1 Sampling and Scheduling for Minimizing Age of Information of Multiple Sources

Ahmed M. Bedewy, Yin Sun, Sastry Kompella, and Ness B. Shroff,

This work has been supported in part by ONR grants N00014-17-1-2417 and N00014-15-1-2166, Army Research Office grants W911NF-14-1-0368 and MURI W911NF-12-1-0385, National Science Foundation grants CNS-1446582, CNS-1421576, CNS-1518829, and CCF-1813050, and a grant from the Defense Thrust Reduction Agency HDTRA1-14-1-0058.
A. M. Bedewy is with the Department of ECE, The Ohio State University, Columbus, OH 43210 USA (e-mail: bedewy.2@osu.edu).
Y. Sun is with the Department of ECE, Auburn University, Auburn, AL 36849 USA (e-mail: yzs0078@auburn.edu).
S. Kompella is with Information Technology Division, Naval Research Laboratory, Washington, DC 20375 USA (e-mail: sk@ieee.org).
N. B. Shroff is with the Department of ECE and the Department of CSE, The Ohio State University, Columbus, OH 43210 USA (e-mail: shroff.11@osu.edu).

1.1 Abstract

In this chapter, 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 *Age of Information*, where all sources have the same age-penalty function. Sources take turns to generate update samples, 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 sample in its turn. We aim to find the optimal scheduler-sampler pairs that minimize the total-average age-penalty (Ta-AP).

We start the chapter by providing a brief explanation of the sampling problem in the light of single-source networks, as well as some useful insights and applications on age of information and its penalty functions. Then, we move on to the multi-source networks, where the problem becomes more challenging. We provide a detailed explanation of the model and the solution in this case. Finally, we conclude this chapter by providing an open question in this area and its inherent challenges.

1.2 Introduction

1.2.1 Sampling Problem in Single-source Networks

In many applications, a controller observes one or more continuous-time processes and takes proper actions depending on these processes' states. However, continuous observation and processing of these continuous-time processes could be costly and not always available. For example, in autonomous vehicles, a control unit may observe multiple processes simultaneously as shown in Fig. 1.1, where sources transmit information to a control unite via a wireless channel. Continuous observation of these processes is difficult because of the limited resources in terms of available bands and energy. One way to overcome this difficulty is to deal with samples from these observed processes. As a result, controlling the sampling times is very important to save precious system resources such as energy, channel use, computation resources, etc. Also, with the influence of uncertain factors, the channel delay between the controller and the observed processes varies with time. Hence, it is necessary to introduce a flexible sampling control. One importing factor in this control system is the design of inter-sampling times. In particular, as the inter-sampling time increases, information at the destination about the observed processes becomes stale. This pushes us to

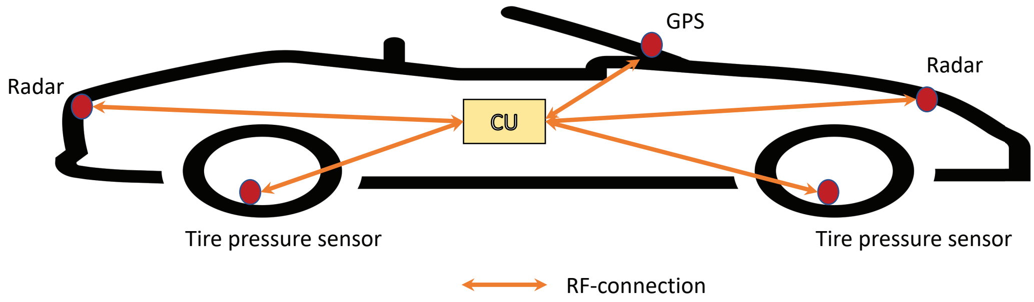


Figure 1.1 A control unite (CU) observes multiple processes, e.g., tires pressure, location, orientation, speed, etc., in an autonomous vehicle.

ask the following question: Is it true that increasing the sampling rate (i.e., decreasing the inter-sampling times) always improves information freshness? We will show that contrary to conventional wisdom this is not always the case. To measure data freshness, age of information metric is used, which is defined as the time elapsed since the most recently received sample was generated. Next, we discuss the sampling problem in the presence of a single source.

1.2.2 Sampling Problem in Single-source Networks

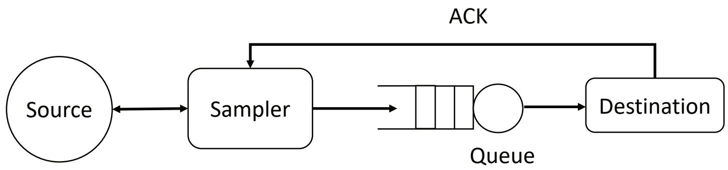


Figure 1.2 A single-source information update system.

In this section, we will focus on a simple setting which includes a single source transmitting sensed information to a receiver. Our goal will be to understand the sampling phenomena in the single source setting first and then use the insights to extend the analysis in Section 1.3 to the multi-source setting. In particular, consider a single-source information update system as shown in Fig. 1.2, where the source (e.g. a sensor) observes a time varying process. There is a sampler that determines when the source should generate samples from the observed process. This is known as the "generate-at-will" model [1–3] (i.e., samples can be generated at any time). The generated samples are, thereafter, sent to the destination via a channel with random delay. The channel is modeled as a single server First-Come, First-Served (FCFS) queue with i.i.d. service times. We use S_i and D_i to denote the generation time and the delivery time of the i-th generated sample, respectively. We suppose that the sampler has full knowledge of the idle/busy state of the server via acknowledgments (ACKs) from the destination with zero delay. Our target is to design the sampling times $(S_1, S_2, \dots)$ such that age of information at the destination (or a non-decreasing penalty function of it) is minimized. Age of information is defined as follows [4–7]: At time t, if the freshest sample at the destination was generated at time $U(t) = \max\{S_i : D_i \leq t\}$, then the age $\Delta(t)$ is defined as

$$\Delta(t) = t - U(t). \tag{1.1}$$

As shown in Fig. 1.3, the age increases linearly with t but is reset to a smaller value with the delivery of a fresher sample. Next, we show that this sampling problem is not trivial and the solution may counter the common wisdom.

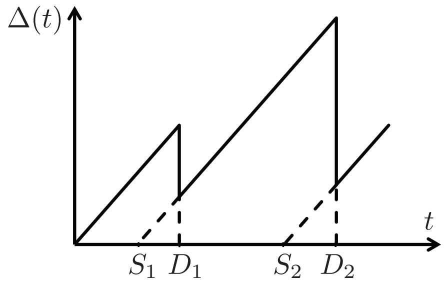


Figure 1.3 Sample path of the age process $\Delta(t)$.

1.2.3 Counter Intuitive Phenomenon of the Optimal Sampler

One may think that, as long as the sampler knows the idle/busy state of the server in real-time, then an intuitive solution is the zero-wait sampler, in which a sample is generated as soon as the server becomes idle (i.e., $S_{i+1} = D_i$). Clearly, this zero-wait sampler is throughput and delay optimal. However, surprisingly, this policy does not always minimize age of information. The following example reveals the reasons behind this counter intuitive phenomenon:

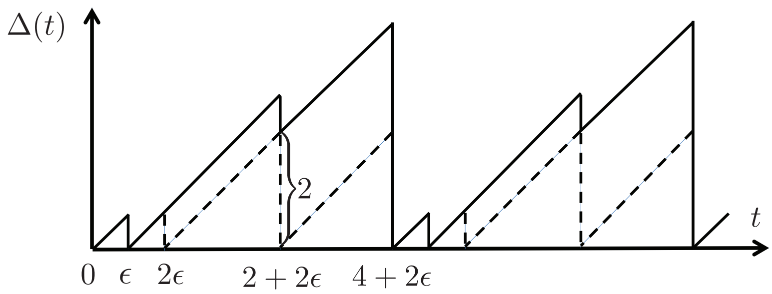


Figure 1.4 Evolution of the age $\Delta(t)$ under the ϵ-wait policy in the example [1].

Example [1]: Consider an information update system as shown in Fig. 1.2. Suppose that the sample transmission times are i.i.d. across the samples and are either 0 or 2 with probability 0.5[1]. For simplicity, consider the following realization of the sample transmission times:

$$0, 0, 2, 2, 0, 0, 2, 2, \ldots$$

Now, if the zero-wait sampler is followed, then Sample 1 is generated at time 0 and delivered at time 0. After the delivery of Sample 1 at time 0, Sample 2 is generated which occurs at time 0 as well. Hence, Samples 1 and 2 are both generated at the same time (time 0). As a result, Sample 2 does not bring any new information to the destination after the delivery of Sample 1. In other words, under zero-wait sampler, having zero transmission times has not been exploited well, but rather has resulted in wasted system resources (some samples with stale information are generated, and hence are wasted). This issue is repeated as follows: Whenever a sample has a zero transmission time, the next generated sample does not carry new information, and hence, is wasted. This raises an important question: Can we do better?

The answer is yes. For the sake of comparison, let us consider an ϵ-wait policy, in which the sampler waits for ϵ seconds after each sample with a zero transmission time, and does not wait otherwise. Note that this policy is a causal policy as it just needs to know the transmission time of the last delivered sample to specify the generation time of the next one. Fig. 1.4 illustrates the age evolution under the ϵ-wait policy for the above realization of the sample transmission times. From Fig. 1.4, we can compute the time-average age of the ϵ-wait policy, which is given by

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

[1] The 0 transmission time here is just chosen for the simplicity of the illustration. Indeed, it represents the transmission times that are extremely small.

If we set the waiting time $\epsilon = .5$, then the time-average age of the ϵ-wait policy is 1.85 seconds. Observe that, if we set $\epsilon = 0$, the ϵ-wait policy reduces to the zero-wait one, whose time-average age is 2 seconds. Hence, the zero-wait sampler is not age-optimal in this case and we can do better. This makes the problem non-trivial which needs to be attacked carefully. Next, we provide some motivation for using age of information as a metric and the benefit of using penalty functions.

1.2.4 Data Freshness in Real-time Applications

In applications such as networked monitoring and control systems, wireless sensor networks, and autonomous vehicles, the destination node must receive timely status updates so that it can make accurate decisions. For example, real-time knowledge about the location, orientation, and speed of motor vehicles is crucial to avoid collisions and reduce traffic congestion. In addition, fresh information about stock price and interest-rate is of ultimate importance in developing efficient business plans in the stock market. In light of this, age of information has emerged to provide a mathematical framework for modeling the data freshness at a particular destination. For this framework to be complete, it must capture the variation of the stale information harmful impact from one application to another. Here comes the importance of using penalty functions of the age. Next, we discuss some examples of the used age-penalty functions and their applications.

Examples of Age-penalty Functions and their Applications:
The impact of stale information depends on how fast the information source varies with time. For instance, the location of a motor vehicle is considered an information source that may vary quickly with time. In particular, a moving car with a speed of 65 mph will traverse almost 29 meters for 1 second. Hence, stale information has a dramatic serious impact on this situation. Meanwhile, the engine temperature in a vehicle, for example, is one of the information sources that may vary slowly with time, and hence a reading that was taken a few minutes ago is still valid for observing the engine health. This illustrates how the value of the fresh information varies from one application to another. Unfortunately, age alone cannot capture such variation. Thus, we desperately need an age-penalty function in such applications.

In this chapter, we use $g : [0, \infty) \to \mathbb{R}$ to denote the non-decreasing age-penalty function in use. Note that $g(\cdot)$ does not have to be convex or continuous. As mentioned, $g(\cdot)$ is used to represent the level of dissatisfaction of data staleness in different applications based on their demands. Here are some examples of $g(\cdot)$:

- 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.
- An indicator function $g(x) = \mathbb{1}(x > q)$ can be used to indicate the dissatisfaction of the violation of an age threshold q.

Besides the above examples, a recent survey [8] showed that, under certain conditions, information freshness metrics expressed in terms of auto-correlation functions, the estimation error of signal values, and mutual information, are monotonic functions of the age. To make it clearer, let us consider the following example:

Auto-correlation function of signals: Consider the model that is illustrated in Fig. 1.2. Let $X_t \in \mathbb{R}$ represent the process that is being observed by the source. The transmitted samples from the sampler are used for estimating the process X_t at the destination. At time t, the most recent received sample is generated at time $t - \Delta(t)$, where $\Delta(t)$ is the age of the most recent received sample. If the samples at t and $t - \Delta(t)$ are correlated, then the accuracy of the estimation will depend on the value of $\Delta(t)$. In other words, the smaller the value of $\Delta(t)$, the more accurate the estimate. Hence, the auto-correlation function $\mathbb{E}[X_t X_{t-\Delta(t)}]$ can be used to evaluate the freshness of the sample $X_{t-\Delta(t)}$ [9]. For some stationary processes, $|\mathbb{E}[X_t X_{t-\Delta(t)}]|$ becomes a non-negative, non-increasing function of the age $\Delta(t)$, which can be considered an age-penalty function.

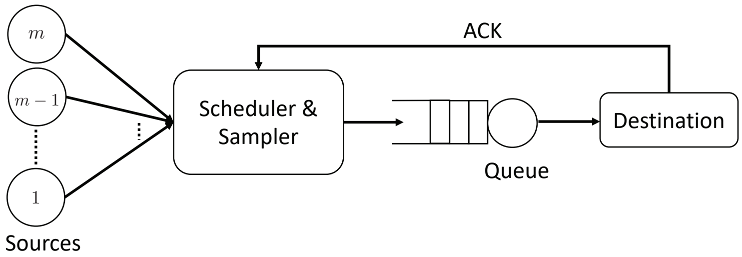


Figure 1.5 System model

1.3 Multi-source Sampling and Scheduling Problem

In real life applications, a controller may need observations from multiple sources in order to make accurate decisions. In autonomous vehicles, for example, many electronic control units (ECUs) are connected to one or more sensors and actuators via a controller area network (CAN) bus [10, 11]. As the vehicles and commercial trucks get smarter, the number of needed sensors increases and can reach up to 200 sensors per vehicle [12]. With such a large number of sensors, some of them may send their information via a shared channel. Hence, the considered single-source model above cannot be fully applied in such applications. In other words, besides the sampler, a scheduler is needed to handle the transmission order of the sources. This leads to the need for designing a sampler and a scheduler that can jointly optimize the data freshness. This makes the optimization problem more challenging. In this section, we present in detail the sampling problem in multi-source networks [13,14]. We start by providing some useful notations and definitions.

1.3.1 Notations 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 1.1. **Univariate Stochastic Ordering:** [15] Let X and Y be two random variables. Then, X is said to be stochastically smaller than Y (denoted as $X \leq_{\text{st}} Y$), if

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

Definition 1.2. **Multivariate Stochastic Ordering:** [15] Let $\mathbf{X}$ and $\mathbf{Y}$ be two random vectors. Then, $\mathbf{X}$ is said to be stochastically smaller than $\mathbf{Y}$ (denoted as $\mathbf{X} \leq_{\text{st}} \mathbf{Y}$), 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 1.3. **Stochastic Ordering of Stochastic Processes:** [15] 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 $\{X(t), t \in [0, \infty)\} \leq_{\text{st}} \{Y(t), t \in [0, \infty)\}$), if, for all choices of an integer n and $t_1 < t_2 < \ldots < t_n$ in $[0, \infty)$, it holds that

$$(X(t_1), X(t_2), \ldots, X(t_n)) \leq_{\text{st}} (Y(t_1), Y(t_2), \ldots, Y(t_n)), \tag{1.2}$$

where the multivariate stochastic ordering in (1.2) was defined in Definition 1.2.

1.3.2 Multi-source Network Model

We consider a status update system with m sources as shown in Fig. 1.5, where each source observes a time-varying process. Sources take turns to generate samples, and forward the samples to their destinations one-by-one through a shared error-free channel with random delay. Hence, a decision maker consists

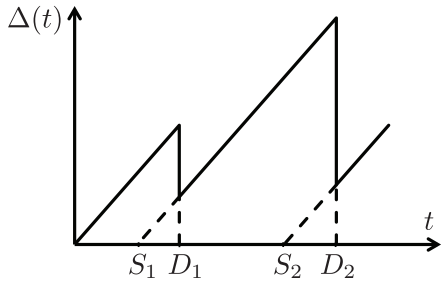


Figure 1.6 The age $\Delta_l(t)$ of source l, where we suppose that the first and third samples are generated from source l, i.e., $r_1 = r_3 = l$.

of a scheduler, that chooses the update order of the sources, and a sampler, that determines when a source should generate a new sample in its turn.

We use S_i to denote the generation time of the i-th generated sample from all sources, called sample i. Moreover, we use r_i to represent the source index from which sample i is generated. The channel is modeled as an FCFS queue with random *i.i.d.* service time Y_i, where Y_i represents the service time of sample 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 a sample is generated while the server is busy, it needs to wait in the queue until its transmission opportunity, and becomes stale while waiting. Hence, there is no loss of optimality to avoid generating a new sample during the busy periods. As a result, a sample is served immediately once it is generated. Let D_i denote the delivery time of sample i, where $D_i = S_i + Y_i$. After the delivery of sample i at time D_i, the decision maker may insert a waiting time Z_i before generating a new sample (hence, $S_{i+1} = D_i + Z_i$)[2], where $Z_i \in \mathcal{Z}$, and $\mathcal{Z} \subset \mathbb{R}^+$ is a finite and bounded set[3].

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

$$U_l(t) = \max\{S_i : r_i = l, D_i \le t\}. \tag{1.3}$$

Hence, the age of source l is defined as

$$\Delta_l(t) = t - U_l(t), \tag{1.4}$$

which is plotted in Fig. 1.6. We suppose that the age $\Delta_l(t)$ is right-continuous. Moreover, we assume that the initial age values $\Delta_l(0^-)$ at time $t = 0^-$ for all l are known to the system. The age process for source l is given by $\{\Delta_l(t), t \ge 0\}$. 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 $x < \infty$.

1.3.3 Decision Policies in Multi-source Networks

A decision policy, denoted by d, controls the following: i) the scheduler, denoted by π, that determines the update order of the sources $\pi \triangleq (r_1, r_2, \ldots)$, ii) the sampler, denoted by f, that controls the sampling 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.

A decision policy acts as follows: After each delivery, the decision maker determines the source to be served, and then imposes a waiting time before the generation of the new sample. Next, we present our optimization problem.

[2] We assume that $D_0 = 0$. Thus, $S_1 = Z_0$.
[3] We assume that $0 \in \mathcal{Z}$.

1.3.4 Problem Formulation and Challenges

In this chapter, we aim to minimize the total-average age-penalty per unit time (Ta-AP). Consider the time interval $[0, D_n]$, the Ta-AP is defined for any decision policy $d = (\pi, f)$ as

$$\Delta_{\text{avg}}(\pi, f) = \limsup_{n\to\infty} \frac{\mathbb{E}\left[\sum_{l=1}^{m} \int_0^{D_n} g\left(\Delta_l(t)\right) dt\right]}{\mathbb{E}\left[D_n\right]}. \tag{1.5}$$

Hence, our optimization problem can be cast as

$$\bar{\Delta}_{\text{avg-opt}} \triangleq \min_{\pi\in\Pi, f\in\mathcal{F}} \Delta_{\text{avg}}(\pi, f), \tag{1.6}$$

where $\bar{\Delta}_{\text{avg-opt}}$ is the optimum objective values of Problem (1.6).

Due to the large decision policy space, the optimization problem is quite challenging. In other words, we need to seek the optimal decision policy that controls both the scheduler and sampler to minimize the Ta-AP. Moreover, the possible correlation between the optimal actions of the sampler and the scheduler makes this challenge more difficult. Thus, we have to find a way to tackle this challenge. To that end, we develop an important separation principle that helps us to bypass this difficulty, which is presented next.

1.4 Optimal Sampling and Scheduling Design

1.4.1 Separation Principle and Optimal Scheduler

We show that our optimization problem in (1.6) has an important separation principle: Given the sampling times, the Maximum Age First (MAF) scheduler provides the best age performance compared to any other scheduler. What then remains to be addressed is the question of finding the best sampler that solves Problem (1.6), given that the scheduler is fixed to the MAF. We start by defining the MAF scheduler as follows:

Definition 1.4 ([16–20]). Maximum Age First (MAF) 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 $f \in \mathcal{F}$*

$$\Delta_{\text{avg}}(\pi_{MAF}, f) = \min_{\pi\in\Pi} \Delta_{\text{avg}}(\pi, f). \tag{1.7}$$

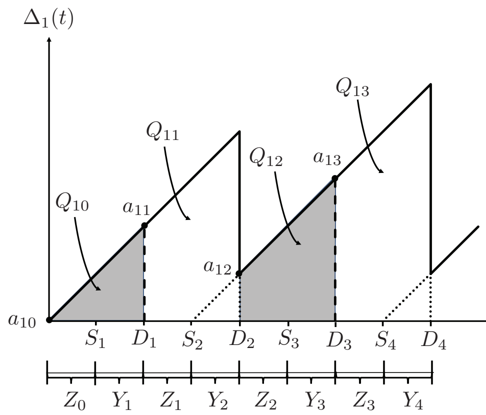
That is, the MAF scheduler minimizes the Ta-AP in (1.5) *among all schedulers in* Π.

Proof We use a sample-path method to prove Proposition 1 as follows: Given any sampler, that controls the sampling times, the scheduler only controls from which source a sample is generated. We couple the policies such that the sample delivery times are fixed under all decision policies. In the MAF scheduler, a source with maximum age becomes the source with minimum age among the m sources after each delivery. Under any arbitrary scheduler, a sample can be generated from any source, which is not necessarily the one with the 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 1.7.1. □

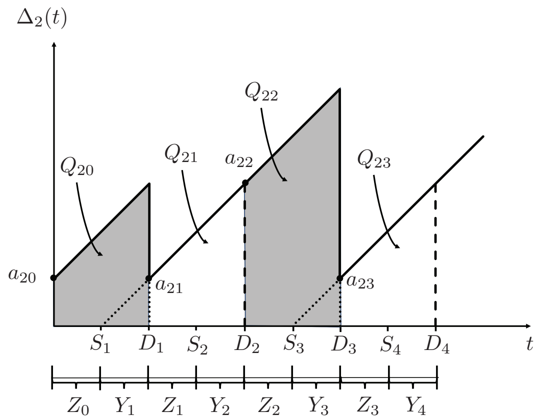
Proposition 1 concludes the separation principle that the optimal sampler can be optimized separately, given that the scheduling policy is fixed to the MAF scheduler. Hence, the optimization problem (1.6) can be rewritten as:

$$\bar{\Delta}_{\text{avg-opt}} \triangleq \min_{f\in\mathcal{F}} \Delta_{\text{avg}}(\pi_{\text{MAF}}, f). \tag{1.8}$$

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 sample $i+1$ at time $S_{i+1} = D_i+Z_i$ from the source with the maximum age at time $t = D_i$. Sample $i+1$ 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



(a) The age evolution of source 1.



(b) The age evolution of source 2.

Figure 1.7 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 Sample 1 is generated from Source 2, which is delivered at time D_1. Then, Source 1 is served and Sample 2 is generated from Source 1, which is delivered at time D_2. The same operation is repeated over time.

with time and the age processes evolve accordingly. An example of age processes evolution is shown in Fig. 1.7. Next, we shift our focus to obtain the optimal sampler for Problem (1.8).

1.4.2 Optimal sampler in Multi-source Networks

Now, we fix the scheduling policy to the MAF scheduler, and seek for the optimal sampler for minimizing the Ta-AP. A naive solution for the optimal sampler would be the zero-wait policy, i.e., $Z_i = 0$ for all i. However, we showed with a counter example in Section 1.2.3 that this is not necessarily true. Thus, we need to be careful in obtaining this optimal sampler. Since solving Problem (1.8) in the current form is challenging, we reformulate it as an equivalent semi-Markov decision problem (SMDP). Next, we discuss this reformulation in detail.

Reformulation of the Optimal Sampling Problem

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. 1.7 [4], 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,

$$\int_0^{D_n} g(\Delta_l(t))dt = \sum_{i=0}^{n-1} Q_{li}, \tag{1.9}$$

where

$$Q_{li} = \int_{D_i}^{D_{i+1}} g(\Delta_l(t))dt = \int_{D_i}^{D_i+Z_i+Y_{i+1}} g(\Delta_l(t))dt. \tag{1.10}$$

Let a_{li} denote the age value of source l at time D_i, i.e., $a_{li} = \Delta_l(D_i)^5$. Hence, for $t \in [D_i, D_{i+1})$, we have

$$\Delta_l(t) = t - U_l(t) = t - (D_i - a_{li}), \tag{1.11}$$

where $(D_i - a_{li})$ represents the generation time of the last delivered sample from source l before time D_{i+1}. By performing change of variables in (1.10), we get

$$Q_{li} = \int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau. \tag{1.12}$$

[4] Observe that a special age-penalty function is depicted in Fig. 1.7, where we choose $g(x) = x$ for simplicity.
[5] Since the age process is right-continuous, if sample 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.

Hence, the Ta-AP can be rewritten as

$$\limsup_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}\left[\sum_{l=1}^{m} \int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau\right]}{\sum_{i=0}^{n-1} \mathbb{E}\left[Z_i+Y_{i+1}\right]}. \tag{1.13}$$

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

$$\bar{\Delta}_{\text{avg-opt}} \triangleq \min_{f\in\mathcal{F}} \limsup_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}\left[\sum_{l=1}^{m} \int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau\right]}{\sum_{i=0}^{n-1} \mathbb{E}\left[Z_i+Y_{i+1}\right]}. \tag{1.14}$$

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 (1.14) is hard to solve in the current form. Therefore, we reformulate it. We consider the following optimization problem with a parameter $\beta \geq 0$:

$$\Theta(\beta) \triangleq \min_{f\in\mathcal{F}} \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\left[\sum_{l=1}^{m} \int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau - \beta(Z_i+Y_{i+1})\right], \tag{1.15}$$

where $\Theta(\beta)$ is the optimal value of (1.15). The following lemma is motivated by [21]. It explains the relationship between Problems (1.14) and (1.15).

Lemma 1.5. *The following assertions are true:*

(i) $\bar{\Delta}_{\text{avg-opt}} \lesseqqgtr \beta$ *if and only if* $\Theta(\beta) \lesseqqgtr 0$.
(ii) If $\Theta(\beta) = 0$, *then the optimal sampling policies that solve* (1.14) *and* (1.15) *are identical.*

Proof See Appendix 1.7.2. □

As a result of Lemma 1.5, the solution to (1.14) can be obtained by solving (1.15) in a multi-layer manner: In the inner layer, we optimize Z_i's for a given β. Then, in the outer layer, we seek a $\beta = \bar{\Delta}_{\text{avg-opt}} \geq 0$ such that $\Theta(\bar{\Delta}_{\text{avg-opt}}) = 0$. Lemma 1.5 helps us to utilize the DP technique to obtain the optimal sampler. Note that without Lemma 1.5, it would be quite difficult to use the DP technique to solve (1.14) optimally.

Existence of an Optimal Stationary Deterministic Policy

We resort to the methodology proposed in [22]. When $\beta = \bar{\Delta}_{\text{avg-opt}}$, Problem (1.15) is equivalent to an average cost per stage problem. According to [22], we describe the components of this problem in detail below.

- **States:** At stage[6] i, the system state is specified by

$$\mathbf{s}(i) = (a_{[1]i}, \ldots, a_{[m]i}), \tag{1.16}$$

  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 sample $i+1$ is $Y_{i+1} = y$, then the evolution of the system state from $\mathbf{s}(i)$ to $\mathbf{s}(i+1)$ is as follows:

$$\begin{aligned} &a_{[m]i+1} = y, \\ &a_{[l]i+1} = a_{[l+1]i} + z + y, \ l = 1, \ldots, m-1. \end{aligned} \tag{1.17}$$

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

$$\mathbb{P}_{\mathbf{ss}'}(z) = \mathbb{P}(\mathbf{s}(i+1) = \mathbf{s}' | \mathbf{s}(i) = \mathbf{s}, Z_i = z), \ \mathbf{s}, \mathbf{s}' \in \mathcal{S}. \tag{1.18}$$

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

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

$$\mathbb{P}_{\mathbf{s}\mathbf{s}'}(z) = \begin{cases} \mathbb{P}(Y_{i+1} = y) & \text{if } a'_{[m]} = y \text{ and} \\ & a'_{[l]} = a_{[l+1]} + z + y \text{ for } l \neq m; \\ 0 & \text{else.} \end{cases} \tag{1.19}$$

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

$$C(\mathbf{s}(i), Z_i, Y_{i+1}) = \sum_{l=1}^{m} \int_{a_{[l]i}}^{a_{[l]i}+Z_i+Y_{i+1}} g(\tau)d\tau - \bar{\Delta}_{\text{avg-opt}}(Z_i + Y_{i+1}). \tag{1.20}$$

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

$$C(\mathbf{s}(i), Z_i) = \mathbb{E}_{Y_{i+1}} \left[C(\mathbf{s}(i), Z_i, Y_{i+1}) \right], \tag{1.21}$$

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

$$\limsup_{n\to\infty} \frac{1}{n} \mathbb{E} \left[\sum_{i=0}^{n-1} C(\mathbf{s}(i), Z_i) \right]. \tag{1.22}$$

We say that a sampling policy $f \in \mathcal{F}$ is *average-optimal* if it minimizes the average cost per stage in (1.22). 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 [22], there may not exist a stationary deterministic policy that is average-optimal. However, in the next proposition, we are able to show that there is a stationary deterministic policy that is average-optimal.

Proposition 2. *There exist a scalar λ and a function h that satisfy the following Bellman's equation:*

$$\lambda + h(\mathbf{s}) = \min_{z\in\mathcal{Z}} \left(C(\mathbf{s}, z) + \sum_{\mathbf{s}'\in\mathcal{S}} \mathbb{P}_{\mathbf{s}\mathbf{s}'}(z) h(\mathbf{s}') \right), \tag{1.23}$$

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

$$\lambda = \lim_{\alpha\to1} (1-\alpha) J_\alpha(\mathbf{s}), \forall \mathbf{s} \in \mathcal{S}, \tag{1.24}$$

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

$$h(\mathbf{s}) = \lim_{\alpha\to1} (J_\alpha(\mathbf{s}) - J_\alpha(\mathbf{o})), \forall \mathbf{s} \in \mathcal{S}, \tag{1.25}$$

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

$$J_\alpha(\mathbf{s}) = \min_{f\in\mathcal{F}} \limsup_{n\to\infty} \mathbb{E} \left[\sum_{i=0}^{n-1} \alpha^i C(\mathbf{s}(i), Z_i) \right], \mathbf{s}(0) = \mathbf{s} \in \mathcal{S}, \tag{1.26}$$

where $0 < \alpha < 1$ is the discount factor. Furthermore, there exists a stationary deterministic policy that attains the minimum in (1.23) *for each $\mathbf{s} \in \mathcal{S}$ and is average-optimal.*

Proof According to [22, 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). For details, see Appendix 1.7.3. □

We can deduce from Proposition 2 that the optimal waiting time is a fixed function of the state $\mathbf{s}$. One possible way to solve this SMDP is by using the relative value iteration (RVI) algorithm. Next, we present this algorithm and reveal a useful structure of the optimal sampler that helps in simplifying this algorithm.

Algorithm 1: RVI algorithm with reduced complexity.

given $l = 0$, sufficiently large u, tolerance $\epsilon_1 > 0$, tolerance $\epsilon_2 > 0$;
while $u - l > \epsilon_1$ **do**
  $\beta = \frac{l+u}{2}$;
  $J(\mathbf{s}) = 0$, $h(\mathbf{s}) = 0$, $h_{\text{last}}(\mathbf{s}) = 0$ for all states $\mathbf{s} \in \mathcal{S}$;
  while $\max_{\mathbf{s}\in\mathcal{S}} |h(\mathbf{s}) - h_{last}(\mathbf{s})| > \epsilon_2$ **do**
    for ***each*** $\mathbf{s} \in \mathcal{S}$ **do**
      if $\mathbb{E}_Y \left[\sum_{l=1}^{m} g(a_{[l]} + Y)\right] \geq \beta$ **then**
        $z_s^\star = 0$;
      else
        $z_s^\star = \text{argmin}_{z\in\mathcal{Z}} C(\mathbf{s}, z) + \sum_{\mathbf{s}'\in\mathcal{S}} \mathbb{P}_{\mathbf{s}\mathbf{s}'}(z) h(\mathbf{s}')$;
      end
      $J(\mathbf{s}) = C(\mathbf{s}, z_s^\star) + \sum_{\mathbf{s}'\in\mathcal{S}} \mathbb{P}_{\mathbf{s}\mathbf{s}'}(z_s^\star) h(\mathbf{s}')$;
    end
    $h_{\text{last}}(\mathbf{s}) = h(\mathbf{s})$;
    $h(\mathbf{s}) = J(\mathbf{s}) - J(\mathbf{o})$;
  end
  if $J(\mathbf{o}) \geq 0$ **then**
    $u = \beta$;
  else
    $l = \beta$;
  end
end

Simplified Relative Value Iteration Using Optimal Sampler Structure:
The RVI algorithm [23, Section 9.5.3], [24, Page 171] can be used to solve Bellman's equation (1.23). Starting with an arbitrary state $\mathbf{o}$, a single iteration for the RVI algorithm is given as follows:

$$\begin{aligned} Q_{n+1}(\mathbf{s}, z) &= C(\mathbf{s}, z) + \sum_{\mathbf{s}'\in\mathcal{S}} \mathbb{P}_{\mathbf{s}\mathbf{s}'}(z) h_n(\mathbf{s}'), \\ J_{n+1}(\mathbf{s}) &= \min_{z\in\mathcal{Z}} (Q_{n+1}(\mathbf{s}, z)), \\ h_{n+1}(\mathbf{s}) &= J_{n+1}(\mathbf{s}) - J_{n+1}(\mathbf{o}), \end{aligned} \tag{1.27}$$

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 we repeat the iteration of the RVI algorithm as described before[7].

The complexity of the RVI algorithm is high due to many sources (i.e., the curse of dimensionality [25]). 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 3. *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* $z_s^\star = 0$.

Proof See Appendix 1.7.4. □

We can exploit the threshold test in Proposition 3 to reduce the complexity of the RVI algorithm as follows: The optimal waiting time for any state $\mathbf{s}$ that satisfies $\mathbb{E}_Y \left[\sum_{l=1}^{m} g(a_{[l]} + Y)\right] \geq \bar{\Delta}_{\text{avg-opt}}$ is zero. Thus, we need to solve (1.27) 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 algorithm in return. 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 3 and the RVI algorithm,

[7] According to [23, 24], 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 [23, Section 8.5.4], which is a simple transformation. However, This is not always necessary to be done.

we propose the "RVI with reduced complexity (RVI-RC) sampler" in Algorithm 1. According to [23, 24], $J(\mathbf{o})$ in Algorithm 1 converges to the optimal average cost per stage. Moreover, in the outer layer of Algorithm 1, bisection is employed to obtain $\bar{\Delta}_{\text{avg-opt}}$, where β converges to $\bar{\Delta}_{\text{avg-opt}}$. Finally, 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.

1.4.3 Jointly Optimal Scheduler and Sampler

Thanks to the separation principle, we are able to design the scheduler and sampler separately, which are presented in Sections 1.4.1 and 1.4.2, respectively. We here conclude our presented results so far, where an optimal solution for Problem (1.6) is manifested in the following theorem:

THEOREM 1.6. *The MAF scheduler and the RVI-RC sampler form an optimal solution for Problem* (1.6).

Proof The theorem follows directly from Proposition 1, Proposition 2, and Proposition 3. □

1.5 Final Remarks and Open Questions

1.5.1 Low-complexity Sampler Design via Bellman's Equation Approximation

In this section, we devise a low-complexity sampler via an approximate analysis for Bellman's equation in (1.23) whose solution is the RVI-RC sampler. During our discussion here and in what follows, we fix the scheduler to the MAF. The obtained low-complexity sampler in this section will be shown to have near optimal age performance in our numerical results in Section 1.5.3. 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 (1.19) depends only on the distribution of the service time Y which is independent of the system state and the control action. Hence, the second term in Bellman's equation in (1.23) can be rewritten as

$$\sum_{\mathbf{s}'\in\mathcal{S}} \mathbb{P}_{\mathbf{s}\mathbf{s}'}(z)h(\mathbf{s}'(z,y)) = \sum_{y\in\mathcal{Y}} \mathbb{P}(Y=y)h(\mathbf{s}'(z,y)). \tag{1.28}$$

As a result, Bellman's equation in (1.23) can be rewritten as

$$\lambda = \min_{z} \left(C(\mathbf{s}, z) + \sum_{y\in\mathcal{Y}} \mathbb{P}(Y=y)(h(\mathbf{s}'(z,y)) - h(\mathbf{s})) \right). \tag{1.29}$$

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 [26] and [27]. Let $\mathbf{s} = (a_{[1]}, \ldots, a_{[m]})$, then using the first order Taylor approximation around a state $\mathbf{v} = (a^v_{[1]}, \ldots, a^v_{[m]})$ (some fixed state), we get

$$h(\mathbf{s}) \approx h(\mathbf{v}) + \sum_{l=1}^{m} (a_{[l]} - a^v_{[l]}) \frac{\partial h(\mathbf{v})}{\partial a_{[l]}}. \tag{1.30}$$

Again, we use the first order Taylor approximation around the state $\mathbf{v}$, together with the state evolution in (1.17), to get

$$h(\mathbf{s}'(z,y)) \approx h(\mathbf{v}) + (y - a^v_{[m]}) \frac{\partial h(\mathbf{v})}{\partial a_{[m]}} + \sum_{l=1}^{m-1} (a_{[l+1]} - a^v_{[l]} + z + y) \frac{\partial h(\mathbf{v})}{\partial a_{[l]}}. \tag{1.31}$$

From (1.30) and (1.31), we get

$$h(\mathbf{s}'(z,y)) - h(\mathbf{s}) \approx (y - a_{[m]}) \frac{\partial h(\mathbf{v})}{\partial a_{[m]}} + \sum_{l=1}^{m-1} (a_{[l+1]} - a_{[l]} + z + y) \frac{\partial h(\mathbf{v})}{\partial a_{[l]}}. \tag{1.32}$$

This implies that

$$\begin{aligned} \sum_{y\in\mathcal{Y}} \mathbb{P}(Y=y)(h(\mathbf{s}'(z,y)) - h(\mathbf{s})) \approx & (\mathbb{E}[Y] - a_{[m]}) \frac{\partial h(\mathbf{v})}{\partial a_{[m]}} \\ & + \sum_{l=1}^{m-1} (a_{[l+1]} - a_{[l]} + z + \mathbb{E}[Y]) \frac{\partial h(\mathbf{v})}{\partial a_{[l]}}. \end{aligned} \tag{1.33}$$

Using (1.29) with (1.33), we can get the following approximated Bellman's equation:

$$\lambda \approx \min_{z} \left(C(\mathbf{s}, z) + (\mathbb{E}[Y] - a_{[m]}) \frac{\partial h(\mathbf{v})}{\partial a_{[m]}} + \sum_{l=1}^{m-1} (a_{[l+1]} - a_{[l]} + z + \mathbb{E}[Y]) \frac{\partial h(\mathbf{v})}{\partial a_{[l]}} \right). \tag{1.34}$$

By following the same steps as in Appendix 1.7.4 to get the optimal z that minimizes the objective function in (1.34), we get the following condition: The optimal z, for a given state $\mathbf{s}$, must satisfy

$$\mathbb{E}_Y \left[\sum_{l=1}^{m} g(a_{[l]} + t + Y) \right] - \bar{\Delta}_{\text{avg-opt}} + \sum_{l=1}^{m-1} \frac{\partial h(\mathbf{v})}{\partial a_{[l]}} \geq 0 \tag{1.35}$$

for all $t > z$, and

$$\mathbb{E}_Y \left[\sum_{l=1}^{m} g(a_{[l]} + t + Y) \right] - \bar{\Delta}_{\text{avg-opt}} + \sum_{l=1}^{m-1} \frac{\partial h(\mathbf{v})}{\partial a_{[l]}} \leq 0 \tag{1.36}$$

for all $t < z$. The smallest z that satisfies (1.35)-(1.36) is

$$\hat{z}_s^{\star} = \inf \left\{ t \geq 0 : \mathbb{E}_Y \left[\sum_{l=1}^{m} g(a_{[l]} + t + Y) \right] \geq \bar{\Delta}_{\text{avg-opt}} - \sum_{l=1}^{m-1} \frac{\partial h(\mathbf{v})}{\partial a_{[l]}} \right\}, \tag{1.37}$$

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, (1.37) can be rewritten as

$$\hat{z}_s^{\star} = \inf \left\{ t \geq 0 : \mathbb{E}_Y \left[\sum_{l=1}^{m} g(a_{[l]} + t + Y) \right] \geq T \right\}. \tag{1.38}$$

This simple threshold sampler can approximate the optimal sampler for the original Bellman's equation in (1.23). The optimal threshold (T) in (1.38) can be obtained using a golden-section method [28]. Moreover, for a given state $\mathbf{s}$ and the threshold (T), (1.38) can be solved using the bisection method or any other one-dimensional search method. As we mentioned before, we will show in our numerical results in Section 1.5.3 that, when the scheduling policy is fixed to the MAF scheduler, the performance of this approximated sampler is almost the same as that of the RVI-RC sampler.

1.5.2 Linear Age-penalty Function and Useful Insights

In this section, we consider the special case when the age-penalty function is linear, i.e., $g(x) = b_1 x + b_2$ for $b_1 > 0$. Such a simplification in the penalty function will further simplify the obtained samplers.

Simplification in the Optimal Sampler

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 3 is simplified as follows:

PROPOSITION 4. *If the state* $\mathbf{s} = (a_{[1]}, \ldots, a_{[m]})$ *satisfies* $A_s \geq \frac{\bar{\Delta}_{avg\text{-}opt} - m b_1 \mathbb{E}[Y] - m b_2}{b_1}$*, then we have* $z_s^{\star} = 0$.

Proof The proposition follows directly by substituting $g(x) = b_1 x + b_2$ into the threshold test in Proposition 3. □

Hence, the change in Algorithm 1 is to replace the threshold test in Step 7 by $A_s \geq (\bar{\Delta}_{\text{avg-opt}} - m b_1 \mathbb{E}[Y] - m b_2)/b_1$. This threshold test is easier to check than that in Proposition 3. This further simplifies the RVI-RC algorithm.

Optimality of the Zero-wait Sampler

It is obvious that the zero-wait sampler is throughput and delay optimal. However, as explained in Section 1.2.3 with a counter example, the zero-wait sampler does not necessarily minimize the average age. For a special case when $g(x) = b_1 x + b_2$, we provide a sufficient condition for the optimality of the zero-wait sampler for minimizing the Ta-AP. 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 4, the sufficient condition for the optimality of the zero-wait sampler is manifested in the following theorem:

THEOREM 1.7. *If*

$$y_{inf} \geq \frac{(m-1)\mathbb{E}[Y]^2 + \mathbb{E}[Y^2]}{(m+1)\mathbb{E}[Y]}, \tag{1.39}$$

then the zero-wait sampler is optimal for Problem (1.15).

Proof See Appendix 1.7.5 □

From Theorem 1.7, it immediately follows that:

COROLLARY. *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* (1.15).

Proof The corollary follows directly from Theorem 1.7 by showing that (1.39) always holds in this case. □

Corollary 1.5.2 suggests that designing the optimal schedulers could be enough if the transmission times are deterministic, where a source generates a sample whenever this source is scheduled for transmission. 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. In practice, such random transmission times occur in many applications, such as autonomous vehicles. In particular, as we mentioned before, there are many ECUs in a vehicle, that are connected to one or more sensors and actuators via a CAN bus [10, 11]. 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 samples usually have hard deadlines, the transmissions of low priority samples are interrupted whenever the higher priority ones are transmitted. Therefore, update samples with lower priority see a time-varying bandwidth, and hence encounter a random transmission time.

Low-complexity Water-filling Sampler

We here investigate how the designed low-complexity sampler in Section 1.5.1 can be further simplified when $g(x) = b_1 x + b_2$. In particular, we will show that the threshold sampler in Section 1.5.1 reduces to the water-filling sampler in this case.

Substituting $g(x) = b_1 x + b_2$ into (1.37), 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

$$b_1 A_s - \bar{\Delta}_{\text{avg-opt}} + mb_1 z + mb_1 \mathbb{E}[Y] + mb_2 + \sum_{l=1}^{m-1} \frac{\partial h(\mathbf{v})}{\partial a_{[l]}} = 0, \tag{1.40}$$

where A_s is the sum of the age values of state $\mathbf{s}$. Rearranging (1.40), we get

$$\hat{z}_s^* = \left[\frac{\bar{\Delta}_{\text{avg-opt}} - mb_1 \mathbb{E}[Y] - \sum_{l=1}^{m-1} \frac{\partial h(\mathbf{v})}{\partial a_{[l]}} - mb_2}{mb_1} - \frac{A_s}{m} \right]^+. \tag{1.41}$$

By observing that the term $\sum_{i=1}^{m-1} \frac{\partial h(\mathbf{v})}{\partial a_{[i]}}$ is constant, (1.41) can be rewritten as

$$\hat{z}_s^* = \left[T - \frac{A_s}{m} \right]^+, \tag{1.42}$$

The solution in (1.42) 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 (1.42) suggests that this simple water-filling sampler can approximate the optimal solution of the original Bellman's equation in (1.23) when $g(x) = b_1 x + b_2$ for some $b_1 > 0$ and b_2. Similar to the general case, the optimal threshold (T) in (1.42) can be obtained using a golden-section method. We will also evaluate the performance of these water-filling samplers in the next section and show that its performance is almost the same as RVI-RC sampler.

1.5.3 Numerical Results

We consider an information update system with $m = 3$ sources. We use "RAND" to represent a random scheduler, where sources are chosen to be served with equal probability. By "constant-wait", we refer to

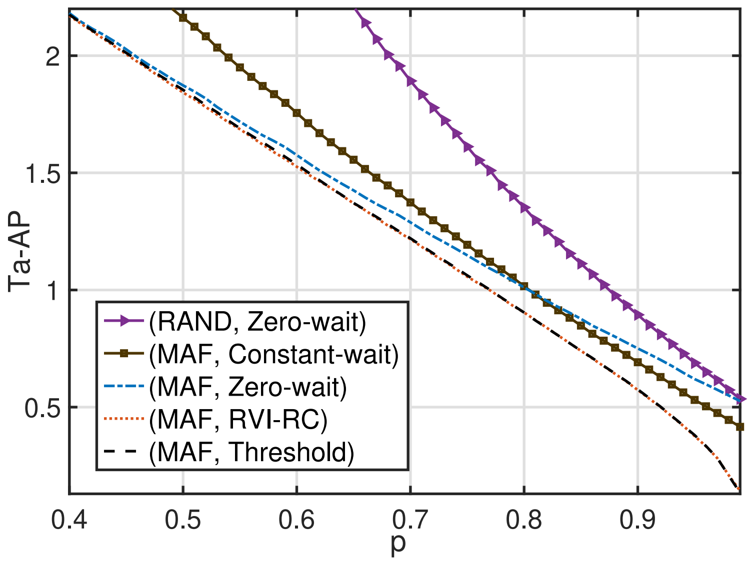


Figure 1.8 Ta-AP versus transmission probability p for an update system with $m = 3$ sources, where $g(x) = e^{0.1x} - 1$.

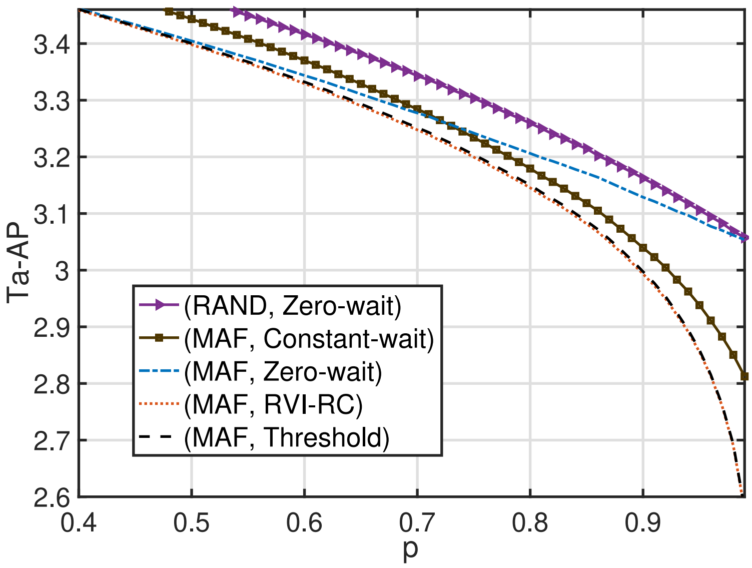


Figure 1.9 Ta-AP versus transmission probability p for an update system with $m = 3$ sources, where $g(x) = x^{0.1}$.

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" and "water-filling" to denote the proposed samplers in (1.38) and (1.42), respectively.

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. Figs. 1.8, 1.9, and 1.10 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. The range of the probability p is $[0.4; 0.99]$ in Figs. 1.8, 1.9, and 1.10. When $p = 1$, $\mathbb{E}[Y] = \mathbb{E}[Y^2] = 0$ and hence the Ta-AP of the zero-wait sampler (for any scheduler) at $p = 1$ is undefined. Therefore, the point $p = 1$ is not plotted in Figs. 1.8, 1.9, and 1.10. 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 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

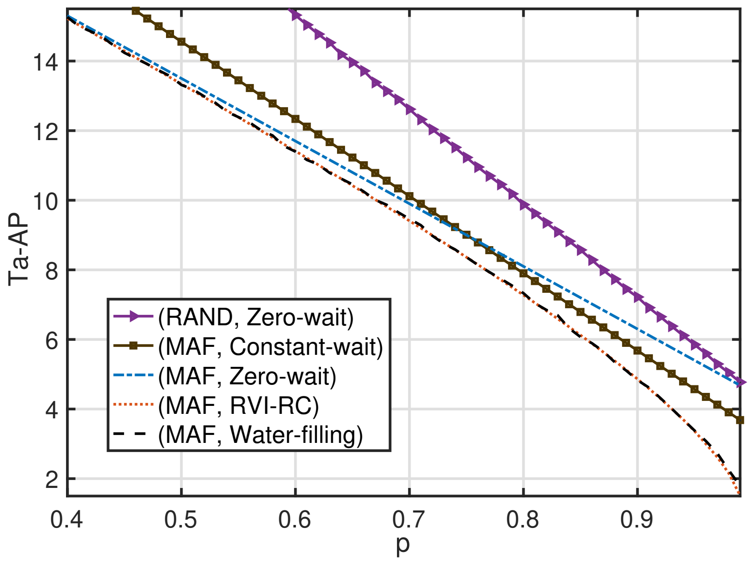


Figure 1.10 Ta-AP versus transmission probability p for an update system with $m = 3$ sources, where $g(x) = x$.

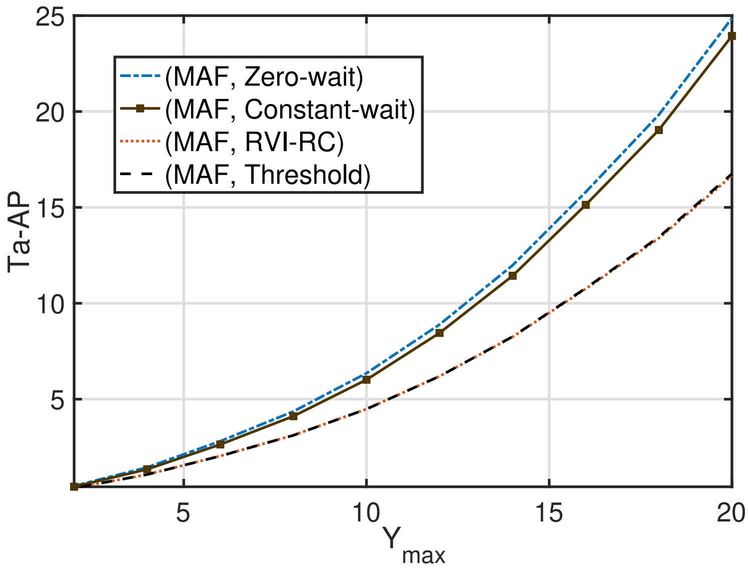


Figure 1.11 Ta-AP versus the maximum service time $Y_{\max}$ for an update system with $m = 3$ sources, where $g(x) = e^{0.1x} - 1$.

threshold sampler in Fig. 1.8 and 1.9, and the water-filling sampler in Fig. 1.10 almost coincides with 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 Figs. 1.11, 1.12, and 1.13, 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 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 Figs. 1.11 and 1.12, and the water-filling sampler in Fig. 1.13 almost coincides with that of the RVI-RC sampler.

1.5.4 Open Questions

So far, we have considered only homogeneous age-penalty functions, i.e., all sources have the same penalty on the age. This technical assumption allowed us to develop the separation principle. Using this principle,

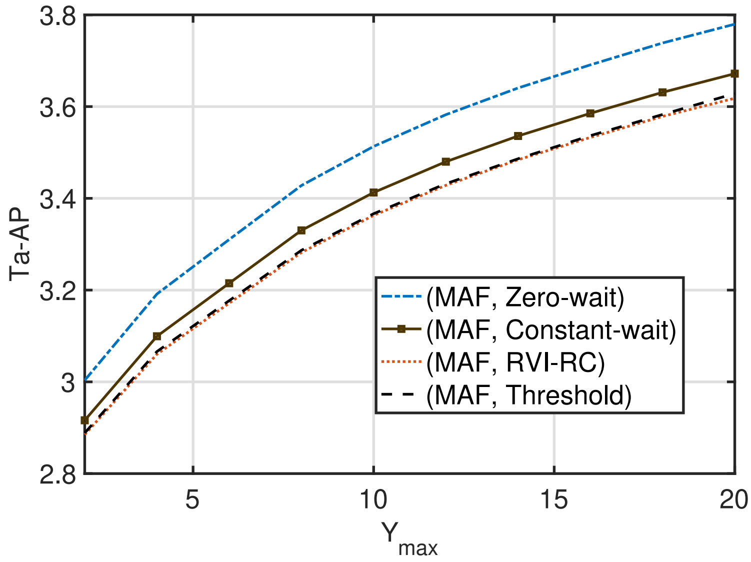


Figure 1.12 Ta-AP versus the maximum service time $Y_{\max}$ for an update system with $m = 3$ sources, where $g(x) = x^{0.1}$.

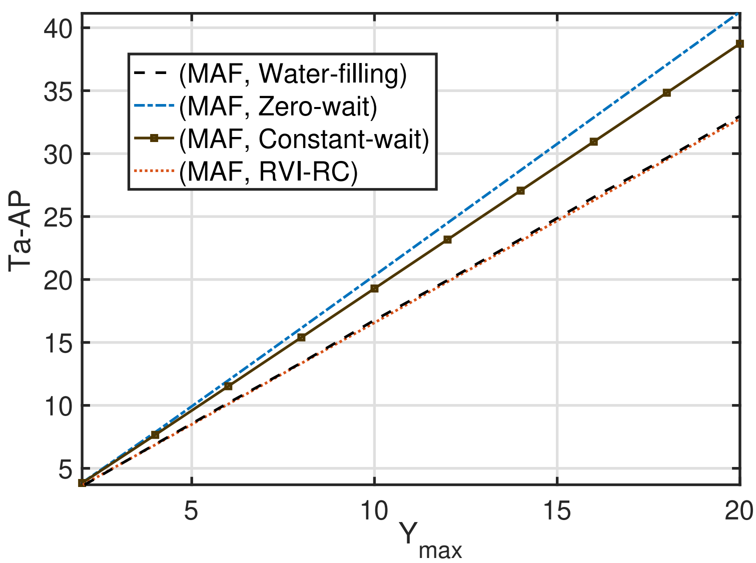


Figure 1.13 Ta-AP versus the maximum service time $Y_{\max}$ for an update system with $m = 3$ sources, where $g(x) = x$.

we were able to decouple the scheduler decisions and the sampler decisions, which allowed us to design each separately from the other. However, in practice, it is quite possible that different sources may incur different penalties on the age, e.g., for certain types of information may be quite sensitive to the age (e.g., in autonomous driving applications), while others may be far more age-tolerant (e.g., measuring the humidity in a region). Hence, for various practical problems, this condition may be violated, in which case the separation principle itself may not hold any longer. This is because of the variation in the sources' order, according to their age-penalty values, with time. This results in a strong correlation between the scheduler and sampler actions. Hence, our technique here may not be suitable for this case. Indeed, other tools and techniques, which could be more sophisticated, are needed to answer such a challenging question. We raised this question to the reader to convey that this research area is still at an early stage and expected to grow considerably in the next few years because various research challenges have to be addressed and solved. The aforementioned challenge is one of them that needs to be addressed.

1.6 Summary

In this chapter, we have investigated how controlling the sampling times can further improve the data freshness in information update systems. We started by providing a brief explanation of the sampling problem in single-source networks, supported by an example which shows that the optimal sampler is not trivial and has a counter intuitive phenomenon. Moreover, we provided real-time applications on age of information and its penalty functions.

Later on, we shifted our focus to study the optimal sampling problem in multi-source networks. It turned out that the optimal sampling problem in multi-source networks is more challenging. This is because the sources are communicating their generated samples to the destination via a shared channel, and hence the decision policy does not only control a sampler, but also a scheduler. Our target was to study the problem of finding the optimal decision policy that controls the sampling times, the sampler, and the transmission order of the sources, the scheduler, to minimize the Ta-AP in multi-source networks. We showed that the MAF scheduler and the RVI-RC sampler, which results from reducing the computation complexity of the RVI algorithm, are jointly optimal for minimizing the Ta-AP. In addition, we devised a low-complexity threshold sampler via an approximate analysis of Bellman's equation. Finally, we considered a special case when the age-penalty function is linear and obtained a sufficient condition for the optimality of the zero-wait sampler in this case. We also showed that the approximated 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.

1.7 Appendix

1.7.1 Proof of Proposition 1

Let the vector $\mathbf{\Delta}_\pi(t) = (\Delta_{[1],\pi}(t), \ldots, \Delta_{[m],\pi}(t))$ 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 $\{\mathbf{\Delta}_\pi(t), t \geq 0\}$ denote the state process of the scheduler π. For notational simplicity, let P represent the MAF scheduler. Throughout the proof, we assume that $\mathbf{\Delta}_\pi(0^-) = \mathbf{\Delta}_P(0^-)$ 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 1.8. *Suppose that $\mathbf{\Delta}_\pi(0^-) = \mathbf{\Delta}_P(0^-)$ for all scheduler π and the sampler is fixed, then we have*

$$\{\mathbf{\Delta}_P(t), t \geq 0\} \leq_{st} \{\mathbf{\Delta}_\pi(t), t \geq 0\} \tag{1.43}$$

We use a coupling and forward induction to prove Lemma 1.8. For any scheduler π, suppose that the stochastic processes $\widetilde{\mathbf{\Delta}}_P(t)$ and $\widetilde{\mathbf{\Delta}}_\pi(t)$ have the same stochastic laws as $\mathbf{\Delta}_P(t)$ and $\mathbf{\Delta}_\pi(t)$. The state processes $\widetilde{\mathbf{\Delta}}_P(t)$ and $\widetilde{\mathbf{\Delta}}_\pi(t)$ are coupled such that the service times of the samples 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 sample generation and delivery times are the same under both schedulers. According to Theorem 6.B.30 of [15], if we can show

$$\mathbb{P}\left[\widetilde{\mathbf{\Delta}}_P(t) \leq \widetilde{\mathbf{\Delta}}_\pi(t), t \geq 0\right] = 1, \tag{1.44}$$

then (1.43) is proven. To ease the notational burden, we will omit the tildes on the coupled versions in this proof and just use $\mathbf{\Delta}_P(t)$ and $\mathbf{\Delta}_\pi(t)$. Next, we compare scheduler P and scheduler π on a sample path and prove (1.43) using the following lemma:

Lemma 1.9 (Inductive Comparison). *Suppose that a sample 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 $\mathbf{\Delta}_P$ before the sample delivery, which becomes $\mathbf{\Delta}'_P$ after the sample delivery. The system state of the scheduler π is $\mathbf{\Delta}_\pi$ before the sample delivery, which becomes $\mathbf{\Delta}'_\pi$ after the sample delivery. If*

$$\Delta_{[i],P} \leq \Delta_{[i],\pi}, i = 1, \ldots, m, \tag{1.45}$$

then

$$\Delta'_{[i],P} \le \Delta'_{[i],\pi}, i = 1, \dots, m. \tag{1.46}$$

Proof Since only one source can be scheduled at a time and the scheduler P is the MAF one, the sample 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,

$$\begin{aligned} &\Delta'_{[i],P} = \Delta_{[i+1],P}, i = 1, \dots, m-1, \\ &\Delta'_{[m],P} = t - S. \end{aligned} \tag{1.47}$$

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

$$\Delta'_{[i],\pi} \ge \Delta_{[i+1],\pi}, i = 1, \dots, m-1. \tag{1.48}$$

By combining (1.45), (1.47), and (1.48), we have

$$\Delta'_{[i],\pi} \ge \Delta_{[i+1],\pi} \ge \Delta_{[i+1],P} = \Delta'_{[i],P}, i = 1, \dots, m-1. \tag{1.49}$$

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

$$\Delta'_{[m],\pi} = t - S = \Delta'_{[m],P}. \tag{1.50}$$

By this, (1.46) is proven. □

Proof of Lemma 1.8 Using the coupling between the system state processes, and for any given sample path of the service times, we consider two cases:

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

Case 2: When a sample is delivered, the ages of the sources evolve according to Lemma 1.9.

By induction over time, we obtain

$$\Delta_{[i],P}(t) \le \Delta_{[i],\pi}(t), i = 1, \dots, m, t \ge 0. \tag{1.51}$$

Hence, (1.44) follows which implies (1.43) by Theorem 6.B.30 of [15]. This completes the proof. □

Proof of Proposition 1 Since the Ta-AP for any scheduling policy π is the expectation of non-decreasing functional of the process $\{\boldsymbol{\Delta}_\pi(t), t \ge 0\}$, (1.43) implies (1.7) using the properties of stochastic ordering [15]. This completes the proof. □

1.7.2 Proof of Lemma 1.5

Part (i) is proven in two steps:

Step 1: We will prove that $\bar{\Delta}_{\text{avg-opt}} \le \beta$ if and only if $\Theta(\beta) \le 0$. If $\bar{\Delta}_{\text{avg-opt}} \le \beta$, there exists a sampling policy $f = (Z_0, Z_1, \dots) \in \mathcal{F}$ that is feasible for (1.14) and (1.15), which satisfies

$$\limsup_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}\left[\sum_{l=1}^{m} \int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau\right]}{\sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}]} \le \beta. \tag{1.52}$$

Hence,

$$\limsup_{n\to\infty} \frac{\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\left[\sum_{l=1}^{m} \int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau - \beta(Z_i+Y_{i+1})\right]}{\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}]} \le 0. \tag{1.53}$$

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

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\left[\sum_{l=1}^{m} \int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau - \beta(Z_i+Y_{i+1})\right] \le 0. \tag{1.54}$$

Therefore, $\Theta(\beta) \le 0$.

In the reverse direction, if $\Theta(\beta) \le 0$, then there exists a sampling policy $f = (Z_0, Z_1, \dots) \in \mathcal{F}$ that is feasible for (1.14) and (1.15), which satisfies (1.54). 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 (1.54) by $\liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}[Z_i+Y_{i+1}]$ to get (1.53), which implies (1.52). 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 $\Theta(\beta) \leq 0$.

Step 2: We need to prove that $\bar{\Delta}_{\text{avg-opt}} < \beta$ if and only if $\Theta(\beta) < 0$. 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 $\Theta(\beta) > 0$. This completes part (i).

Part(ii): We first show that each optimal solution to (1.14) is an optimal solution to (1.15). By the claim of part (i), $\Theta(\beta) = 0$ 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 (1.14). Then, $\Delta_{\text{avg}(\pi_{\text{MAF}},f)} = \bar{\Delta}_{\text{avg-opt}} = \beta$. Applying this in the arguments of (1.52)-(1.54), we can show that policy f satisfies

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\left[\sum_{l=1}^{m}\int_{a_{li}}^{a_{li}+Z_i+Y_{i+1}} g(\tau)d\tau - \beta(Z_i+Y_{i+1})\right] = 0. \tag{1.55}$$

This and $\Theta(\beta) = 0$ imply that policy f is an optimal solution to (1.15).

Similarly, we can prove that each optimal solution to (1.15) is an optimal solution to (1.14). By this, part (ii) is proven. □

1.7.3 Proof of Proposition 2

According to [22, 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

$$\mathbb{P}\left[\mathbf{s}(k) = \mathbf{s}' | \mathbf{s}(0) = \mathbf{s}, f\right] > 0. \tag{1.56}$$

From the state evolution equation (1.17), we can observe that any state in $\mathcal{S}$ can be represented in terms of the waiting and service times. This implies (1.56). To clarify this, let us consider a system with 3 sources. Assume that the elements of state $\mathbf{s}'$ are as follows:

$$\begin{aligned} a'_{[1]} &= y_3 + z_2 + y_2 + z_1 + y_1, \\ a'_{[2]} &= y_3 + z_2 + y_2, \\ a'_{[3]} &= y_3, \end{aligned} \tag{1.57}$$

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 (1.17), we have in the first stage

$$\begin{aligned} a_{[1]1} &= a_{[2]} + z_1 + y_1, \\ a_{[2]1} &= a_{[3]} + z_1 + y_1, \\ a_{[3]1} &= y_1, \end{aligned} \tag{1.58}$$

and in the second stage, we have

$$\begin{aligned} a_{[1]2} &= a_{[3]} + z_1 + y_2 + z_1 + y_1, \\ a_{[2]2} &= y_2 + z_1 + y_1, \\ a_{[3]2} &= y_2, \end{aligned} \tag{1.59}$$

and in the third stage, we have

$$\begin{aligned} a_{[1]3} &= y_3 + z_2 + y_2 + z_1 + y_1 = a'_{[1]}, \\ a_{[2]3} &= y_3 + z_2 + y_2 = a'_{[2]}, \\ a_{[3]3} &= y_3 = a'_{[3]}. \end{aligned} \tag{1.60}$$

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

$$\mathbb{P}\left[\mathbf{s}(3) = \mathbf{s}' | \mathbf{s}(0) = \mathbf{s}, f\right] = \prod_{i=1}^{3}\mathbb{P}(Y_i = y_i) > 0, \tag{1.61}$$

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 (1.56) trivially holds for $m = 1$, and the previous argument can be used to show that (1.56) holds for any general m. This completes the proof. □

1.7.4 Proof of Proposition 3

We prove Proposition 3 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 (1.26), is non-decreasing in $\mathbf{s}$, which together with (1.25) 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

$$J_\alpha(\mathbf{s}(0); f) = \limsup_{n\to\infty} \mathbb{E}\left[\sum_{i=0}^{n-1} \alpha^i C(\mathbf{s}(i), Z_i)\right], \tag{1.62}$$

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

$$J_\alpha(\mathbf{s}) = \min_{f\in\mathcal{F}} J_\alpha(\mathbf{s}; f), \ \ \mathbf{s} \in \mathcal{S}. \tag{1.63}$$

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.

PROPOSITION 5. *The optimal total expected α-discounted cost $J_\alpha(\mathbf{s})$ satisfies*

$$J_\alpha(\mathbf{s}) = \min_{z\in\mathcal{Z}} C(\mathbf{s}, z) + \alpha \sum_{\mathbf{s}'\in\mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z) J_\alpha(\mathbf{s}'). \tag{1.64}$$

Moreover, a stationary deterministic policy that attains the minimum in equation (1.64) 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 $n \geq 0$,

$$J_{\alpha,n+1}(\mathbf{s}) = \min_{z\in\mathcal{Z}} C(\mathbf{s}, z) + \alpha \sum_{\mathbf{s}'\in\mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z) J_{\alpha,n}(\mathbf{s}'). \tag{1.65}$$

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 [22, Proposition 1.2.2 and Proposition 1.2.3], and [29]. □

Next, we use the optimality equation (1.64) and the value iteration in (1.65) to prove that $J_\alpha(\mathbf{s})$ is non-decreasing in $\mathbf{s}$.

LEMMA 1.10. *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 (1.65) to prove Lemma 1.10. 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

$$C(\mathbf{s}_1, z) \leq C(\mathbf{s}_2, z). \tag{1.66}$$

From the state evolution equation (1.17) and the transition probability equation (1.19), the second term of the right-hand side (RHS) of (1.65) can be rewritten as

$$\sum_{\mathbf{s}'\in\mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z) J_{\alpha,n}(\mathbf{s}') = \sum_{y\in\mathcal{Y}} \mathbb{P}(Y = y) J_{\alpha,n}(\mathbf{s}'(z, y)), \tag{1.67}$$

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

[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.

evolution equation (1.17), 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 $\mathbf{s}_1'(z,y) \le \mathbf{s}_2'(z,y)$. This implies that

$$\sum_{y\in\mathcal{Y}} \mathbb{P}(Y=y) J_{\alpha,n}(\mathbf{s}_1'(z,y)) \le \sum_{y\in\mathcal{Y}} \mathbb{P}(Y=y) J_{\alpha,n}(\mathbf{s}_2'(z,y)), \tag{1.68}$$

where we have used the induction assumption that $J_{\alpha,n}(\mathbf{s})$ is non-decreasing in $\mathbf{s}$. Using (1.66), (1.68), and the fact that the minimum operator in (1.65) retains the non-decreasing property, we conclude that

$$J_{\alpha,n+1}(\mathbf{s}_1) \le J_{\alpha,n+1}(\mathbf{s}_2). \tag{1.69}$$

This completes the proof. □

Step 2: We use Step 1 to prove Proposition 3. 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 (1.23) ($\sum_{\mathbf{s}'\in\mathcal{S}} \mathbb{P}_{\mathbf{s}\mathbf{s}'}(z) h(\mathbf{s}')$) is non-decreasing in $\mathbf{s}'$. Moreover, from the state evolution (1.17), 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 (1.23) ($\sum_{\mathbf{s}'\in\mathcal{S}} \mathbb{P}_{\mathbf{s}\mathbf{s}'}(z) h(\mathbf{s}')$) 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 (1.23). 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

$$\delta q(z;\omega) \triangleq \lim_{\epsilon\to 0^+} \frac{q(z+\epsilon\omega)-q(z)}{\epsilon}. \tag{1.70}$$

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 [1, 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 [30, 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 $\theta>0$ [31, Proposition 1.1.2(i)]. Using the monotone convergence theorem [32, 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 [30, p.710] and the convexity of $q(z)$, z is optimal to the cost function $C(\mathbf{s},z)$ if and only if

$$\delta q(z;\omega) - \bar{\Delta}_{\text{avg-opt}}\omega \ge 0,\ \forall\omega\in\mathbb{R}. \tag{1.71}$$

As ω in (1.71) is an arbitrary real number, considering $\omega=1$, (1.71) becomes

$$\lim_{t\to z^+}\mathbb{E}_Y\left[\sum_{l=1}^{m} g(a_{[l]}+t+Y)\right] - \bar{\Delta}_{\text{avg-opt}} \ge 0. \tag{1.72}$$

Likewise, considering $\omega=-1$, (1.71) implies

$$\lim_{t\to z^-}\mathbb{E}_Y\left[\sum_{l=1}^{m} g(a_{[l]}+t+Y)\right] - \bar{\Delta}_{\text{avg-opt}} \le 0. \tag{1.73}$$

Since $g(\cdot)$ is non-decreasing, we get from (1.71)-(1.73) that z must satisfy

$$\mathbb{E}_Y\left[\sum_{l=1}^{m} g(a_{[l]}+t+Y)\right]-\bar{\Delta}_{\text{avg-opt}}\geq 0,\ \text{if } t>z, \tag{1.74}$$

$$\mathbb{E}_Y\left[\sum_{l=1}^{m} g(a_{[l]}+t+Y)\right]-\bar{\Delta}_{\text{avg-opt}}\leq 0,\ \text{if } t<z. \tag{1.75}$$

Subsequently, the smallest z that satisfies (1.74)-(1.75) is

$$z=\inf\left\{t\geq 0:\mathbb{E}_Y\left[\sum_{l=1}^{m} g(a_{[l]}+t+Y)\right]\geq\bar{\Delta}_{\text{avg-opt}}\right\}. \tag{1.76}$$

According to (1.76), 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. □

1.7.5 Proof of Theorem 1.7

We use the threshold test $A_s\geq(\bar{\Delta}_{\text{avg-opt}}-mb_1\mathbb{E}[Y]-mb_2)/b_1$, in Proposition 4, to prove Theorem 1.7. We will show that the condition in (1.39) implies that $A_s\geq(\bar{\Delta}_{\text{avg-opt}}-mb_1\mathbb{E}[Y]-mb_2)/b_1$ holds for all states $\mathbf{s}\in\mathcal{S}$, and hence the zero-wait sampler is optimal under this condition. From the state evolution (1.17), we can deduce that for any state $\mathbf{s}\in\mathcal{S}$, we have

$$a_{[l]}\geq(m-l+1)y_{\text{inf}},\ \forall l=1,\ldots,m. \tag{1.77}$$

This implies

$$A_s\geq\sum_{l=1}^{m} l y_{\text{inf}}=\frac{m(m+1)}{2}y_{\text{inf}},\ \forall\mathbf{s}\in\mathcal{S}. \tag{1.78}$$

Moreover, it is easy to show that the Ta-AP of the zero-wait sampler, when the scheduling policy is fixed to the MAF scheduler and $g(x)=b_1x+b_s$, is given by

$$\bar{\Delta}_0=\frac{\frac{m(m+1)}{2}b_1\mathbb{E}[Y]^2+\frac{m}{2}b_1\mathbb{E}[Y^2]+mb_2\mathbb{E}[Y]}{\mathbb{E}[Y]}. \tag{1.79}$$

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

$$\frac{\bar{\Delta}_0-mb_1\mathbb{E}[Y]-mb_2}{b_1}\geq\frac{\bar{\Delta}_{\text{avg-opt}}-mb_1\mathbb{E}[Y]-mb_2}{b_1}. \tag{1.80}$$

Hence, if the following condition holds

$$\frac{m(m+1)}{2}y_{\text{inf}}\geq\frac{\frac{m(m+1)}{2}b_1\mathbb{E}[Y]^2+\frac{m}{2}b_1\mathbb{E}[Y^2]+mb_2\mathbb{E}[Y]}{b_1\mathbb{E}[Y]}-m\mathbb{E}[Y]-\frac{mb_2}{b_1}, \tag{1.81}$$

which is equivalent to

$$y_{\text{inf}}\geq\frac{(m-1)\mathbb{E}[Y]^2+\mathbb{E}[Y^2]}{(m+1)\mathbb{E}[Y]}, \tag{1.82}$$

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. □

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," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7492–7508, Nov 2017.

[2]B. T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu, "Age of information under energy replenishment constraints," in *Proc. ITA*, Feb. 2015.

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

[4]B. Adelberg, H. Garcia-Molina, and B. Kao, "Applying update streams in a soft real-time database system," in *ACM SIGMOD Record*, vol. 24, no. 2, 1995, pp. 245–256.

[5]J. Cho and H. Garcia-Molina, "Synchronizing a database to improve freshness," in *ACM SIGMOD Record*, vol. 29, no. 2, 2000, pp. 117–128.

[6]L. Golab, T. Johnson, and V. Shkapenyuk, "Scheduling updates in a real-time stream warehouse," in *Proc. IEEE 25th Int'l Conf. Data Eng. (ICDE)*, March 2009, pp. 1207–1210.

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

[8]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.

[9]A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "Age and value of information: Non-linear age case," in *ISIT*. IEEE, 2017, pp. 326–330.

[10]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.

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

[12]A. sensors and electronics expo 2017. [Online]. Available: http://www.automotivesensors2017.com/

[13]A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, "Age-optimal sampling and transmission scheduling in multi-source systems," in *Proc. MobiHoc*, ser. Mobihoc '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 121–130. [Online]. Available: https://doi.org/10.1145/3323679.3326510

[14]A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, "Optimal sampling and scheduling for timely status updates in multi-source networks," *IEEE Trans. Inf. Theory*, pp. 1–1, 2021.

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

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

[17]Y. Hsu, E. Modiano, and L. Duan, "Scheduling algorithms for minimizing age of information in wireless broadcast networks with random arrivals," *arXiv preprint arXiv:1712.07419*, 2017.

[18]I. Kadota, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Minimizing the age of information in broadcast wireless networks," in *Communication, Control, and Computing (Allerton), 2016 54th Annual Allerton Conference on*. IEEE, 2016, pp. 844–851.

[19]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.

[20]Y. Sun, E. Uysal-Biyikoglu, and S. Kompella, "Age-optimal updates of multiple information flows," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2018, pp. 136–141.

[21]Y. Sun, Y. Polyanskiy, and E. Uysal, "Sampling of the wiener process for remote estimation over a channel with random delay," *IEEE Transactions on Information Theory*, vol. 66, no. 2, pp. 1118–1135, 2019.

[22]D. P. Bertsekas, *Dynamic Programming and Optimal Control, 2nd ed.* Belmont, MA: Athena Scientific, 2001, vol. 2.

[23]M. L. Puterman, "Markov decision processes: Discrete stochastic dynamic programming (wiley series in probability and statistics)," 2005.

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

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

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

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

[28]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.

[29]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.

[30]D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.

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

[32]R. Durrett, *Probability: theory and examples.* Cambridge university press, 2010.