y]$ for an update system with $m = 3$ sources, where $g(x) = x$.

By observing that the term $\sum_{i=1}^{m-1} \frac{\partial h(\mathbf{v})}{\partial a_{[i]}}$ is constant, (44) can be rewritten as

$$\hat{z}_s^* = \left[T - \frac{A_s}{m} \right]^+ , \tag{45}$$

The solution in (45) is in the form of the water-filling solution as we compare a fixed threshold (T) with the average age of a state $\mathbf{s}$. The solution in (45) suggests that this simple water-filling sampler can approximate the optimal solution of the original Bellman's equation in (26) when $g(x) = x$. Similar to the general case, the optimal threshold (T) in (45) can be obtained using a golden-section method. We evaluate the performance of the approximated samplers in the next section.

V. NUMERICAL RESULTS

We present numerical results to evaluate our proposed solutions. We consider an information update system with $m = 3$ sources. We use "RAND" to represent the random scheduler in which a source is chosen randomly to be served. By "constant-wait sampler", we refer to the sampler that imposes a constant waiting time after each delivery with $Z_i = 0.3\mathbb{E}[Y]$, $\forall i$. Moreover, we use "threshold sampler" and "water-filling sampler" to denote the proposed samplers in (42) and (45), respectively.

Fig. 4 illustrates the Ta-PAP versus the mean transmission time $\mathbb{E}[Y]$, where we have $g(x) = x$. As we can observe, with fixing the scheduling policy to the MAF scheduler, the zero-wait sampler provides a lower Ta-PAP compared to the constant-wait sampler. This observation agrees with Theorem 2. However, as we will see later, zero-wait sampler does not always minimize the Ta-AP.

We now evaluate the performance of our proposed solutions for minimizing the Ta-AP. We set the transmission times to be either 0 or 3 with probability p and $1 - p$, respectively. Fig. 5, 6, and 7 illustrate the Ta-AP versus the transmission probability p, where we set the age-penalty function $g(x)$ to be $e^{0.1x} - 1$, $x^{0.1}$, and x, respectively. For the zero-wait sampler, we find that the MAF scheduler provides a lower Ta-AP than that of the RAND scheduler. This agrees with Proposition 1. Moreover, when the scheduling policy is fixed to the MAF scheduler, we find that the Ta-AP resulting from the
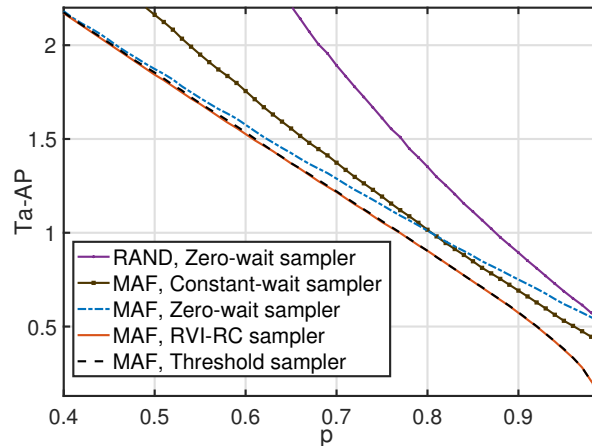
Figure 5: Ta-AP versus transmission probability p for an update system with $m = 3$ sources, where $g(x) = e^{0.1x} - 1$.
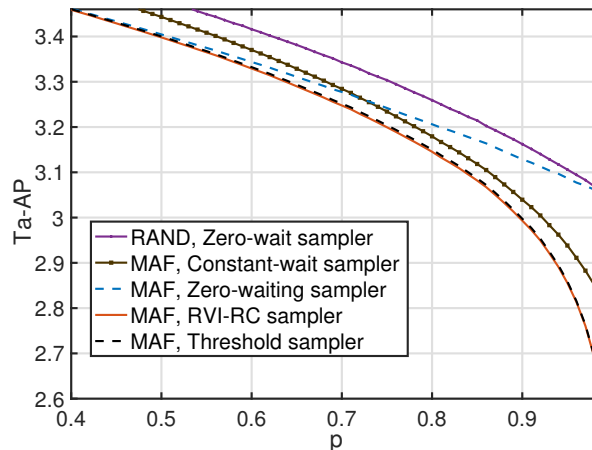

Figure 6: Ta-AP versus transmission probability p for an update system with $m = 3$ sources, where $g(x) = x^{0.1}$.

RVI-RC sampler is lower than those resulting from the zero-wait sampler and the constant-wait sampler. This observation suggests the following: i) The zero-wait sampler does not necessarily minimize the Ta-AP, ii) optimizing the scheduling policy only is not enough to minimize the Ta-AP, but we have to optimize both the scheduling policy and the sampling policy together to minimize the Ta-AP. In addition, as we can observe, the Ta-AP resulting from the threshold sampler in Fig. 5 and 6, and the water-filling sampler in Fig. 7 almost coincides on the Ta-AP resulting from the RVI-RC sampler.

We then set the transmission times to be either 0 or $Y_{\max}$ with probability 0.9 and 0.1, respectively. We vary the maximum transmission time $Y_{\max}$ and plot the Ta-AP in Fig. 8, 9, and 10, where $g(x)$ is set to be $e^{0.1x} - 1$, $x^{0.1}$, and x, respectively. The scheduling policy is fixed to the MAF scheduler in all plotted curves. We can observe in all figures that the Ta-AP resulting from the RVI-RC sampler is lower than those resulting from the zero-wait sampler and the constant-wait sampler, and the gap between them
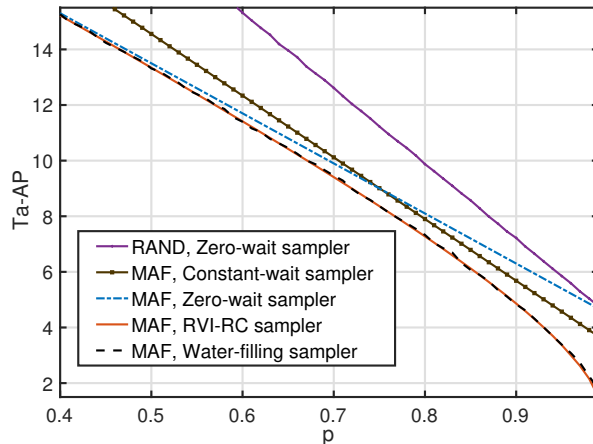
Figure 7: Ta-AP versus transmission probability p for an update system with $m = 3$ sources, where $g(x) = x$.
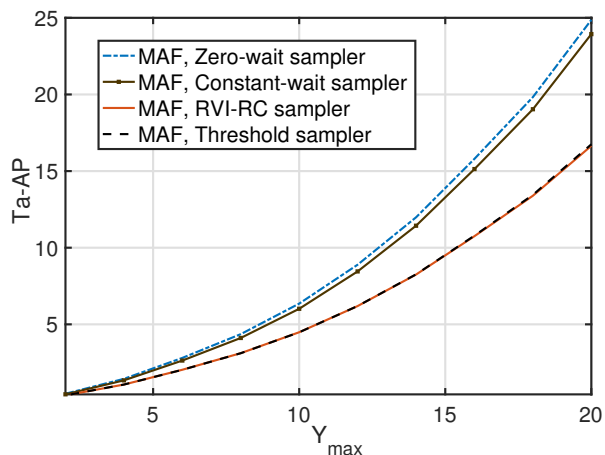

Figure 8: Ta-AP versus the maximum service time $Y_{\max}$ for an update system with $m = 3$ sources, where $g(x) = e^{0.1x} - 1$.

increases as the variability (variance) of the transmission times increases. This suggests that when the transmission times have a big variation, we have to optimize the scheduler and the sampler together to minimize the Ta-AP. Finally, as we can observe, the Ta-AP of the threshold sampler in Fig. 8 and 9, and the water-filling sampler in Fig. 10 almost coincides on that of the RVI-RC sampler.

VI. CONCLUSION

In this work, we studied the problem of finding the optimal decision policy that controls the packet generation times and transmission order of the sources to minimize the Ta-PAP and Ta-AP in a multi-source information update system. We showed that the MAF scheduler and the zero-wait sampler are jointly optimal for minimizing the Ta-PAP. Moreover, we showed that the MAF scheduler and the RVI-RC sampler, that results from reducing the computation complexity of the RVI algorithm, are jointly optimal for minimizing the Ta-AP. Finally, we devised a low-complexity threshold sampler via an
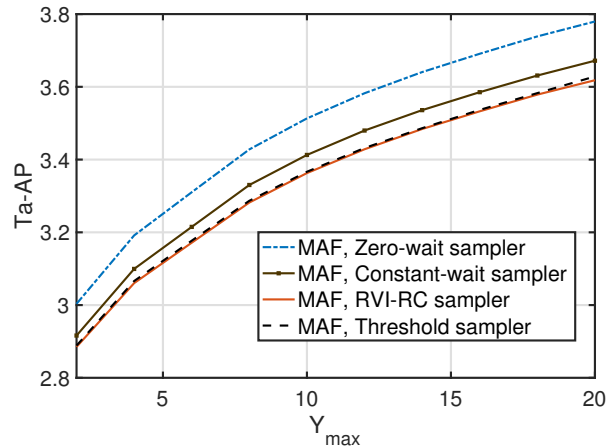
Figure 9: Ta-AP versus the maximum service time $Y_{\max}$ for an update system with $m = 3$ sources, where $g(x) = x^{0.1}$.
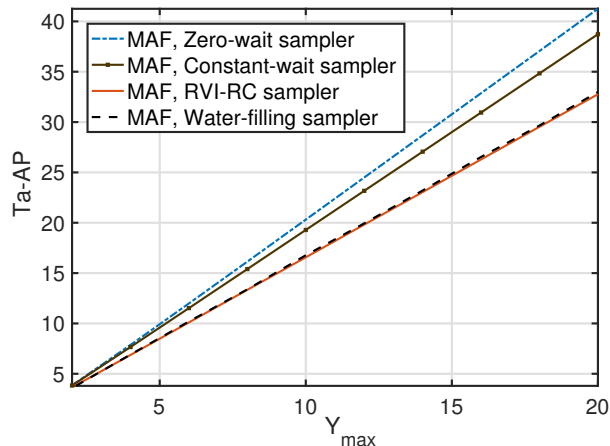

Figure 10: Ta-AP versus the maximum service time $Y_{\max}$ for an update system with $m = 3$ sources, where $g(x) = x$.

approximate analysis of Bellman's equation. This threshold sampler is further simplified to a simple water-filling sampler in the special case of linear age-penalty function. The numerical results showed that the performance of these approximated samplers is almost the same as that of the RVI-RC sampler.

VII. APPENDIX

APPENDIX A

PROOF OF PROPOSITION 1

Let the vector $\boldsymbol{\Delta}_\pi(t) = (\Delta_{[1],\pi}(t), \ldots, \Delta_{[m],\pi}(t))$ denote the system state at time t of the scheduler π, where $\Delta_{[l],\pi}(t)$ is the l-th largest age of the sources at time t under the scheduler π. Let $\{\boldsymbol{\Delta}_\pi(t), t \geq 0\}$ denote the state process of the scheduler π. For notational simplicity, let P represent the MAF scheduler. Throughout the proof, we assume that $\boldsymbol{\Delta}_\pi(0^-) = \boldsymbol{\Delta}_P(0^-)$ for all π and the sampler is fixed to an

arbitrarily chosen one. The key step in the proof of Proposition 1 is the following lemma, where we compare the scheduler P with any arbitrary scheduler π.

Lemma 10. *Suppose that* $\boldsymbol{\Delta}_\pi(0^-) = \boldsymbol{\Delta}_P(0^-)$ *for all scheduler π and the sampler is fixed, then we have*

$$\{\boldsymbol{\Delta}_P(t), t \geq 0\} \leq_{st} \{\boldsymbol{\Delta}_\pi(t), t \geq 0\} \tag{46}$$

We use a coupling and forward induction to prove Lemma 10. For any scheduler π, suppose that the stochastic processes $\widetilde{\boldsymbol{\Delta}}_P(t)$ and $\widetilde{\boldsymbol{\Delta}}_\pi(t)$ have the same stochastic laws as $\boldsymbol{\Delta}_P(t)$ and $\boldsymbol{\Delta}_\pi(t)$. The state processes $\widetilde{\boldsymbol{\Delta}}_P(t)$ and $\widetilde{\boldsymbol{\Delta}}_\pi(t)$ are coupled such that the packet service times are equal under both scheduling policies, i.e., Y_i's are the same under both scheduling policies. Such a coupling is valid since the service time distribution is fixed under all policies. Since the sampler is fixed, such a coupling implies that the packet generation and delivery times are the same under both schedulers. According to Theorem 6.B.30 of [37], if we can show

$$\mathbb{P}\left[\widetilde{\boldsymbol{\Delta}}_P(t) \leq \widetilde{\boldsymbol{\Delta}}_\pi(t), t \geq 0\right] = 1, \tag{47}$$

then (46) is proven. To ease the notational burden, we will omit the tildes on the coupled versions in this proof and just use $\boldsymbol{\Delta}_P(t)$ and $\boldsymbol{\Delta}_\pi(t)$. Next, we compare scheduler P and scheduler π on a sample path and prove (46) using the following lemma:

Lemma 11 (Inductive Comparison). *Suppose that a packet with generation time S is delivered under the scheduler P and the scheduler π at the same time t. The system state of the scheduler P is $\boldsymbol{\Delta}_P$ before the packet delivery, which becomes $\boldsymbol{\Delta}'_P$ after the packet delivery. The system state of the scheduler π is $\boldsymbol{\Delta}_\pi$ before the packet delivery, which becomes $\boldsymbol{\Delta}'_\pi$ after the packet delivery. If*

$$\Delta_{[i],P} \leq \Delta_{[i],\pi}, i = 1, \ldots, m, \tag{48}$$

then

$$\Delta'_{[i],P} \leq \Delta'_{[i],\pi}, i = 1, \ldots, m. \tag{49}$$

Lemma 11 is proven by following the proof idea of [31, Lemma 2]. For the sake of completeness, we provide the proof of Lemma 11 as follows:

Proof. Since only one source can be scheduled at a time and the scheduler P is the MAF one, the packet with generation time S must be generated from the source with maximum age $\Delta_{[1],P}$, call it source l^*. In other words, the age of source l^* is reduced from the maximum age $\Delta_{[1],P}$ to the minimum age $\Delta'_{[m],P} = t - S$, and the age of the other $(m-1)$ sources remain unchanged. Hence,

$$\begin{aligned} \Delta'_{[i],P} &= \Delta_{[i+1],P}, i = 1, \ldots, m-1, \\ \Delta'_{[m],P} &= t - S. \end{aligned} \tag{50}$$

In the scheduler π, this packet can be generated from any source. Thus, for all cases of scheduler π, it must hold that

$$\Delta'_{[i],\pi} \geq \Delta_{[i+1],\pi}, i = 1, \ldots, m-1. \tag{51}$$

By combining (48), (50), and (51), we have

$$\Delta'_{[i],\pi} \geq \Delta_{[i+1],\pi} \geq \Delta_{[i+1],P} = \Delta'_{[i],P}, i = 1, \ldots, m-1. \tag{52}$$

In addition, since the same packet is also delivered under the scheduler π, the source from which this packet is generated under policy π will have the minimum age after the delivery, i.e., we have

$$\Delta'_{[m],\pi} = t - S = \Delta'_{[m],P}. \tag{53}$$

By this, (49) is proven. $\qquad\square$

Proof of Lemma 10. Using the coupling between the system state processes, and for any given sample path of the packet service times, we consider two cases:

Case 1: When there is no packet delivery, the age of each source grows linearly with a slope 1.

Case 2: When a packet is delivered, the ages of the sources evolve according to Lemma 11.

By induction over time, we obtain

$$\Delta_{[i],P}(t) \leq \Delta_{[i],\pi}(t), i = 1, \ldots, m, t \geq 0. \tag{54}$$

Hence, (47) follows which implies (46) by Theorem 6.B.30 of [37]. This completes the proof. $\qquad\square$

Proof of Proposition 1. Since the Ta-PAP and Ta-AP for any scheduling policy π are the expectation of non-decreasing functional of the process $\{\boldsymbol{\Delta}_\pi(t), t \geq 0\}$, (46) implies (8) and (9) using the properties of stochastic ordering [37]. This completes the proof. $\qquad\square$

APPENDIX B
PROOF OF THEOREM 2

According to Proposition 1, we need to fix the scheduler to the MAF, and show that the Ta-PAP is an increasing function of the packets waiting times Z_i's. For notational simplicity, let a_{li} denote the age value of source l at time D_i, i.e., $a_{li} = \Delta_l(D_i)$[7]. Since the age process increases linearly with time when there is no packet delivery, we have

$$\Delta_{r_i}(D_i^-) = a_{r_i(i-1)} + Z_{i-1} + Y_i, \tag{55}$$

where $a_{r_i(i-1)} = \Delta_{r_i}(D_{i-1})$. Substituting this into (4), we get

$$\Delta_{\text{peak}}(\pi_{\text{MAF}}, f) = \limsup_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^{n} g(a_{r_i(i-1)} + Z_{i-1} + Y_i)\right]. \tag{56}$$

Since the scheduling policy is fixed to the MAF scheduler, the last service that source r_i has received before time D_{i-1} was at time D_{i-m}. Since the age process increases linearly if there is no packet delivery, we have

$$a_{r_i(i-1)} = D_{i-1} - D_{i-m} + Y_{i-m}, \tag{57}$$

---

[7]Since the age process is right-continuous, if packet i is delivered from source l, then $\Delta_l(D_i)$ is the age value of source l just after the delivery time D_i.

where Y_{i-m} is the age value of the source r_i at time D_{i-m}, i.e., $\Delta_{r_i}(D_{i-m}) = Y_{i-m}$. Note that $D_{i-1} = Y_{i-1} + Z_{i-2} + D_{i-2}$. Repeating this, we can express $(D_{i-1} - D_{i-m})$ in terms of Z_i's and Y_i's, and hence we get

$$a_{r_i(i-1)} = \sum_{k=1}^{m} Y_{i-k} + \sum_{k=2}^{m} Z_{i-k}. \tag{58}$$

For example, in Fig. 3, we have $a_{22} = Y_1 + Z_1 + Y_2$. Note that (58) holds after the first m transmissions, i.e., for $i > m$. In a similar way, we can express $a_{r_i(i-1)}$ in terms of Z_i's and Y_i's for $i \leq m$ as follows:

$$a_{r_i(i-1)} = a_{r_i 0} + \sum_{k=1}^{i-1} Y_{i-k} + \sum_{k=2}^{i} Z_{i-k}, \tag{59}$$

where $a_{r_i 0}$ is the initial age value of source r_i. In Fig. 3, For example, we have $a_{11} = a_{10} + Z_0 + Y_1$. Substituting (58) and (59) into (56), we get

$$\Delta_{\text{peak}}(\pi_{\text{MAF}}, f) = \limsup_{n \to \infty} \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^{m} g\left(a_{r_i 0} + \sum_{k=0}^{i-1} Y_{i-k} + \sum_{k=1}^{i} Z_{i-k} \right) + \sum_{i=m+1}^{n} g\left(\sum_{k=0}^{m} Y_{i-k} + \sum_{k=1}^{m} Z_{i-k} \right) \right]. \tag{60}$$

Since the function $g(\cdot)$ is non-decreasing, (60) implies that the Ta-PAP is a non-decreasing function of the waiting times. This completes the proof.

APPENDIX C
PROOF OF LEMMA 3

Part (i) is proven in two steps:

Step 1: We will prove that $\bar{\Delta}_{\text{avg-opt}} \leq \beta$ if and only if $p(\beta) \leq 0$. If $\bar{\Delta}_{\text{avg-opt}} \leq \beta$, there exists a sampling policy $f = (Z_0, Z_1, \ldots) \in \mathcal{F}$ that is feasible for (17) and (18), which satisfies

$$\limsup_{n \to \infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}\left[\sum_{l=1}^{m} \int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau \right]}{\sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}]} \leq \beta. \tag{61}$$

Hence,

$$\limsup_{n \to \infty} \frac{\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\left[\sum_{l=1}^{m} \int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau - \beta(Z_i + Y_{i+1}) \right]}{\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}]} \leq 0. \tag{62}$$

Since Z_i's and Y_i's are bounded and positive and $\mathbb{E}[Y_i] > 0$ for all i, we have $0 < \liminf_{n \to \infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}] \leq \limsup_{n \to \infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}] \leq q$ for some $q \in \mathbb{R}^+$. By this, we get

$$\limsup_{n \to \infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\left[\sum_{l=1}^{m} \int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau - \beta(Z_i + Y_{i+1}) \right] \leq 0. \tag{63}$$

Therefore, $p(\beta) \leq 0$.

In the reverse direction, if $p(\beta) \leq 0$, then there exists a sampling policy $f = (Z_0, Z_1, \ldots) \in \mathcal{F}$ that is feasible for (17) and (18), which satisfies (63). Since we have $0 < \liminf_{n \to \infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}] \leq \limsup_{n \to \infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}] \leq q$, we can divide (63) by $\liminf_{n \to \infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}]$ to get

(62), which implies (61). Hence, $\bar{\Delta}_{\text{avg-opt}} \leq \beta$. By this, we have proven that $\bar{\Delta}_{\text{avg-opt}} \leq \beta$ if and only if $p(\beta) \leq 0$.

Step 2: We need to prove that $\bar{\Delta}_{\text{avg-opt}} < \beta$ if and only if $p(\beta) < 0$. This statement can be proven by using the arguments in Step 1, in which "$\leq$" should be replaced by "$<$". Finally, from the statement of Step 1, it immediately follows that $\bar{\Delta}_{\text{avg-opt}} > \beta$ if and only if $p(\beta) > 0$. This completes part (i).

Part(ii): We first show that each optimal solution to (17) is an optimal solution to (18). By the claim of part (i), $p(\beta) = 0$ is equivalent to $\bar{\Delta}_{\text{avg-opt}} = \beta$. Suppose that policy $f = (Z_0, Z_1, \ldots) \in \mathcal{F}$ is an optimal solution to (17). Then, $\Delta_{\text{avg}(\pi_{\text{MAF}}, f)} = \bar{\Delta}_{\text{avg-opt}} = \beta$. Applying this in the arguments of (61)-(63), we can show that policy f satisfies

$$\limsup_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}\left[\sum_{l=1}^{m} \int_{a_{li}}^{a_{li} + Z_i + Y_{i+1}} g(\tau)d\tau - \beta(Z_i + Y_{i+1}) \right] = 0. \tag{64}$$

This and $p(\beta) = 0$ imply that policy f is an optimal solution to (18).

Similarly, we can prove that each optimal solution to (18) is an optimal solution to (17). By this, part (ii) is proven.

APPENDIX D
PROOF OF PROPOSITION 4

According to [39, Proposition 4.2.1 and Proposition 4.2.6], it is enough to show that for every two states $\mathbf{s}$ and $\mathbf{s}'$, there exists a stationary deterministic policy f such that for some k, we have

$$\mathbb{P}\left[\mathbf{s}(k) = \mathbf{s}' | \mathbf{s}(0) = \mathbf{s}, f \right] > 0. \tag{65}$$

From the state evolution equation (20), we can observe that any state in $\mathcal{S}$ can be represented in terms of the packet waiting and service times. This implies (65). To clarify this, let us consider a system with 3 sources. Assume that the elements of state $\mathbf{s}'$ are as follows:

$$\begin{aligned}
a'_{[1]} &= y_3 + z_2 + y_2 + z_1 + y_1, \\
a'_{[2]} &= y_3 + z_2 + y_2, \\
a'_{[3]} &= y_3,
\end{aligned} \tag{66}$$

where y_i's and z_i's are any arbitrary elements in $\mathcal{Y}$ and $\mathcal{Z}$, respectively. Then, we will show that from any arbitrary state $\mathbf{s} = (a_{[1]}, a_{[2]}, a_{[3]})$, a sequence of service and waiting times can be followed to reach state $\mathbf{s}'$. If we have $Z_0 = z_1$, $Y_1 = y_1$, $Z_1 = z_1$, $Y_2 = y_2$, $Z_2 = z_2$, and $Y_3 = y_3$, then according to (20), we have in the first stage

$$\begin{aligned}
a_{[1]1} &= a_{[2]} + z_1 + y_1, \\
a_{[2]1} &= a_{[3]} + z_1 + y_1, \\
a_{[3]1} &= y_1,
\end{aligned} \tag{67}$$

and in the second stage, we have

$$
\begin{aligned}
a_{[1]2} &= a_{[3]} + z_1 + y_2 + z_1 + y_1, \\
a_{[2]2} &= y_2 + z_1 + y_1, \\
a_{[3]2} &= y_2,
\end{aligned}
\tag{68}
$$

and in the third stage, we have

$$
\begin{aligned}
a_{[1]3} &= y_3 + z_2 + y_2 + z_1 + y_1 = a'_{[1]}, \\
a_{[2]3} &= y_3 + z_2 + y_2 = a'_{[2]}, \\
a_{[3]3} &= y_3 = a'_{[3]}.
\end{aligned}
\tag{69}
$$

Hence, a stationary deterministic policy f can be designed to reach state $\mathbf{s}'$ from state $\mathbf{s}$ in 3 stages, if the aforementioned sequence of service times occurs. This implies that

$$
\mathbb{P}\left[\mathbf{s}(3) = \mathbf{s}'|\mathbf{s}(0) = \mathbf{s}, f\right] = \prod_{i=1}^{3} \mathbb{P}(Y_i = y_i) > 0,
\tag{70}
$$

where we have used that Y_i's are *i.i.d.*[8] The previous argument can be generalized to any number of sources. In particular, a forward induction over m can be used to show the result, where (65) trivially holds for $m = 1$, and the previous argument can be used to show that (65) holds for any general m. This completes the proof.

APPENDIX E
PROOF OF PROPOSITION 5

We prove Proposition 5 into two steps:

Step 1: We first show that $h(\mathbf{s})$ is non-decreasing in $\mathbf{s}$. To do so, we show that $J_\alpha(\mathbf{s})$, defined in (29), is non-decreasing in $\mathbf{s}$, which together with (28) imply that $h(\mathbf{s})$ is non-decreasing in $\mathbf{s}$.

Given an initial state $\mathbf{s}(0)$, the total expected discounted cost under a sampling policy $f \in \mathcal{F}$ is given by

$$
J_\alpha(\mathbf{s}(0); f) = \limsup_{n \to \infty} \mathbb{E}\left[\sum_{i=0}^{n-1} \alpha^i C(\mathbf{s}(i), Z_i)\right],
\tag{71}
$$

where $0 < \alpha < 1$ is the discount factor. The optimal total expected α-discounted cost function is defined by

$$
J_\alpha(\mathbf{s}) = \min_{f \in \mathcal{F}} J_\alpha(\mathbf{s}; f), \quad \mathbf{s} \in \mathcal{S}.
\tag{72}
$$

A policy is said to be α-optimal if it minimizes the total expected α-discounted cost. The discounted cost optimality equation of $J_\alpha(\mathbf{s})$ is discussed below.

---

[8]We assume that all elements in $\mathcal{Y}$ have a strictly positive probability, where the elements with zero probability can be removed without affecting the proof.

Proposition 12. *The optimal total expected α-discounted cost $J_\alpha(\mathbf{s})$ satisfies*

$$J_\alpha(\mathbf{s}) = \min_{z \in \mathcal{Z}} C(\mathbf{s}, z) + \alpha \sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z) J_\alpha(\mathbf{s}'). \tag{73}$$

Moreover, a stationary deterministic policy that attains the minimum in equation (73) for each $\mathbf{s} \in \mathcal{S}$ will be an α-optimal policy. Also, let $J_{\alpha,0}(\mathbf{s}) = 0$ for all $\mathbf{s}$ and any $n \geq 0$,

$$J_{\alpha,n+1}(\mathbf{s}) = \min_{z \in \mathcal{Z}} C(\mathbf{s}, z) + \alpha \sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z) J_{\alpha,n}(\mathbf{s}'). \tag{74}$$

Then, we have $J_{\alpha,n}(\mathbf{s}) \to J_\alpha(\mathbf{s})$ as $n \to \infty$ for every $\mathbf{s}$, and α.

Proof. Since we have bounded cost per stage, the proposition follows directly from [39, Proposition 1.2.2 and Proposition 1.2.3], and [46]. $\square$

Next, we use the optimality equation (73) and the value iteration in (74) to prove that $J_\alpha(\mathbf{s})$ is non-decreasing in $\mathbf{s}$.

Lemma 13. *The optimal total expected α-discounted cost function $J_\alpha(\mathbf{s})$ is non-decreasing in $\mathbf{s}$.*

Proof. We use induction on n in equation (74) to prove Lemma 13. Obviously, the result holds for $J_{\alpha,0}(\mathbf{s})$.

Now, assume that $J_{\alpha,n}(\mathbf{s})$ is non-decreasing in $\mathbf{s}$. We need to show that for any two states $\mathbf{s}_1$ and $\mathbf{s}_2$ with $\mathbf{s}_1 \leq \mathbf{s}_2$, we have $J_{\alpha,n+1}(\mathbf{s}_1) \leq J_{\alpha,n+1}(\mathbf{s}_2)$. First, we note that, since the age-penalty function $g(\cdot)$ is non-decreasing, the expected cost per stage $C(\mathbf{s}, z)$ is non-decreasing in $\mathbf{s}$, i.e., we have

$$C(\mathbf{s}_1, z) \leq C(\mathbf{s}_2, z). \tag{75}$$

From the state evolution equation (20) and the transition probability equation (22), the second term of the right-hand side (RHS) of (74) can be rewritten as

$$\sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z) J_{\alpha,n}(\mathbf{s}') = \sum_{y \in \mathcal{Y}} \mathbb{P}(Y = y) J_{\alpha,n}(\mathbf{s}'(z, y)), \tag{76}$$

where $\mathbf{s}'(z, y)$ is the next state from state $\mathbf{s}$ given the values of z and y. Also, according to the state evolution equation (20), if the next states of $\mathbf{s}_1$ and $\mathbf{s}_2$ for given values of z and y are $\mathbf{s}'_1(z, y)$ and $\mathbf{s}'_2(z, y)$, respectively, then we have $\mathbf{s}'_1(z, y) \leq \mathbf{s}'_2(z, y)$. This implies that

$$\sum_{y \in \mathcal{Y}} \mathbb{P}(Y = y) J_{\alpha,n}(\mathbf{s}'_1(z, y)) \leq \sum_{y \in \mathcal{Y}} \mathbb{P}(Y = y) J_{\alpha,n}(\mathbf{s}'_2(z, y)), \tag{77}$$

where we have used the induction assumption that $J_{\alpha,n}(\mathbf{s})$ is non-decreasing in $\mathbf{s}$. Using (75), (77), and the fact that the minimum operator in (74) retains the non-decreasing property, we conclude that

$$J_{\alpha,n+1}(\mathbf{s}_1) \leq J_{\alpha,n+1}(\mathbf{s}_2). \tag{78}$$

This completes the proof. $\square$

Step 2: We use Step 1 to prove Proposition 5. From Step 1, we have that $h(\mathbf{s})$ is non-decreasing in $\mathbf{s}$.

Similar to Step 1, this implies that the second term of the right-hand side (RHS) of (26) ($\sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z)h(\mathbf{s}')$) is non-decreasing in $\mathbf{s}'$. Moreover, from the state evolution (20), we can notice that, for any state $\mathbf{s}$, the next state $\mathbf{s}'$ is increasing in z. This argument implies that the second term of the right-hand side (RHS) of (26) ($\sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z)h(\mathbf{s}')$) is increasing in z. Thus, the value of $z \in \mathcal{Z}$ that achieves the minimum value of this term is zero. If, for a given state $\mathbf{s}$, the value of $z \in \mathcal{Z}$ that achieves the minimum value of the cost function $C(\mathbf{s}, z)$ is zero, then $z = 0$ solves the RHS of (26). In the sequel, we obtain the condition on $\mathbf{s}$ under which $z = 0$ minimizes the cost function $C(\mathbf{s}, z)$.

Now, we focus on the cost function $C(\mathbf{s}, z)$. In order to obtain the optimal z that minimizes this cost function, we need to obtain the one-sided derivative of it. The one-sided derivative of a function q in the direction of ω at z is given by

$$\delta q(z; \omega) \triangleq \lim_{\epsilon \to 0^+} \frac{q(z + \epsilon\omega) - q(z)}{\epsilon}. \tag{79}$$

Let $r(\mathbf{s}, z, Y) = \sum_{l=1}^{m} \int_{a_{[l]}}^{a_{[l]}+z+Y} g(\tau)d\tau$. Since $r(\mathbf{s}, z, Y)$ is the sum of integration of a non-decreasing function $g(\cdot)$, it is easy to show that $r(\mathbf{s}, z, Y)$ is convex. According to [15, Lemma 4], the function $q(z) = \mathbb{E}_Y[r(\mathbf{s}, z, Y)]$ is convex as well. Hence, the one-sided derivative $\delta q(z; \omega)$ of $q(z)$ exists [47, p.709]. Moreover, since $z \to r(\mathbf{s}, z, Y)$ is convex, the function $\epsilon \to [r(\mathbf{s}, z + \epsilon\omega, Y) - r(\mathbf{s}, z, Y)]/\epsilon$ is non-decreasing and bounded from above on $(0, \theta]$ for some $\theta > 0$ [48, Proposition 1.1.2(i)]. Using the monotone convergence theorem [49, Theorem 1.5.6], we can interchange the limit and integral operators in $\delta q(z; \omega)$ such that

$$\begin{aligned}
\delta q(z; \omega) &= \lim_{\epsilon \to 0^+} \frac{1}{\epsilon} \mathbb{E}_Y[r(\mathbf{s}, z + \epsilon\omega, Y) - r(\mathbf{s}, z, Y)] \\
&= \mathbb{E}_Y\left[\lim_{\epsilon \to 0^+} \frac{1}{\epsilon}\{r(\mathbf{s}, z + \epsilon\omega, Y) - r(\mathbf{s}, z, Y)\}\right] \\
&= \mathbb{E}_Y\left[\lim_{t \to z^+} \sum_{l=1}^{m} g(a_{[l]} + t + Y)w\mathbb{1}_{\{\omega > 0\}} + \lim_{t \to z^-} \sum_{l=1}^{m} g(a_{[l]} + t + Y)w\mathbb{1}_{\{\omega < 0\}}\right] \\
&= \lim_{t \to z^+} \mathbb{E}_Y\left[\sum_{l=1}^{m} g(a_{[l]} + t + Y)w\mathbb{1}_{\{\omega > 0\}}\right] + \lim_{t \to z^-} \mathbb{E}_Y\left[\sum_{l=1}^{m} g(a_{[l]} + t + Y)w\mathbb{1}_{\{\omega < 0\}}\right],
\end{aligned}$$

where $\mathbb{1}_E$ is the indicator function of event E. According to [47, p.710] and the convexity of $q(z)$, z is optimal to the cost function $C(\mathbf{s}, z)$ if and only if

$$\delta q(z; \omega) - \bar{\Delta}_{\text{avg-opt}}\omega \geq 0, \ \forall \omega \in \mathbb{R}. \tag{80}$$

As ω in (80) is an arbitrary real number, considering $\omega = 1$, (80) becomes

$$\lim_{t \to z^+} \mathbb{E}_Y\left[\sum_{l=1}^{m} g(a_{[l]} + t + Y)\right] - \bar{\Delta}_{\text{avg-opt}} \geq 0. \tag{81}$$

Likewise, considering $\omega = -1$, (80) implies

$$\lim_{t \to z^-} \mathbb{E}_Y\left[\sum_{l=1}^{m} g(a_{[l]} + t + Y)\right] - \bar{\Delta}_{\text{avg-opt}} \leq 0. \tag{82}$$

Since $g(\cdot)$ is non-decreasing, we get from (80)-(82) that z must satisfy

$$\mathbb{E}_Y\left[\sum_{l=1}^m g(a_{[l]} + t + Y)\right] - \bar{\Delta}_{\text{avg-opt}} \geq 0, \text{ if } t > z, \tag{83}$$

$$\mathbb{E}_Y\left[\sum_{l=1}^m g(a_{[l]} + t + Y)\right] - \bar{\Delta}_{\text{avg-opt}} \leq 0, \text{ if } t < z. \tag{84}$$

Subsequently, the smallest z that satisfies (83)-(84) is

$$z = \inf\left\{t \geq 0 : \mathbb{E}_Y\left[\sum_{l=1}^m g(a_{[l]} + t + Y)\right] \geq \bar{\Delta}_{\text{avg-opt}}\right\}. \tag{85}$$

According to (85), Since $g(\cdot)$ is non-decreasing, if $\mathbb{E}_Y\left[\sum_{l=1}^m g(a_{[l]} + Y)\right] \geq \bar{\Delta}_{\text{avg-opt}}$, then $z = 0$ minimizes $C(\mathbf{s}, z)$. This completes the proof.

APPENDIX F
PROOF OF THEOREM 8

We use the threshold test $A_s \geq (\bar{\Delta}_{\text{avg-opt}} - m\mathbb{E}[Y])$, in Proposition 7, to prove Theorem 8. From the state evolution (20), we can deduce that for any state $\mathbf{s} \in \mathcal{S}$, we have

$$a_{[l]} \geq (m - l + 1)y_{\text{inf}}, \ \forall l = 1, \dots, m. \tag{86}$$

This implies

$$A_s \geq \sum_{l=1}^m l y_{\text{inf}} = \frac{m(m+1)}{2} y_{\text{inf}}, \ \forall \mathbf{s} \in \mathcal{S}. \tag{87}$$

Moreover, it is easy to show that the total-average age of the zero-wait sampler, when the scheduling policy is fixed to the MAF scheduler, is given by

$$\bar{\Delta}_0 = \frac{\frac{m(m+1)}{2}\mathbb{E}[Y]^2 + \frac{m}{2}\mathbb{E}[Y^2]}{\mathbb{E}[Y]}. \tag{88}$$

Since $\bar{\Delta}_0 \geq \bar{\Delta}_{\text{avg-opt}}$, we have

$$\bar{\Delta}_0 - m\mathbb{E}[Y] \geq \bar{\Delta}_{\text{avg-opt}} - m\mathbb{E}[Y]. \tag{89}$$

Hence, if the following condition holds

$$\frac{m(m+1)}{2}y_{\text{inf}} \geq \frac{\frac{m(m+1)}{2}\mathbb{E}[Y]^2 + \frac{m}{2}\mathbb{E}[Y^2]}{\mathbb{E}[Y]} - m\mathbb{E}[Y], \tag{90}$$

which is equivalent to

$$y_{\text{inf}} \geq \frac{m(m-1)\mathbb{E}[Y]^2 + m\mathbb{E}[Y^2]}{m(m+1)\mathbb{E}[Y]}, \tag{91}$$

then we have $A_s \geq (\bar{\Delta}_{\text{avg-opt}} - m\mathbb{E}[Y])$ for all states $\mathbf{s} \in \mathcal{S}$. This implies that the zero-wait sampler is optimal under this condition. This completes the proof.

R<small>EFERENCES</small>

[1] A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, "Age-optimal sampling and transmission scheduling in multi-source systems," in *Proc. MobiHoc*, pp. 121–130.

[2] B. Adelberg, H. Garcia-Molina, and B. Kao, "Applying update streams in a soft real-time database system," in *ACM SIGMOD Record*, 1995, vol. 24, pp. 245–256.

[3] J. Cho and H. Garcia-Molina, "Synchronizing a database to improve freshness," in *ACM SIGMOD Record*, 2000, vol. 29, pp. 117–128.

[4] L. Golab, T. Johnson, and V. Shkapenyuk, "Scheduling updates in a real-time stream warehouse," in *Proc. IEEE ICDE*, 2009, pp. 1207–1210.

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

[6] S. Kaul, R. D. Yates, and M. Gruteser, "Status updates through queues," in *Conf. on Info. Sciences and Systems*, Mar. 2012, pp. 1–6.

[7] M. Costa, M. Codreanu, and A. Ephremides, "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.

[8] 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.

[9] A. M. Bedewy, Y. Sun, and N. B. Shroff, "Optimizing data freshness, throughput, and delay in multi-server information-update systems," in *Proc. IEEE ISIT*, 2016, pp. 2569–2573.

[10] A. M. Bedewy, Y. Sun, and N. B. Shroff, "Minimizing the age of information through queues," *IEEE Trans. Inf. Theory*, vol. 65, no. 8, pp. 5215–5232, Aug 2019.

[11] A. M. Bedewy, Y. Sun, and N. B. Shroff, "Age-optimal information updates in multihop networks," in *Proc. IEEE ISIT*, 2017, pp. 576–580.

[12] A. M. Bedewy, Y. Sun, and N. B. Shroff, "The age of information in multihop networks," *IEEE/ACM Trans. Netw.*, vol. 27, no. 3, pp. 1248–1257, June 2019.

[13] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *Proc. IEEE ISIT*, 2015, pp. 3008–3012.

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

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

[16] Y. Sun and B. Cyr, "Sampling for data freshness optimization: Non-linear age functions," *Journal of Communications and Networks - special issue on the Age of Information*, vol. 21, no. 3, pp. 204–219, June 2019.

[17] L. Ran, W. Junfeng, W. Haiying, and L. Gechen, "Design method of can bus network communication structure for electric vehicle," in *International Forum on Strategic Technology 2010*. IEEE, 2010, pp. 326–329.

[18] K. H. Johansson, M. Törngren, and L. Nielsen, "Vehicle applications of controller area network," in *Handbook of networked and embedded control systems*, pp. 741–765. Springer, 2005.

[19] R. D. Yates and S. Kaul, "Real-time status updating: Multiple sources," in *Proc. IEEE ISIT*, 2012, pp. 2666–2670.

[20] L. Huang and E. Modiano, "Optimizing age-of-information in a multi-class queueing system," in *Proc. IEEE ISIT*, 2015, pp. 1681–1685.

[21] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Trans. Inf. Theory*, vol. 65, no. 3, pp. 1807–1827, March 2018.

[22] N. Pappas, J. Gunnarsson, L. Kratz, M. Kountouris, and V. Angelakis, "Age of information of multiple sources with queue management," in *Proc. IEEE ICC*, 2015, pp. 5935–5940.

[23] A. M. Bedewy, Y. Sun, R. Singh, and N. B. Shroff, "Optimizing information freshness using low-power status updates via sleep-wake scheduling," *arXiv preprint arXiv:1910.00205*, 2019.

[24] Q. He, D. Yuan, and A. Ephremides, "Optimal link scheduling for age minimization in wireless systems," *IEEE Trans. Inf. Theory*, vol. 64, no. 7, pp. 5381–5394, July 2018.

[25] I. Kadota, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Minimizing the age of information in broadcast wireless networks," in *Proc. Allerton Conf.*, 2016, pp. 844–851.

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

[27] Y. Hsu, E. Modiano, and L. Duan, "Scheduling algorithms for minimizing age of information in wireless broadcast networks with random arrivals," *IEEE Trans. Mobile Comput.*, 2019.

[28] I. Kadota, A. Sinha, and E. Modiano, "Optimizing age of information in wireless networks with throughput constraints," in *Proc. IEEE INFOCOM*, 2018, pp. 1844–1852.

[29] Y. Hsu, "Age of information: Whittle index for scheduling stochastic arrivals," in *Proc. IEEE ISIT*, 2018, pp. 2634–2638.

[30] R. Li, A. Eryilmaz, and B. Li, "Throughput-optimal wireless scheduling with regulated inter-service times," in *Proc. IEEE INFOCOM*, 2013, pp. 2616–2624.

[31] Y. Sun, E. Uysal-Biyikoglu, and S. Kompella, "Age-optimal updates of multiple information flows," in *IEEE INFOCOM - the 1st Workshop on the Age of Information (AoI Workshop)*, 2018, pp. 136–141.

[32] R. D. Yates and S. K. Kaul, "Status updates over unreliable multiaccess channels," in *Proc. IEEE ISIT*, 2017, pp. 331–335.

[33] R. Talak, S. Karaman, and E. Modiano, "Distributed scheduling algorithms for optimizing information freshness in wireless networks," in *Proc. IEEE SPAWC*, 2018, pp. 1–5.

[34] R. Talak, S. Karaman, and E. Modiano, "Optimizing age of information in wireless networks with perfect channel state information," in *Proc. WiOpt*, 2018, pp. 1–8.

[35] R. Talak, I. Kadota, S. Karaman, and E. Modiano, "Scheduling policies for age minimization in wireless networks with unknown channel state," in *Proc. IEEE ISIT*, 2018, pp. 2564–2568.

[36] R. Talak, S. Karaman, and E. Modiano, "Optimizing information freshness in wireless networks under general interference constraints," in *Proc. MobiHoc*, 2018, pp. 61–70.

[37] M. Shaked and J. G. Shanthikumar, *Stochastic orders*, Springer Science & Business Media, 2007.

[38] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the wiener process over a channel with random delay," in *Proc. IEEE ISIT*, June 2017, pp. 321–325.

[39] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 2, Athena Scientific, Belmont, MA, 2 edition, 2001.

[40] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley & Sons, Hoboken, NJ, 2005.

[41] L. P. Kaelbling and L. P. Kaelbling, *Recent advances in reinforcement learning*, Springer, Boston, 1996.

[42] W. B. Powell, *Approximate Dynamic Programming: Solving the curses of dimensionality*, vol. 703, John Wiley & Sons, Hoboken, NJ, 2007.

[43] I. Bettesh and S. S. Shamai, "Optimal power and rate control for minimal average delay: The single-user case," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 4115–4141, Sep. 2006.

[44] R. Wang and V. K. Lau, "Delay-optimal two-hop cooperative relay communications via approximate mdp and distributive stochastic learning," *IEEE Trans. Inf. Theory*, Nov 2013.

[45] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Golden section search in one dimension," *Numerical Recipes in C: The Art of Scientific Computing*, p. 2, 1992.

[46] L. I. Sennott, "Average cost optimal stationary policies in infinite state markov decision processes with unbounded costs," *Operations Research*, vol. 37, no. 4, pp. 626–633, 1989.

[47] Dimitri P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Belmont, MA, 2 edition, 1999.

[48] D. Butnariu and A. N. Iusem, *Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization*, Kluwer Academic Publisher, Norwell, MA, USA, 2000.

[49] R. Durrett, *Probability: theory and examples*, Cambridge Univ. Press, Cambridge, U.K., 4 edition, 2010.