Optimal Sampling for Data Freshness: Unreliable Transmissions with Random Two-way Delay

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

***Abstract*—In this paper, we aim to design an optimal sampler for a system in which fresh samples of a signal (source) are sent through an unreliable channel to a remote estimator, and acknowledgments are sent back over a feedback channel. Both the forward and feedback channels could have random transmission times due to time varying channel conditions. Motivated by distributed sensing, the estimator can estimate the real-time value of the source signal by combining the signal samples received through the channel and the noisy signal observations collected from a local sensor. We prove that the estimation error is a non-decreasing function of the Age of Information (AoI) for the received signal samples and design an optimal sampling strategy that minimizes the long-term average estimation error subject to a sampling rate constraint. The sampling strategy is also optimal for minimizing the long-term average of general non-decreasing functions of the AoI. The optimal sampler design follows a randomized threshold strategy: If the last transmission was successful, the source waits until the expected estimation error upon delivery exceeds a threshold and then sends out a new sample. If the last transmission fails, the source immediately sends out a new sample without waiting. The threshold is the root of a fixed-point equation and can be solved with low complexity (e.g., by bisection search). The optimal sampling strategy holds for general transmission time distributions of the forward and feedback channels. Numerical simulations are provided to compare different sampling policies.**

***Index Terms*—Age of information, unreliable transmissions, two-way delay, and sampling.**

I. Introduction

Timely updates are crucial in many applications such as vehicular networks, wireless sensor networks, and UAV navigations. To achieve timely updates, we require the destination to receive fresh information from the remote source as quickly as possible. The information freshness is measured by age of information, or simply age, which has been widely explored in recent years (e.g., [2]–[23]). Age of information with the function of current time t is defined as $\Delta_t = t - U_t$, where U_t is the generation time of the freshest information data. In several different queueing systems, the Last-Generated, First-Served (LGFS) policy is shown to achieve age-optimality [2]–[4]. Scheduling policies in various wireless networks are studied to minimize age [5]–[9]. A literature review of recent works in age of information is provided in [10].

In [11] and [12], a connection between age of information and remote estimation of time-varying processes (e.g., Wiener process or Ornstein-Uhlenbeck (OU) process) was established. One of the remote estimation objectives in these early studies was to design an optimal sampling policy to minimize the long-term average minimum mean square error (MMSE). The MMSE is a function of the age if the sampling policy is independent of the signal being sampled [11]–[14]. Among these studies, the estimator obtains the exact signal samples subject to delay. However, the estimator neglects the instant and inexact signal samples. For example, in vehicular networks, the estimator can estimate a signal via both the exact signal samples from the remote sensor and the instant camera streaming from the close vehicle sensor over time. To consider both the delayed and instant signal samples, we will apply the Kalman Filter [24, Chapter 7] and study the relationship between the MMSE and age of information.

The desire for timely updates and the study of the new remote estimation problem necessitates considering general nonlinear age functions in the development of optimal sampling policies. To reduce the age, we may require the source to wait before submitting a new sample, i.e., the zero-wait policy may not be age-optimal[1] [15]. The study in [16] generalized the result in [15], proposed an optimal sampling policy under a Markov channel with sampling rate constraint, and observed that the zero-wait policy is far from optimal if, for example, the transmission times are heavy-tail distributed or positively correlated. In [17], the authors provided a survey of the age penalty functions related to autocorrelation, remote estimation, and mutual information. The optimal sampling solution is a deterministic or randomized threshold policy based on the objective value and the sampling rate constraint. However, in real-time network systems, both the forward direction and the feedback direction have a random delay. Such a random two-way delay model was considered in e.g., [18], [19]. In [18], the paper proposed a low complexity algorithm with a quadratic convergence rate to compute the optimal threshold. In [19], an optimal joint cost-and-AoI minimization solution was provided for multiple coexisting source-destination pairs with heterogeneous AoI penalty functions. Although the above studies have developed optimal sampling strategies, they assume that the transmission process is reliable. However, due to the channel

This paper was presented in part at IEEE INFOCOM 2022 [1].

This work has been supported in part by NSF grants: 2112471, CNS-2106932, CNS- 2106933, CNS-1955535, CNS-1901057, and CCF-1813050, and a grant from the Army Research Office: W911NF-21-1-0244.

J. Pan is with the Department of ECE, The Ohio State University, Columbus, OH 43210 USA (e-mail: pan.743@osu.edu).

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).

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]We refer to a policy as *zero-wait* if the source takes the sample and transmits as soon as it receives the acknowledgement.

fading, the channel conditions are time-varying, and thus the transmission process is unreliable.

Recent studies [20], [21] investigate sampling strategies while considering unreliable transmissions. In [20], the authors considered quantization errors, noisy channel, and non-zero receiver processing time, and they established the relationship between the MMSE and age. For general age functions, they provided optimal sampling policies, given that the sampler needs to wait before receiving feedback. When the sampler does not need to wait, they provided enhanced sampling policies that perform better than previous ones. In [21], the authors chose *idle* or *transmit* at each time slot to minimize joint age penalty and transmission cost. The optimality of a threshold-based policy is shown, and the policy's threshold is computed efficiently. Nevertheless, in practice, transmission delays are random rather than constant because of congestion, random sample sizes, etc, which is a critical challenge facing the design of sampling strategies.

To address the aforementioned challenges, we investigate how to design optimal sampling strategies in wireless networks under the following more realistic (and general) conditions that have largely been unexplored: unreliable transmissions and random delay in both forward and feedback directions. Early studies on optimizing sampling assuming reliable channels with random delays have shown that the sampling problem is decomposed into a per-sample problem. The per-sample problem can be further solved by optimization theory (e.g., [15]–[18]) or optimal stopping rules (e.g., [11]–[13]). Similarly, our problem assuming an unreliable channel is equivalent to a per-epoch problem containing multiple samples until successful packet delivery. Therefore, the per-epoch problem is a Markov Decision Process (MDP) with an uncountable state space, which is a key difference with past works, (e.g., [11]–[13], [15]–[18]) and faces the curse of dimensionality.[2] The main contributions of this paper are stated as follows:

- We first formulate the problem where the estimator estimates a signal in real-time by combining noisy signal observations from a local sensor and accurate signal samples received from a remote sensor. We show that if the sampling policy is made independently of the signal being sampled, the MMSE equals an increasing function of the age of the received signal samples.
- For general nonlinear age functions, or simply age penalty functions, we provide an exact solution for minimizing these data freshness metrics. The optimal sampling policy has a simple threshold-type structure, and the threshold can be efficiently computed by bisection search and fixed-point iterations. We uncover the following interesting property: if the last transmission is successful, the optimal policy may wait for a positive time period before generating the next sample and sending it out; otherwise, no waiting time should be added. The key technical approach developed in our results is given as follows: (i) The value function of the proposed policy is an exact solution to the Bellman equation. (ii) Under the contraction mapping assumption, the solution to the Bellman equation is unique, which guarantees optimality of our proposed threshold-based policy. Our results hold for (i) general non-decreasing age penalty functions, (ii) general delay distributions of both the forward and feedback channels, (iii) sampling problems both with or without a sampling rate constraint. Therefore, our paper extends previous studies on sampling for optimizing age (e.g., [15]–[18], [20], [21]). Although our sampling problem is in continuous time, it can be easily reduced to be in discrete time.
- When there is no sampling rate constraint, we provide necessary and sufficient conditions on the optimality of the zero-wait sampling policy [10] based on the choice of age penalty function, forward and feedback channels. Finally, numerical simulations show that our optimal policy can reduce the age compared with other approaches.

[2]We further compare our technical differences with past works in Section V-D.

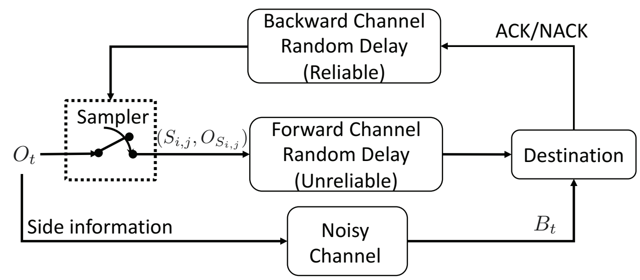


Figure 1: System model.

II. Estimation and the AoI

A. System Model

Consider a status update system that is composed of a source, a destination, a source-to-destination channel, and a destination-to-source channel, as is illustrated in Fig. 1. The source process O_t is sampled and delivered to the destination via the forward channel. The forward channel suffers from i.i.d. transmission failures, where $\alpha \in [0, 1)$ is the probability of failure. Upon each delivery, the destination then sends an 1-bit feedback message denoting whether the transmission is successful (ACK) or unsuccessful (NACK). The feedback is sent via the feedback channel that is reliable with an i.i.d. random delay.

To clarify the system model, we set $i \in \{1, 2, \ldots\}$ as the label of a successful delivery in chronological order. Let us denote the ith *epoch* to be the time period between the $(i-1)$th and the ith successful deliveries. We denote M_i as the total number of samples attempted during the ith epoch. Then, the M_i's are i.i.d. and has a geometric distribution with parameter $1 - \alpha$. We use j to describe the indices of samples at the ith epoch, where we have $1 \leq j \leq M_i$. The case $j = 1$ implies that the previous sample is successfully transmitted to the destination. Upon delivery, the destination immediately sends the feedback to the sampler and arrives at time $A_{i,j}$ via the backward channel with an i.i.d. delay $X_{i,j}$, which satisfies $\mathbb{E}[X_{i,j}] < \infty$. Then, the jth sample in the ith epoch is

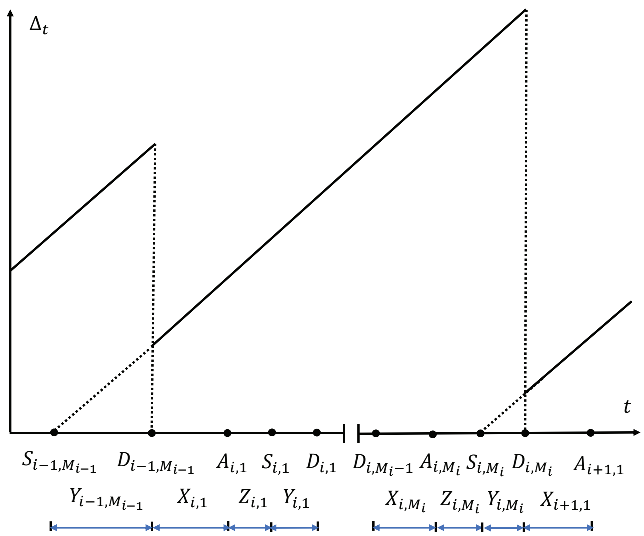


Figure 2: Evolution of the age Δ_t over time. The ith epoch starts from $D_{i-1,M_{i-1}}$ to D_{i,M_i}.

generated at $S_{i,j}$ and is delivered at $D_{i,j}$ through the forward channel with an i.i.d. delay $Y_{i,j}$, which satisfies $\mathbb{E}[Y_{i,j}] < \infty$.

We assume that the backward delays $X_{i,j}$'s and forward delays $Y_{i,j}$'s are mutually independent. In addition, the source generates a sample after receiving the feedback of the previous sample[3], i.e., $S_{i,j} \geq A_{i,j}$. In other words, we have a non-negative waiting time $Z_{i,j} \triangleq S_{i,j} - A_{i,j}$ for all epoch i and sample index j. Thus, the forward channel is always available for transmission at $S_{i,j}$, and the delivery time $D_{i,j}$ satisfies $D_{i,j} = S_{i,j} + Y_{i,j}$. By Wald's equation, the total transmission delay needed in each epoch has a finite expectation:

$$\mathbb{E}\left[\sum_{j=1}^{M_i}(X_{i,j}+Y_{i,j})\right] = \mathbb{E}\left[X_{i,j}+Y_{i,j}\right]\mathbb{E}\left[M_i\right] < \infty. \quad (1)$$

Age of information (or simply *age*) is the metric for evaluating data freshness and is equal to the time elapsed between the current time t and the generation time of the freshest delivered packet [23]. Let $U_t = \max_i\{S_{i,M_i} : D_{i,M_i} \leq t\}$. Note that only the M_ith sample is successfully delivered for the ith epoch. Then, the age of information Δ_t at the current time t is defined as

$$\Delta_t = t - U_t. \quad (2)$$

We plot the evolution of the age (2) in Fig. 2. Upon each successful delivery time D_{i,M_i}, the age decreases to Y_{i,M_i}, the transmission delay of the newly generated packet. At other time, the age increases linearly over time. The age is updated at the beginning of each epoch and keeps increasing during the epoch. Hence, the age is also determined by

$$\Delta_t = t - S_{i,M_i}, \text{ if } D_{i,M_i} \leq t < D_{i+1,M_{i+1}}. \quad (3)$$

[3]This assumption arises from the stop-and-wait mechanism. When the backward delay $X_{i,j} = 0$, the policy that samples ahead of receiving feedback is always suboptimal. The reason is that such a policy takes a new sample when the channel is busy and can be replaced by another policy that samples at the exact time of receiving feedback [17]. When $X_{i,j} \neq 0$, however, it may be optimal to transmit before receiving feedback, which is out of the scope of this paper.

B. Remote Estimation and Kalman Filter

We first introduce some notations. For any multi-dimensional vector O, we denote O^T as the transpose of O. We denote $\boldsymbol{I}_{n\times n}, \boldsymbol{0}_{n\times m}$ as the $n\times n$ identity matrix and $n\times m$ zero matrix, respectively. For a given $n\times n$ matrix $\boldsymbol{N}$, we set $tr(\boldsymbol{N})$ as the trace of $\boldsymbol{N}$, i.e., the summation of the diagonal elements of $\boldsymbol{N}$.

In this subsection, the source process O_t is an n-dimensional diffusion process that is defined as the solution to the following stochastic differential equation:

$$dO_t = -\boldsymbol{\Theta} O_t dt + \boldsymbol{\Sigma} dW_t, \quad (4)$$

where $\boldsymbol{\Theta}$ and $\boldsymbol{\Sigma}$ are $n\times n$ matrices, and W_t is the n-dimensional Wiener process such that $\mathbb{E}[W_t W_s^T] = \boldsymbol{I}_{n\times n}\min\{s,t\}$ for all $0 \leq t,s \leq \infty$. The process O_t represents the behavior of many physical systems such as the motion of a Brownian particle under friction and the motion of the monomers in dilute solutions [25]. At the destination, there is an estimator that provides estimations according to the received samples. One key difference from previous works (e.g., [11]–[13], [21]) is that the estimator not only receives the accurate samples $O_{S_{i,j}}$ at time $S_{i,j}$ but also has an instant noisy observation B_t of the process O_t, as is illustrated in Fig. 1. The observation process B_t is an m-dimensional vector, modeled as

$$B_t = \boldsymbol{H} O_t + V_t, \quad (5)$$

where $\boldsymbol{H}$ is an $n\times m$ matrix and V_t is a zero mean white noise process such that for all $t,s \geq 0$,

$$\mathbb{E}[V_t V_s^T] = \begin{cases} \boldsymbol{R} & t = s; \\ \boldsymbol{0}_{m\times m} & t \neq s, \end{cases} \quad (6)$$

$\boldsymbol{R}$ is an $m\times m$ positive definite matrix. We suppose that W_t and V_t are uncorrelated such that for all $t,s \geq 0$, $\mathbb{E}[W_t V_s^T] = \boldsymbol{0}_{n\times m}$.

The estimator provides an estimate $\hat{O}_t$ for the minimum mean squared error (MMSE) $\mathbb{E}[||O_t - \hat{O}_t||^2]$ based on the causally received information. Compared to [12], the MMSE in our study can be reduced due to the additional observation process B_t. Using the strong Markov property of O_t [26, Eq. (4.3.27)] and the assumption that the sampling times are independent of O_t, as is shown by Appendix A in our supplementary material, the MMSE estimator is determined by

$$\hat{O}_t = \mathbb{E}\left[O_t|\{B_\tau\}_{S_{i,M_i}\leq\tau\leq t}, O_{S_{i,M_i}}\right], t \in [D_{i,M_i}, D_{i+1,M_{i+1}}). \quad (7)$$

By (7), we find that $\hat{O}_t$ is equal to the estimate produced by the *Kalman filter* [24, Chapter 7]. Therefore, in this work, we use the Kalman filter as the estimator. At time t, the Kalman filter utilizes both the exact sample $O_{S_{i,M_i}}$ and noisy observation B_t and provides the minimum mean squared error (MMSE) estimation $\hat{O}_t$. Let $\boldsymbol{N}_t \triangleq \mathbb{E}[(O_t - \hat{O}_t)(O_t - \hat{O}_t)^T]$ be the covariance matrix of the estimation error $O_t - \hat{O}_t$. Hence, $\mathbb{E}[||O_t - \hat{O}_t||^2] = tr(\boldsymbol{N}_t)$.

According to (7), the estimation process works as follows: Once a sample is delivered to the Kalman filter at time D_{i,M_i}, the Kalman filter re-initiates itself with the initial condition

$\boldsymbol{N}_t = \mathbf{0}_{n\times n}$ when $t = S_{i,M_i}$ and starts a new estimation session. Then, during the time period $[D_{i,M_i}, D_{i+1,M_{i+1}})$, the Kalman filter uses the causal observations $\{B_\tau : S_{i,M_i} \leq \tau \leq t\}$ to estimate the process O_t.

Proposition 1. *The MMSE $tr(\boldsymbol{N}_t)$ of the process O_t is a non-decreasing function of the age Δ_t.*

Proof. See Appendix B in our supplementary material. □

As a result of Proposition 1, when the sampling times $S_{i,j}$'s are independent of O_t, the MMSE is still a non-decreasing function of the age Δ_t. When $S_{i,j}$'s are correlated to O_t, the MMSE is not necessary a function of Δ_t.

In the one-dimensional case, where $n = m = 1$, we use scalars $\theta, \sigma, h, r, n_t$ to replace the matrices $\boldsymbol{\Theta}, \boldsymbol{\Sigma}, \boldsymbol{H}, \boldsymbol{R}, \boldsymbol{N}_t$, respectively. The Ornstein–Uhlenbeck (OU) process is defined as a one-dimensional special case of diffusion process (4) where $\theta > 0$ [27]. Then, we have

Proposition 2. *Suppose that $n = m = 1$ and $\theta > 0$. Then, for $t \in [D_{i,M_i}, D_{i+1,M_{i+1}})$ and $i = 0, 1, 2, \ldots$, the MMSE n_t of the OU process O_t is given by*

$$n_t = \bar{n} - \frac{1}{l + \left(\frac{1}{\bar{n}} - l\right) e^{2\sqrt{\theta^2 + \frac{\sigma^2 h^2}{r}}\Delta_t}}, \tag{8}$$

where $\Delta_t = t - S_{i,M_i}$,

$$\bar{n} = \frac{-\theta r + \sqrt{(\theta r)^2 + \sigma^2 r h^2}}{h^2}, \tag{9}$$

$$l = \frac{h^2}{2\sqrt{(\theta r)^2 + \sigma^2 r h^2}}. \tag{10}$$

Moreover, n_t in (8) *is a bounded and non-decreasing function of the age Δ_t.*

Proof. See Appendix C in our supplementary material. □

When the side observation has zero knowledge of O_t, i.e., $h = 0$ for $t \geq 0$, then the estimator $\hat{O}_t$ is equal to that in [12]. Therefore, Proposition 2 reduces to [12, Lemma 4], i.e., the MMSE n_t is given by

$$n_t = \frac{\sigma^2}{2\theta}\left(1 - e^{-2\theta\Delta_t}\right), \tag{11}$$

moreover, n_t for $h = 0$ is a bounded and non-decreasing function of age Δ_t.

III. Problem Formulation for General Age Penalty

The function in Proposition 2 is not the only choice of nonlinear age functions. In this paper, to achieve data freshness in various applications, we consider a general type of age penalty function. The age penalty function $p : [0, \infty) \to \mathbb{R}$ is assumed to be non-decreasing and need not be continuous or convex. We further assume that $\mathbb{E}\left[\int_\delta^{\delta + \sum_{j=1}^{M_i}(X_{i,j}+Y_{i,j})} p(t)dt\right] < \infty$ and $\mathbb{E}\left[p\left(\delta + \sum_{j=1}^{M_i}(X_{i,j} + Y_{i,j})\right) dt\right] < \infty$ for any given δ.

We list another two categories of applications for the age penalty functions. First, the age penalty functions can be linear, polynomial, or exponential, depending on the dissatisfactions of the stale information updates in multiple practical settings such as the Internet of Things [28]. Second, some applications are shown to be closely related to nonlinear age functions, such as auto-correlation function of the source, remote estimation, and information based data freshness metric [17].

We then define the sampling policies below. We denote $\mathcal{H}_{i,j}$ as the sample path of the history information previous to $A_{i,j}$, including sampling times, forward channel conditions, and channels delays. We denote Π as the collection of sampling policies $\{S_{i,j}\}_{i,j}$ such that $S_{i,j} \geq A_{i,j}$ for each (i, j), and $S_{i,j}(ds_{i,j}|\mathcal{H}_{i,j})$ is a *Borel measurable stochastic kernel* [29, Chapter 7] for any possible $\mathcal{H}_{i,j}$. Further, we assume that $T_i = S_{i,M_i} - S_{i-1,M_{i-1}}$ is a regenerative process: there exists an increasing sequence $0 \leq k_1 < k_2 < \ldots$ of finite random variables such that the post-k_j process $\{T_{k_j+i}, i = 0, 1, \ldots\}$ has the same distribution as the post-k_1 process $\{T_{k_1+i}, i = 0, 1, \ldots\}$ and is independent of the pre-k_j process $\{T_i, i = 1, 2, \ldots, k_j - 1\}$; in addition, $\mathbb{E}\left[k_{j+1} - k_j\right] < \infty$, $\mathbb{E}\left[S_{k_1,M_{k_1}}\right] < \infty$ and $0 < \mathbb{E}\left[S_{k_{j+1},M_{k_{j+1}}} - S_{k_j,M_{k_j}}\right] < \infty$, $j = 1, 2, \ldots$[4]

The authors in [13] have stated that: to reduce the estimation error related to the Wiener process, it may be optimal to wait on both the source and the destination before transmission. However, in this paper, it is sufficient to only wait at the source to minimize the age. To validate this statement, consider any policy that waits on both the source and the destination. We first remove the waiting time at the destination. Then, at the source, we add up the removed waiting time. The replaced policy we propose has the same age performance as the former one.

Our objective in this paper is to optimize the long-term average expected age penalty under a sampling rate constraint:

$$p_{\text{opt}} = \inf_{\pi\in\Pi} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T p(\Delta_t)dt\right], \tag{12}$$

$$\text{s.t.} \ \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[C(T)\right] \leq f_{\max}. \tag{13}$$

Here, $C(T)$ is the total number of samples taken by time T, and $f_{\max}$ is the maximum allowed sampling rate. The constraint (13) is added because in practice, the sensor may need to keep working for a long time with limited amount of energy. To avoid triviality, the optimal objective value p_{opt} in (12) satisfies $p_{\text{opt}} < \bar{p}$, where $\bar{p} = \lim_{\delta\to\infty} p(\delta)$. Note that we allow $\bar{p}$ to be infinite.

A. An Additional Assumption and Its Rationale

We will utilize the following assumption in this paper.

Assumption 1. *If $\alpha > 0$, the backward delay $X_{i,j} \in [0, \bar{x}]$, and the waiting time $Z_{i,j} \in [0, \bar{z}]$ for all i, j. For any positive $\bar{x}, \bar{z}$ (that can be sufficiently large), there exists an increasing positive function $v(\delta)$ such that the function $G(\delta) =$*

[4]In this paper, we will optimize $\limsup_{T\to\infty}(1/T)\mathbb{E}\left[\int_0^T p(\Delta_t)dt\right]$. However, a nicer objective is to optimize $\lim_{n\to\infty}\mathbb{E}\left[\int_0^{D_{n,M_n}} p(\Delta_t)dt\right] / \mathbb{E}\left[D_{n,M_n}\right]$. If T_i is a regenerative process, then the two objective functions are equal [30], [31]. If no conditions are applied, they are different.

$\mathbb{E}\left[\int_{\delta}^{\delta+\bar{x}+\bar{z}+Y_{i,j}} |p(t)|dt\right]$ *satisfies* $\max_{\delta\geq 0}|G(\delta)/v(\delta)|<\infty$. *In addition, there exists* $\rho\in(0,1)$ *and a positive integer* m, *such that*

$$\alpha^m \frac{\mathbb{E}\left[v(\delta+m\bar{x}+m\bar{z}+\sum_{j=1}^{m}Y_j)\right]}{v(\delta)} \leq \rho \tag{14}$$

holds for all $\delta\geq 0$, *where* $Y_1,\ldots,Y_m$ *are an i.i.d. sequence with the same distribution as the* $Y_{i,j}$*'s.*

When the forward channel is reliable, i.e., $\alpha=0$, then Assumption 1 is negligible by letting $v(\delta)=G(\delta)$. Thus, Assumption 1 restricts on the choices of age penalty $p(\cdot)$ when $\alpha>0$. Note that the optimal sampling policy of the cases $\alpha=0$ and $X_{i,j}=0$ has been solved in [16], [17].

In the following corollary, we provide a list of age penalties $p(\cdot)$ that Assumption 1 is satisfied for $\alpha>0$.

Corollary 1. *For any one of the following conditions, Assumption 1 holds:*

(a) The penalty function $p(\cdot)$ *is bounded, i.e.,* $\bar{p}<\infty$.

(b) There exists $n>0$ *such that* $p(\delta)=O(\delta^n)$,[5] *and the* $Y_{i,j}$*'s have a finite* $n+1$*-moment, i.e.,* $\mathbb{E}\left[Y_{i,j}^{n+1}\right]<\infty$.

(c) There exists $a>0$ *and* $b<1$ *such that* $\int_0^{\delta}p(t)dt=O(e^{a\delta^b})$, *and the* $Y_{i,j}$*'s are bounded.*

Proof. See Appendix D in our supplementary material. □

Most of the literatures of MDP have shown that the *value function* of an optimal policy is the solution to the *Bellman equation*. In this paper, we figure out a policy and its value function that is indeed the solution to the Bellman equation. If the Bellman equation has a unique solution, then our proposed policy is optimal. Otherwise, we cannot guarantee the optimality of our proposed policy. Assumption 1 arises from the contraction mapping assumption [32], [33] that guarantees that the Bellman equation has a unique solution. In other words, Assumption 1 is a sufficient condition for the Bellman equation to have a unique solution. Corollary 1 implies that there are a wide range of age penalty functions that satisfy Assumption 1. For example, the age penalty function derived in Proposition 2 satisfies Assumption 1. Indeed, Assumption 1 holds if the age penalty function grows exponentially at some bounded intervals. For all cases of the age penalty functions we have mentioned, the constants $\bar{z},\bar{x}$ can be *sufficiently large*. Therefore, in this paper, we set the constants $\bar{z},\bar{x}$ to be sufficiently large.

IV. Optimal Sampling policy

In this section, we provide an optimal solution to (12). The optimal solution is described by the waiting times $Z_{i,j}'s$ throughout this paper.

A. Optimal Sampling Policy without Sampling Rate Constraint

When there is no sampling rate constraint, i.e., $f_{\max}=\infty$, we have the following result:

Theorem 1. *If* $f_{max}=\infty$, $p(\cdot)$ *is non-decreasing, the* $Y_{i,j}$*'s are i.i.d. with finite mean* $\mathbb{E}[Y_{i,j}]<\infty$, *the* $X_{i,j}$*'s are i.i.d. with finite mean* $\mathbb{E}[X_{i,j}]<\infty$, *the* $Y_{i,j}$*'s and the* $X_{i,j}$*'s are mutually independent, and Assumption 1 holds, then the optimal solution to* (12) *is given by*

$$Z_{i,1}(\beta)=\inf_{z}\Big\{z\geq 0: \mathbb{E}_{Y'}\left[p(Y_{i-1,M_{i-1}}+X_{i,1}+z+Y')\ \middle|\ Y_{i-1,M_{i-1}},X_{i,1}\right]\geq\beta\Big\}, \tag{15}$$

$$Z_{i,j}(\beta)=0 \quad j=2,3,\ldots, \tag{16}$$

$Y'=Y_{i,1}+\sum_{j=2}^{M_i}(X_{i,j}+Y_{i,j})$,[6] *and* β *is the unique solution to*

$$\mathbb{E}\left[\int_{Y_{i-1,M_{i-1}}}^{Y_{i-1,M_{i-1}}+X_{i,1}+Z_{i,1}(\beta)+Y'}p(t)dt\right] - \beta\mathbb{E}\left[X_{i,1}+Z_{i,1}(\beta)+Y'\right]=0. \tag{17}$$

Moreover, $\beta=p_{opt}$ *is the optimal objective value of* (12).

Proof. See Section V. □

In Theorem 1, the case $j=1$ in (15) means that the previous transmission (of the M_{i-1}th sample in the $(i-1)$th epoch) is successful, and the system starts the new epoch from $i-1$ to i. Since the age drops to $Y_{i-1,M_{i-1}}$ at the successful delivery time $D_{i-1,M_{i-1}}$, the current age state at arrival time $A_{i,1}$ is $Y_{i-1,M_{i-1}}+X_{i,1}$. The case $j=2,3,\ldots$ in (16) means that the previous transmission is unsuccessful, and the system stays within epoch i.

Theorem 1 provides an optimal policy with an interesting structure. First, by (15), in each epoch, the optimal waiting time for the first sample $Z_{i,1}(\beta)$ has a simple threshold type structure on the current age $Y_{i-1,M_{i-1}}+X_{i,1}$. Since the waiting times for $j=2,3,\ldots$ are zero, Y' is the remaining transmission delay needed for the next successful delivery. Note that β is equal to the optimal objective value p_{opt} in problem (12). Therefore, the waiting time $Z_{i,1}(\beta)$ in (15) is chosen such that the expected age penalty upon delivery is no smaller than p_{opt}. Second, by (16), the source sends the packet as soon as it receives negative feedback, i.e., the previous transmission is not successful. This is quite different from most of the previous works assuming reliable channels, e.g., [16]–[19], where for all samples, the source may wait for some time before transmitting a new sample.

We call a sampling policy to be *stationary* if each sampling time is decided by the current age state and the previous backward delay. We call a sampling policy to be *deterministic* if each sampling time chooses a value with probability 1 (w.p. 1). We remind that the optimal policy we proposed in Theorem 1 is *stationary and deterministic*. This stationary and deterministic policy depends only on the current age state and the previous backward delay, not on the sample index j. For example, when $j=1$, the previous backward delay is $X_{i,1}$, and the current age state is $Y_{i-1,M_{i-1}}+X_{i,1}$. For general value of j, the previous backward delay is $X_{i,j}$, and we suppose that

[5] We denote $f(\delta)=O(g(\delta))$ if there exists some nonnegative constants c and δ' such that $|f(\delta)|\leq c|g(\delta)|$ for all $\delta>\delta'$.

[6] In this paper, we set the summation operator $\sum_{j=a}^{b}$ to be 0 if $b<a$ for any given integers a,b.

Algorithm 1: Bisection method for solving (17)

Given function $f(\beta) = f_1(\beta) - \beta f_2(\beta)$. k_1 close to $\underline{p}$, k_2 close to $\bar{p}$, $k_1 < k_2$, and tolerance ϵ small.
repeat
  $\beta = \frac{1}{2}(k_1 + k_2)$
  if $f(\beta) < 0$: $k_2 = \beta$. **else** $k_1 = \beta$
until $k_2 - k_1 < \epsilon$
return β

the current age state is $\Delta_{i,j} + X_{i,j}$. Then, the stationary and deterministic policy, which has an equivalent form of (15),(16) in Theorem 1, is as follows:

$$Z_{i,j}(\beta) = \inf_z \Big\{ z \geq 0 : \mathbb{E}_{Y'}\left[p(\Delta_{i,j} + X_{i,j} + z + Y_j') \mid \Delta_{i,j}, X_{i,j}\right] \geq \beta \Big\}. \tag{18}$$

Here, $Y_j' = Y_{i,j} + \sum_{k=j+1}^{M_i}(X_{i,k} + Y_{i,k})$, conditioned that $M_i \geq j$. Since M_i is geometrically distributed, M_i conditioned that $M_i \geq j$ has the same distribution as $M_i + j - 1$. Therefore, Y_j' has the same distribution as Y' defined in Theorem 1, which guarantees that the optimal decision $Z_{i,j}$ is stationary.

The root of β in (17) can be solved efficiently. According to (17), we can use a low complexity algorithm such as bisection search and fixed-point iterations to obtain the optimal objective value p_{opt}. The bisection search approach to solving p_{opt} is illustrated in Algorithm 1. For simplicity, we set

$$f_1(\beta) = \mathbb{E}\left[\int_{Y_{i-1,M_{i-1}}}^{Y_{i-1,M_{i-1}} + X_{i,1} + Z_{i,1}(\beta) + Y'} p(t)dt\right], \tag{19}$$

$$f_2(\beta) = \mathbb{E}\left[X_{i,1} + Z_{i,1}(\beta) + Y'\right]. \tag{20}$$

Then, the function $f(\beta) \triangleq f_1(\beta) - \beta f_2(\beta)$ satisfies the following mathematical property:

Lemma 1. *(1) $f(\beta)$ is concave, and strictly decreasing in $\beta \in [\underline{p}, \bar{p}) \cap \mathbb{R}$, where $\underline{p} = p(0)$ and $\bar{p} = \lim_{\delta \to \infty} p(\delta)$.*

(2) There exists a unique root $\beta \in [\underline{p}, \bar{p}) \cap \mathbb{R}$ such that $f(\beta) = 0$.

Proof. See Appendix L in our supplementary material. □

Therefore, the solution to Algorithm 1 is unique.

One common sampling policy is the zero-wait policy, which samples the packet once it receives the feedback, i.e., $Z_{i,j} = 0$ for all (i, j) [10]. The zero-wait policy maximizes the throughput and minimizes the delay. However, by Theorem 1, the zero-wait policy may be suboptimal on age. The following result provides the necessary and sufficient condition when the zero-wait policy is optimal.

Corollary 2. *If $f_{max} = \infty$, $p(\cdot)$ is non-decreasing, the $Y_{i,j}$'s are i.i.d. with finite mean $\mathbb{E}[Y_{i,j}] < \infty$, the $X_{i,j}$'s are i.i.d. with finite mean $\mathbb{E}[X_{i,j}] < \infty$, the $Y_{i,j}$'s and the $X_{i,j}$'s are mutually independent, and Assumption 1 holds, then the zero-wait policy is optimal if and only if*

$$\text{ess}\inf \mathbb{E}_{Y'}\left[p(Y + X + Y') \mid Y, X\right] \geq \frac{\mathbb{E}\left[\int_Y^{Y+X+Y'} p(t)dt\right]}{\mathbb{E}\left[X + Y'\right]}, \tag{21}$$

where $Y' = Y_{i,1} + \sum_{j=2}^{M_i}(X_{i,j} + Y_{i,j})$, $Y = Y_{i-1,M_{i-1}}$, $X = X_{i,1}$, and we denote $\text{ess}\inf E = \inf\{e : \mathbb{P}(E \leq e) > 0\}$ for any random variable E.

Proof. See Appendix M in our supplementary material. □

When the channel delays are constant, we can get from Corollary 2 that

Corollary 3. *If $f_{max} = \infty$, $p(\cdot)$ is non-decreasing and satisfies Assumption 1, and the $Y_{i,j}$'s, $X_{i,j}$'s are constants, then the zero-wait policy is the solution to problem* (12).

Proof. See Appendix N in our supplementary material. □

Theorem 1 is an extension to [17], [18]. When the forward channel is reliable, i.e., $M_i = 1$ for all i or $\alpha = 0$, Theorem 1 can be reduced to the result in [18]. Further, we extend [18] in two ways: (i) The age penalty $p(\cdot)$ is allowed to be negative or discontinuous. (ii) The channel delays $Y_{i,1}, X_{i,1}$ have a finite expectation and do not need to be bounded. Note that when $M_i = 1$, Assumption 1 is negligible. When $M_i = 1$, and there is no backward delay ($X_{i,1} = 0$), our result reduces to [17, Theorem 1].

The study in [20, Theorem 2] proves the optimality of the zero-wait policy among the deterministic policies under an unreliable forward channel. This result corresponds to Corollary 3, a special case of Theorem 1. Our paper extends [20] in two folds: (i) We allow the policy space Π to be randomized. Among randomized policies, due to the disturbances on the previous sampling times, the current sampling time is dependent on the previous ones, which is different from [20]. (ii) We consider random two-way delays, extending the constant one-way delay in [20].

B. Optimal Sampling Policy with Sampling Rate Constraint

For general values of f_{max}, we propose the following result that extends Theorem 1:

Theorem 2. *If $p(\cdot)$ is non-decreasing, the $Y_{i,j}$'s are i.i.d. with finite mean $\mathbb{E}[Y_{i,j}] < \infty$, the $X_{i,j}$'s are i.i.d. with finite mean $\mathbb{E}[X_{i,j}] < \infty$, the $Y_{i,j}$'s and the $X_{i,j}$'s are mutually independent, and Assumption 1 holds, then* (15)-(17) *is the optimal solution to* (12), *if the following condition holds:*

$$\mathbb{E}\left[X_{i,1} + Z_{i,1}(\beta) + Y'\right] > \frac{1}{f_{max}(1 - \alpha)}, \tag{22}$$

where $Y' = Y_{i,1} + \sum_{j=2}^{M_i}(X_{i,j} + Y_{i,j})$. Otherwise, an optimal solution is as follows:

$$Z_{i,1}(\beta) = \begin{cases} Z_{min}(\beta) & \text{w.p. } \lambda, \\ Z_{max}(\beta) & \text{w.p. } 1 - \lambda. \end{cases} \tag{23}$$

$$Z_{i,j} = 0, \quad j = 2, 3, \ldots, M_i, \tag{24}$$

$Z_{min}(\beta)$ and $Z_{max}(\beta)$ are described as follows:

$$Z_{min}(\beta) = \inf_{z} \Big\{ z \geq 0 : \\ \mathbb{E}_{Y'}\left[p(Y_{i-1,M_{i-1}} + X_{i,1} + z + Y') \;\middle|\; Y_{i-1,M_{i-1}}, X_{i,1} \right] \geq \beta \Big\}, \tag{25}$$

$$Z_{max}(\beta) = \inf_{z} \Big\{ z \geq 0 : \\ \mathbb{E}_{Y'}\left[p(Y_{i-1,M_{i-1}} + X_{i,1} + z + Y') \;\middle|\; Y_{i-1,M_{i-1}}, X_{i,1} \right] > \beta \Big\}. \tag{26}$$

β is determined by

$$\mathbb{E}\left[X_{i,1} + Z_{min}(\beta) + Y'\right] \leq \frac{1}{f_{max}(1-\alpha)} \\ \leq \mathbb{E}\left[X_{i,1} + Z_{max}(\beta) + Y'\right]. \tag{27}$$

The probability λ is given by

$$\lambda = \frac{\mathbb{E}\left[X_{i,1} + Z_{max}(\beta) + Y'\right] - \frac{1}{f_{max}(1-\alpha)}}{\mathbb{E}\left[Z_{max}(\beta) - Z_{min}(\beta)\right]}. \tag{28}$$

Proof. See Section V. □

According to Theorem 2, the proposed optimal policy may be randomized or deterministic. When $p(\cdot)$ is strictly increasing, we have $Z_{\min}(\beta) = Z_{\max}(\beta)$. Similar to Theorem 1, the optimal policy is stationary and deterministic in current age and previous backward delay. When $p(\cdot)$ is not strictly increasing, $Z_{\min}(\beta)$ and $Z_{\max}(\beta)$ may be different, so the optimal policy at $j = 1$ is a random mixture of two deterministic sampling times. Note that when $Z_{\min}(\beta)$ and $Z_{\max}(\beta)$ may be different, the random optimal policy *may be nonstationary*. In addition, we can solve (27) via low complexity algorithms such as bisection search.

When $M_i = 1$ (or $\alpha = 0$) and $X_{i,j} = 0$, Theorem 2 reduces to [17, Theorem2]. Combined with the discussions in Section IV-A, we conclude that our paper is an extension to some recent studies on sampling for optimizing age, e.g., [15]–[18], [20], [21].

V. Proof of the Main Result

In this section, we provide the proof of our main results: Theorem 1 and Theorem 2. In Section V-A, we utilize the Lagrangian dual problem of the original long-term average problem and reformulate the Lagrangian dual problem into a per-epoch MDP problem. In Section V-B, we solve the per-epoch MDP problem by formulating an exact optimal value function to the Bellman Equation, which is the key challenge to this paper. In Section V-C, we established zero duality gap to the Lagrangian problem, which ends our proof. Finally, in Section V-D, we summarize our technical contribution and compare it with some related works.

A. Reformulation of Problem (12)

In this subsection, we decompose the original problem to a per-epoch problem. The idea is motivated by recent studies that reformulate the average problem into a per-sample problem [11], [12], [16]–[18].

Since $\{S_{i,M_i}\}_i$ follows a regenerative process, by renewal theory, [30, Section 6.1], [34],

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T p(\Delta_t)dt\right] \tag{29}$$

$$= \lim_{n\to\infty} \frac{\mathbb{E}\left[\int_0^{D_{n,M_n}} p(\Delta_t)dt\right]}{\mathbb{E}\left[D_{n,M_n}\right]} \tag{30}$$

$$= \lim_{n\to\infty} \frac{\sum_{i=1}^n \mathbb{E}\left[\int_{D_{i-1,M_{i-1}}}^{D_{i,M_i}} p(\Delta_t)dt\right]}{\sum_{i=1}^n \mathbb{E}\left[D_{i,M_i} - D_{i-1,M_{i-1}}\right]}. \tag{31}$$

In addition,

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[C(T)\right] = \lim_{n\to\infty} \frac{\mathbb{E}\left[\sum_{i=1}^n M_i\right]}{\mathbb{E}\left[S_{n,M_n}\right]} \tag{32}$$

$$= \lim_{n\to\infty} \frac{n}{(1-\alpha)\mathbb{E}\left[D_{n,M_n}\right]}. \tag{33}$$

From (29)-(33), the original problem (12) is equivalent to

$$p_{\text{opt}} = \inf_{\pi\in\Pi} \lim_{n\to\infty} \frac{\sum_{i=1}^n \mathbb{E}\left[\int_{D_{i-1,M_{i-1}}}^{D_{i,M_i}} p(\Delta_t)dt\right]}{\sum_{i=1}^n \mathbb{E}\left[D_{i,M_i} - D_{i-1,M_{i-1}}\right]}, \tag{34}$$

$$\text{s.t.} \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[D_{i,M_i} - D_{i-1,M_{i-1}}\right] \geq \frac{1}{f_{\max}(1-\alpha)}. \tag{35}$$

We consider the following MDP with a parameter $c \in \mathbb{R}$:

$$h(c) = \inf_{\pi\in\Pi} \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathbb{E}\Bigg[\int_{D_{i-1,M_{i-1}}}^{D_{i,M_i}} p(\Delta_t)dt \\ - c\left(D_{i,M_i} - D_{i-1,M_{i-1}}\right)\Bigg], \tag{36}$$

$$\text{s.t.} \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[D_{i,M_i} - D_{i-1,M_{i-1}}\right] \geq \frac{1}{f_{\max}(1-\alpha)}. \tag{37}$$

By Dinkelbach's method [35], we have

Lemma 2. *[17, lemma 2]*
(i) $h(c) \lesseqqgtr 0$ if and only if $p_{opt} \lesseqqgtr c$.
(ii) The solution to (34) and (36) are equivalent.

We define the Lagrangian with $c = p_{\text{opt}}$:

$$L(\pi;\gamma) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathbb{E}\Bigg[\int_{D_{i-1,M_{i-1}}}^{D_{i,M_i}} p(\Delta_t)dt \tag{38}$$

$$- (p_{\text{opt}} + \gamma)\left(D_{i,M_i} - D_{i-1,M_{i-1}}\right)\Bigg] + \frac{\gamma}{f_{\max(1-\alpha)}}, \tag{39}$$

where $\gamma \geq 0$ is the dual variable. The primal problem is

$$l(\gamma) \triangleq \inf_{\pi\in\Pi} L(\pi;\gamma). \tag{40}$$

The dual problem is

$$d \triangleq \max_{\gamma\geq 0} l(\gamma). \tag{41}$$

Weak duality theorem [36], [37] implies that $d \le h(p_{\text{opt}})$. We will later show that the duality gap is 0, i.e., $d = h(p_{\text{opt}})$. Note that

$$D_{i,M_i} - D_{i-1,M_{i-1}} = \sum_{j=1}^{M_i}(X_{i,j} + Z_{i,j} + Y_{i,j}), \quad (42)$$

$$\mathbb{E}\left[\int_{D_{i-1,M_{i-1}}}^{D_{i,M_i}} p(\Delta_t)dt\right] \quad (43)$$

$$=\mathbb{E}\left[\int_{Y_{i-1,M_{i-1}}}^{Y_{i-1,M_{i-1}}+\sum_{j=1}^{M_i}(X_{i,j}+Z_{i,j}+Y_{i,j})} p(t)dt\right]. \quad (44)$$

Recall that the age decreases to $Y_{i-1,M_{i-1}}$ at time $D_{i-1,M_{i-1}}$. Note that $Y_{i-1,M_{i-1}}$ is independent of the history information by the sampling time $S_{i-1,M_{i-1}}$. Thus, the age evolution at the i^{th} epoch is independent of the sampling decisions from the previous epochs $0, 1, 2, ..., i-1$. Therefore, to solve (40), minimizing each epoch separately is sufficient. We define the policy space Π_i as the collection of sampling decisions $(Z_{i,1}, Z_{i,2}, ...)$ at epoch i such that the stochastic kernel

$$Z_{i,j}(dz_{i,j}|y_{i-1,M_{i-1}}, x_{i,1}, z_{i,1}, y_{i,1}, \ldots, z_{i,j-1}, y_{i,j-1}, x_{i,j})$$

is Borel measurable. The difference between Π_i and Π is that the sampling decisions in Π_i do not depend on the history information from previous epochs (except $Y_{i-1,M_{i-1}}$). Hence, it is easy to find that $\Pi_i \subset \Pi$.

Then, by the analysis of the previous paragraph, we have the following result:

Lemma 3. *An optimal solution to* (40) *satisfies*

$$\inf_{\pi\in\Pi_i} \mathbb{E}\Bigg[\int_{Y_{i-1,M_{i-1}}}^{Y_{i-1,M_{i-1}}+\sum_{j=1}^{M_i}(X_{i,j}+Z_{i,j}+Y_{i,j})} p(t)dt - (p_{opt}+\gamma)\sum_{j=1}^{M_i}(X_{i,j}+Z_{i,j}+Y_{i,j}) \;\Bigg|\; Y_{i-1,M_{i-1}}, X_{i,1}\Bigg]. \quad (45)$$

Thus, for any epoch i, we will solve $Z_{i,1}, Z_{i,2}, ...$ according to (45).

B. Solution to the Per-epoch Problem (45)

We will solve problem (45) given that $Y_{i-1,M_{i-1}} = \delta$ and $X_{i,1} = x$, where $\delta \ge 0$ and $x \ge 0$. Since the epoch number i does not affect problem (45), in this subsection, we will remove the subscription i from $M_i, X_{i,j}, Y_{i,j}, Z_{i,j}$ and replace them by M, X_j, Y_j, Z_j for the ease of descriptions. In addition, since we want to find out a solution to (34), we need to avoid that $l(\gamma) = -\infty$. Thus, we assume that γ satisfies $\inf_{z\ge 0}\{z : p(z) > p_{\text{opt}} + \gamma\} < \infty$.[7]

Different from [11], [12], [16], [17], the per-epoch problem (45) is an MDP with multiple samples and cannot be reduced to the per-sample problem in the sense that the age is not refreshed under failed transmissions. According to (45), we define the value function $J_{\pi,\gamma}$ under a policy $\pi \in \Pi_i$ with an initial age state $\delta \ge 0$ (at delivery time) and backward delay $x \ge 0$:

$$J_{\pi,\gamma}(\delta, x) =\mathbb{E}\Bigg[\int_{\delta}^{\delta+\sum_{j=1}^{M}(X_j+Z_j+Y_j)} p(t)dt - (p_{\text{opt}}+\gamma)\sum_{j=1}^{M}(X_j+Z_j+Y_j) \;\Bigg|\; X_1 = x\Bigg] \quad (46)$$

$$=\mathbb{E}\left[\sum_{j=1}^{M} g_\gamma(\Delta_j, X_j, Z_j) \;\Bigg|\; \Delta_1 = \delta, X_1 = x\right], \quad (47)$$

where the instant cost function $g_\gamma(\delta, x, z)$ with state (δ, x) and action z is defined as

$$g_\gamma(\delta, x, z) =\mathbb{E}_Y\left[\int_{\delta}^{\delta+x+z+Y} p(t)dt - (p_{\text{opt}}+\gamma)(x+z+Y)\right], \quad (48)$$

where Y has the same delay distribution as the Y_j's and the age state evolution is described as

$$\Delta_{j+1} = \Delta_j + X_j + Z_j + Y_j, \quad j = 1, 2, ...M-1, \quad (49)$$

with initial age state $\Delta_1 = \delta$ and initial backward delay x. Also, the policy $\pi \in \Pi_i$ has a Borel measurable stochastic kernel $Z_j(dz_j|\delta_1, x_1, z_1, \ldots, \delta_j, x_j)$, and thus $J_{\pi,\gamma}(\delta, x)$ is Borel measurable [29, Chapter 9]. The above settings imply that problem (45) is equivalent to a shortest path MDP problem. Solving (45) is equivalent to solving

$$J_\gamma(\delta, x) = \inf_{\pi\in\Pi_i} J_{\pi,\gamma}(\delta, x). \quad (50)$$

When the channel state is reliable, i.e., $\alpha = 0$ or $M = 1$, problem (45) (or equivalently, (50)) becomes a single-sample problem, and there is no bound restriction to the instant cost function $g_\gamma(\delta, x, z)$. However, in the unreliable transmission case where $\alpha > 0$, problem (45) contains multiple samples. In the case of multiple samples, most of the literature of dynamic programming e.g., [29], [32], [33], [38]–[42] requires that the instant cost function $g_\gamma(\delta, x, z)$ is bounded from below. We have such a requirement.

Lemma 4. *There exists a value η such that $g_\gamma(\delta, x, z) \ge -\eta$ and $J_{\pi,\gamma}(\delta, x) \ge -\eta/(1-\alpha)$ for all (δ, x, z) and any policy $\pi \in \Pi_i$.*

Proof. See Appendix E in our supplementary material. □

Using Lemma 4 and Appendix F in our supplementary material, $J_{\pi,\gamma}(\delta, x)$ defined in (47) also equals to a discounted sum with discount factor α:

$$J_{\pi,\gamma}(\delta, x) = \sum_{j=1}^{\infty}\alpha^{j-1}\mathbb{E}\left[g_\gamma(\Delta_j, X_j, Z_j) \;\Big|\; \Delta_1 = \delta, X_1 = x\right]. \quad (51)$$

Note that (51) is motivated by [38, Chapter 5], illustrating that the discounted problem is equivalent to a special case of shortest path problem.

[7]If $\inf_{z\ge 0}\{z : p(z) > p_{\text{opt}}+\gamma\} = \infty$, this subsection implies that waiting for arbitrary large time can optimize (40). If such a policy optimizes (36), we have $p_{\text{opt}} = \bar{p}$, which contradicts to our assumption that $p_{\text{opt}} < \bar{p}$.

Recall that uncountable infimum of Borel measurable functions is not necessary Borel measurable. Problem (45) has an uncountable state space. Thus, the optimal value function $J_\gamma(\delta, x)$ defined in (50) *may not be Borel measurable*[8], despite that $J_{\pi,\gamma}(\delta, x)$ is Borel measurable for all $\pi \in \Pi_i$. Then, some well known theories may not satisfy, such as the optimality of the Bellman equation among Π_i. One of the methods to overcome this challenge is to enlarge the policy spaces. We define a collection of policies Π'_i such that the stochastic kernel $Z_j(dZ_j|\delta_1, x_1, z_1, \ldots, \delta_j, x_j)$ is universally measurable [29]. Note that every Borel measurable stochastic kernel is a universally measurable stochastic kernel, so we have $\Pi_i \subset \Pi'_i$.

Note that if $\pi \in \Pi'_i$, we also denote $J_{\pi,\gamma}(\delta, x)$ as the discounted cost of π given in (51). For all given age state δ and delay x, we define

$$J'_\gamma(\delta, x) = \inf_{\pi \in \Pi'_i} J_{\pi,\gamma}(\delta, x). \tag{52}$$

It is easy to see that $J'_\gamma(\delta, x) \leq J_\gamma(\delta, x)$. In this subsection, we will finally show that $J'_\gamma(\delta, x) = J_\gamma(\delta, x)$.

By Lemma 4, it is easy to show that $J_{\pi,\gamma} \geq -\eta/(1-\alpha)$ for all $\pi \in \Pi'$. Using $J_{\pi,\gamma} \geq -\eta/(1-\alpha)$ and [29, Corollary 9.4.1], $J'(\delta, x)$ is lower semianalytic [29]. Note that any real-valued Borel measurable function is lower semianalytic. This allows us to consider the Bellman operator based on a general lower semianalytic function $u(\delta, x)$. For any deterministic and stationary policy $\pi \in \Pi_i$ with Borel measurable decisions $\pi(\delta, x)$, we define an operator $T_{\pi,\gamma}$ on a function u:

$$\begin{aligned} & T_{\pi,\gamma} u(\delta, x) \\ =& g_\gamma(\delta, x, \pi(\delta, x)) + \alpha \mathbb{E}_{Y,X}\left[u(\delta + x + \pi(\delta, x) + Y, X)\right], \end{aligned} \tag{53}$$

where Y and X have the same distribution as the i.i.d. forward delay Y_j's and backward delay X_j's, respectively. We also define the Bellman operator T_γ on the function u:

$$T_\gamma u(\delta, x) = \inf_{z \in [0, \bar{z}]} g(\delta, x, z) + \alpha \mathbb{E}_{Y,X}\left[u(\delta + x + z + Y, X)\right]. \tag{54}$$

As is described in Assumption 1, the bound $\bar{z}$ is taken sufficiently large. Note that if the function $u(\delta, x)$ is Borel measurable, $T_\gamma u(\delta, x)$ is not necessary Borel measurable in the sense that uncountable infimum of Borel measurable functions is not necessary Borel measurable. However, if we extend $u(\delta, x)$ to be lower semianalytic, then $T_\gamma u(\delta, x)$ is also lower semianalytic [29, Proposition 7.47], i.e., T_γ is well-defined under lower semianalytic functions. Note that the expectation on a lower semianalytic function has the same definition with the expectation on a Borel measurable function. In all, we have

Lemma 5. *If $u(\delta, x)$ is lower semianalytic, then $T_{\pi,\gamma} u(\delta, x)$ and $T_\gamma u(\delta, x)$ are both lower semianalytic.*

Proof. See Appendix G in our supplementary material. □

We denote $u_1 = u_2$ if $u_1(\delta, x) = u_2(\delta, x)$ for all $\delta, x \in [0, \infty)$. Using the definition of $T_{\pi,\gamma}$ and T_γ, the discounted problem (52) has the following properties [29, Chapter 9.4]:

[8] see [29], [32] for counterexamples. In discrete-time system where the system time is slotted, we do not have this challenge.

Lemma 6. *If $p(\cdot)$ is non-decreasing, the Y_j's are i.i.d. with finite mean $\mathbb{E}[Y_j] < \infty$, the X_j's are i.i.d. with finite mean $\mathbb{E}[X_j] < \infty$, the Y_j's and the X_j's are mutually independent, then the optimal value function $J'_\gamma(\delta, x)$ defined in* (52) *satisfies the Bellman equation:*

$$J'_\gamma = T J'_\gamma, \tag{55}$$

i.e., the optimal value function J'_γ is a fixed point of T_γ.

To derive an optimal policy, we first provide two stationary and deterministic policies called $\mu_{\min,\gamma}$ and $\mu_{\max,\gamma}$. Then we will show that both $\mu_{\min,\gamma}$ and $\mu_{\max,\gamma}$ are the solution to problem (45).

Definition 1. *The stationary and deterministic policies $\mu_{min,\gamma}$ and $\mu_{max,\gamma}$ are defined as*

$$\mu_{min,\gamma}(\delta, x) = \max\{b_{min,\gamma} - \delta - x, 0\}, \tag{56}$$

$$\mu_{max,\gamma}(\delta, x) = \max\{b_{max,\gamma} - \delta - x, 0\}, \tag{57}$$

$$b_{min,\gamma} = \inf_c \left\{c \geq 0 : \mathbb{E}\left[p(c + Y')\right] \geq p_{opt} + \gamma\right\}, \tag{58}$$

$$b_{max,\gamma} = \inf_c \left\{c \geq 0 : \mathbb{E}\left[p(c + Y')\right] > p_{opt} + \gamma\right\}, \tag{59}$$

$$Y' \triangleq Y_1 + \sum_{j=2}^{M}(X_j + Y_j). \tag{60}$$

A randomized policy $\tilde{\mu}_{\lambda,\gamma} = \{Z_1, Z_2, ...\}$ with $\lambda \in [0,1]$ satisfies

$$Z_1 = \begin{cases} \mu_{min,\gamma}(\delta, x) & w.p.\ \lambda, \\ \mu_{max,\gamma}(\delta, x) & w.p.\ 1-\lambda. \end{cases} \tag{61}$$

$$Z_j = 0, \quad j = 2, 3, ..., M_i. \tag{62}$$

Using the definition of Π_i, we have $\mu_{\min,\gamma}, \mu_{\max,\gamma} \in \Pi_i$, and $\tilde{\mu}_{\lambda,\gamma} \in \Pi_i$ for all $\lambda \in [0,1]$ [29, Chapter 7].

Upon delivery of the first sample, age of $\mu_{\min,\gamma}$ and $\mu_{\max,\gamma}$ increase to $\Delta_2 = \delta + x + \mu_{\min,\gamma}(\delta, x) + Y_1$ and $\delta + x + \mu_{\max,\gamma}(\delta, x) + Y_1$, which are larger than $\max\{\delta + x, b_{\min,\gamma}\}$, $\max\{\delta + x, b_{\max,\gamma}\}$, respectively. Then, the waiting time for the second sample is $\mu_{\min,\gamma}(\Delta_2, X_2) = 0$ and $\mu_{\max,\gamma}(\Delta_2, X_2) = 0$, respectively. Thus, the waiting time at stage 2,... is 0 under $\mu_{\min,\gamma}$ and $\mu_{\max,\gamma}$. Therefore, we have $\mu_{\min,\gamma} = \tilde{\mu}_{0,\gamma}$ and $\mu_{\max,\gamma} = \tilde{\mu}_{1,\gamma}$. Note that when we do not consider sampling rate constraint, then $\gamma = 0$, and the policy $\mu_{\min,\gamma}$ is equivalent to (15) and (16) in Theorem 1. It remains to show that $\mu_{\min,\gamma}$ and $\mu_{\max,\gamma}$ are indeed optimal to problem (45).

Recall that we denote $J_{\pi,\gamma}(\delta, x)$ to be the value function with initial state δ, x under a policy π. Then, we have the following key result:

Lemma 7. *If $p(\cdot)$ is non-decreasing, the Y_j's are i.i.d. with finite mean $\mathbb{E}[Y_j] < \infty$, the X_j's are i.i.d. with finite mean $\mathbb{E}[X_j] < \infty$, the Y_j's and the X_j's are mutually independent, then the value functions $J_{\mu_{min,\gamma}}(\delta, x)$ and $J_{\mu_{max,\gamma}}(\delta, x)$ satisfy*

$$J_{\mu_{min,\gamma}} = T_\gamma J_{\mu_{min,\gamma}} = J_{\mu_{max,\gamma}} = T_\gamma J_{\mu_{max,\gamma}}. \tag{63}$$

Moreover, for any $\lambda \in [0,1]$, we have $J_{\tilde{\mu}_{\lambda,\gamma}} = J_{\mu_{min,\gamma}} = J_{\mu_{max,\gamma}}$.

Proof. We provide the proof sketch of $J_{\mu_{\min,\gamma}} = T_\gamma J_{\mu_{\min,\gamma}}$ here and replace $\mu_{\min,\gamma}$ by μ for simplicity. We relegate the detailed proof to Appendix H in our supplementary material.

We define the q-function $Q_\mu(\delta, x, z)$ as the cost of starting at state (δ, x), waiting for time z for the first sample, and then following policy μ for the remaining samples [38, Section 6] [43, Chapter 3]. It is easy to find that

$$Q_\mu(\delta, x, z) = g_\gamma(\delta, x, z) + \alpha \mathbb{E}\left[J_\mu(\delta + x + z + Y, X)\right]. \quad (64)$$

From (64) and (54), showing $J_\mu = T_\gamma J_\mu$ is equivalent to showing that $Q_\mu(\delta, x, z) \geq J_\mu(\delta, x)$ for all the waiting time $z \geq 0$, age δ and x. We have stated that μ has a nice structure: for any initial state, the waiting times of stage $2, 3...$ are 0. Thus, we can derive the closed form expression of $J_\mu(\delta, x)$ according to (46) (where $Z_j = 0$ for $j \geq 2$). Also, given the definition of $Q_\mu(\delta, x, z)$, we can derive the expression of $Q_\mu(\delta, x, z)$ with the similar form of (46). By comparing $Q_\mu(\delta, x, z)$ and $J_\mu(\delta, x)$, we can finally show that $Q_\mu(\delta, x, z) \geq J_\mu(\delta, x)$ for all (δ, x, z). □

Lemma 7 tells that $J_{\mu_{\min,\gamma}}$ (or equivalently, $J_{\mu_{\max,\gamma}}$) is a fixed point of T_γ. From Lemma 6, the optimal value function J'_γ is also a fixed point of T_γ. To show that $J'_\gamma = J_{\mu_{\min,\gamma}}$, it remains to show that the fixed point of T_γ is unique. If the age penalty $p(\cdot)$ is bounded, $J_{\pi,\gamma}(\delta, x)$ is bounded for any policy $\pi \in \Pi'_i$. Then, according to the contraction mapping theorem, the bellman equation (55) has a unique bounded solution [32], [34], [42], i.e., $J_{\mu_{\min,\gamma}} = J'_\gamma$. Note that there may be unbounded solutions to (55) [40], [41]. If $p(\cdot)$ is unbounded, we will utilize Assumption 1 to show the uniqueness.

Let us denote $\Lambda = [0, \infty) \times [0, \bar{x}]$, where $\bar{x}$ is the bound of X_j mentioned in Assumption 1. In Assumption 1, we have defined an increasing function $v(\delta) : [0, \infty) \to \mathbb{R}^+$ (also called the *weighted function*). The *weighted sup-norm* $\|u\|$ of a function $u : \Lambda \to \mathbb{R}$ is defined as

$$\|u\| = \max_{(\delta,x)\in\Lambda} \frac{|u(\delta, x)|}{v(\delta)}. \quad (65)$$

Let $B(\Lambda)$ denote the set of all lower semianalytic functions $u : \Lambda \to \mathbb{R}$ such that $\|u\| < \infty$. Note that any real-valued Borel measurable function is lower semianalytic. From [32, p. 47], [29, Lemma 7.30.2], $B(\Lambda)$ is complete under the weighted sup-norm.

Lemma 8. *If $p(\cdot)$ is non-decreasing, the Y_j's are i.i.d. with finite mean $\mathbb{E}[Y_j] < \infty$, the X_j's are i.i.d. with finite mean $\mathbb{E}[X_j] < \infty$, the Y_j's and the X_j's are mutually independent, and Assumption 1 holds, then for all $\pi \in \Pi_i$, $J_{\pi,\gamma} \in B(\Lambda)$.*

Proof. See Appendix I in our supplementary material. □

Then, the following result shows the uniqueness of the Bellman equation $T_\gamma u = u$.

Lemma 9. *If $p(\cdot)$ is non-decreasing, the Y_j's are i.i.d. with finite mean $\mathbb{E}[Y_j] < \infty$, the X_j's are i.i.d. with finite mean $\mathbb{E}[X_j] < \infty$, the Y_j's and the X_j's are mutually independent, and Assumption 1 holds, the following conditions hold:*

(a) For any lower semianalytic function $u : \Lambda \to \mathbb{R}$, if $u \in B(\Lambda)$, then $T_{\pi,\gamma} u \in B(\Lambda)$ for all deterministic and stationary policy $\pi \in \Pi_i$, and $T_\gamma u \in B(\Lambda)$.

(b) The Bellman operator T_γ has an m-stage contraction mapping with modulus ρ, i.e., for all $u_1, u_2 \in B(\Lambda)$,

$$\|T^m_\gamma u_1 - T^m_\gamma u_2\| \leq \rho \|u_1 - u_2\|, \quad (66)$$

where constants $\rho \in (0, 1)$ and m are mentioned in Assumption 1, and the weighted sup-norm $\|\cdot\|$ is defined in (65).

(c) There exists a unique function $u \in B(\Lambda)$ such that $T_\gamma u = u$.

Proof. See Appendix J in our supplementary material. □

From Lemma 8, $J_{\mu_{\min,\gamma}} \in B(\Lambda)$. From Lemma 9(c), Lemma 7 and $J_{\mu_{\min,\gamma}} \in B(\Lambda)$, $J_{\mu_{\min,\gamma}}$ (or equivalently, $J_{\mu_{\max,\gamma}}$) is the unique solution to $T_\gamma u = u$. From Lemma 6, $J_{\mu_{\min,\gamma}} = J'_\gamma$. Since $\mu_{\min,\gamma}, \mu_{\max,\gamma} \in \Pi_i$ and $\Pi_i \subset \Pi'_i$, $\mu_{\min,\gamma}$ and $\mu_{\max,\gamma}$ are the optimal policies in Π_i. Note that $\mu_{\min,\gamma} = \tilde{\mu}_{0,\gamma}$ and $\mu_{\max,\gamma} = \tilde{\mu}_{1,\gamma}$. Using Lemma 7, we immediately get the final result:

Lemma 10. *A collection of optimal policies to problem (45) is $\{\tilde{\mu}_{\lambda,\gamma} : \lambda \in [0, 1]\}$ described in Definition 1.*

C. Optimal Solution to (36) When $c = p_{opt}$

Section V-B provides the optimal solution to (45) given the initial states $Y_{i-1,M_{i-1}} = \delta$ and $X_{i,1} = x$. Using Lemma 10 and strong duality, we have the following result.

Theorem 3. *If $p(\cdot)$ is non-decreasing, the $Y_{i,j}$'s are i.i.d. with finite mean $\mathbb{E}[Y_{i,j}] < \infty$, the $X_{i,j}$'s are i.i.d. with finite mean $\mathbb{E}[X_{i,j}] < \infty$, the $Y_{i,j}$'s and the $X_{i,j}$'s are mutually independent, and Assumption 1 holds, then $\mu_{min,0}$ described in Definition 1 is an optimal solution to (36) with $c = p_{opt}$, if the following condition holds:*

$$\mathbb{E}\left[X_{i,1} + \mu_{min,0}(Y_{i-1,M_{i-1}}, X_{i,1}) + Y'\right] > \frac{1}{f_{max}(1-\alpha)}, \quad (67)$$

where $Y' = Y_{i,1} + \sum_{j=2}^{M_i}(X_{i,j} + Y_{i,j})$. Otherwise, $\tilde{\mu}_{\lambda,\gamma}$ is an optimal solution to (36) with $c = p_{opt}$, where γ is determined by

$$\begin{aligned}\mathbb{E}\left[X_{i,1} + \mu_{min,\gamma}(Y_{i-1,M_{i-1}}, X_{i,1}) + Y'\right] \leq \frac{1}{f_{max}(1-\alpha)} \\ \leq \mathbb{E}\left[X_{i,1} + \mu_{max,\gamma}(Y_{i-1,M_{i-1}}, X_{i,1}) + Y'\right], \quad (68)\end{aligned}$$

and the probability λ is given by

$$\lambda = \frac{\mathbb{E}\left[X_{i,1} + \mu_{max,\gamma}(Y_{i-1,M_{i-1}}, X_{i,1}) + Y'\right] - \frac{1}{f_{max}(1-\alpha)}}{\mathbb{E}\left[\mu_{max,\gamma}(Y_{i-1,M_{i-1}}, X_{i,1}) - \mu_{min,\gamma}(Y_{i-1,M_{i-1}}, X_{i,1})\right]}. \quad (69)$$

Proof. See Appendix K in our supplementary material. □

By taking $\beta = p_{\text{opt}} + \gamma$, Theorem 2 is directly shown by Theorem 3.

In addition, note that Theorem 1 is directly shown by Lemma 10, by taking $\beta = p_{\text{opt}}$ and $\gamma = 0$. In other words, $\mu_{\min,0}$ is an optimal solution to (12) when $f_{\max} = \infty$.

D. Discussion

Many existing studies on AoI sampling assume that the transmission channel is error-free, i.e., $M_i = 1$ for all i, e.g., [11]–[13], [15]–[18]. Due to the renewal property, their original problems are reduced to a per-sample problem. Similarly, our result is equivalent to the per-epoch problem illustrated in (45). If $M_i = 1$, problem (45) reduces to a per-sample

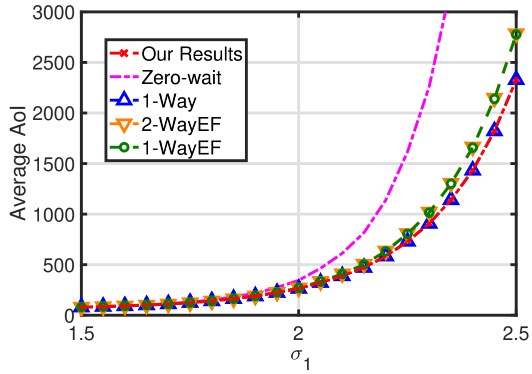


Figure 3: Average AoI versus the parameter σ_1 of the forward channel, where $\sigma_2 = 1.5$ and $\alpha = 0.8$.

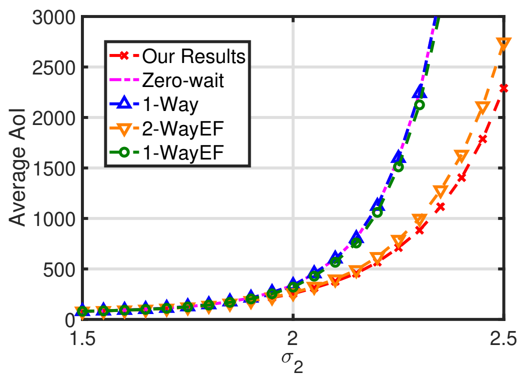


Figure 4: Average AoI versus the parameter σ_2 of the backward channel, where $\sigma_1 = 1.5$ and $\alpha = 0.8$.

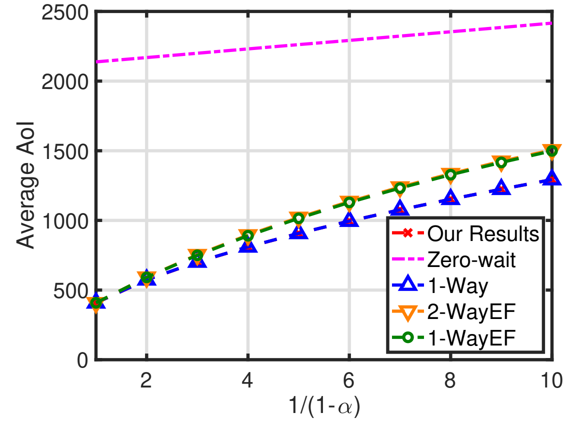


Figure 5: Average AoI versus $1/(1-\alpha)$, where $\sigma_1 = 2.3$ and $\sigma_2 = 1.5$.

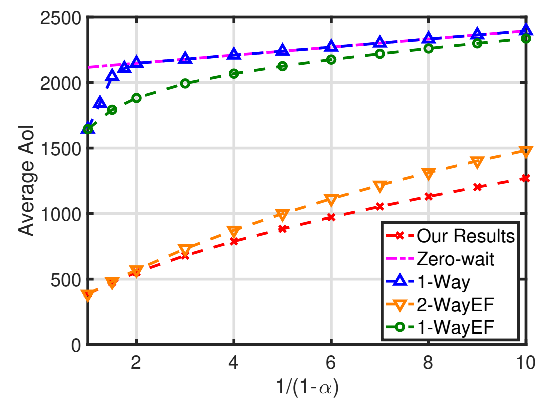


Figure 6: Average AoI versus $1/(1-\alpha)$, where $\sigma_1 = 1.5$ and $\sigma_2 = 2.3$.

problem, where there is only one decision $Z_{i,1}$ and is solved using convex optimization. However, when $M_i \neq 1$, problem (45) is an MDP that contains multiple samples. This MDP cannot be solved by convex optimization (e.g., [15]–[18]) or optimal stopping rules (e.g., [11]–[13]).

Therefore, one of the technical contributions in this paper is to accurately solve the MDP in (45). We summarize the high-level idea of solving (45): First, in Lemma 6, among the extended policy space Π_i' with universally measurable stochastic kernel [29, Chapter 7], the optimal policy satisfies the Bellman equation (55). Then, in Lemma 7, we provide the exact value function that is the solution to the Bellman Equation. Finally, under Assumption 1 and Lemma 9, the uniqueness of the Bellman equation is guaranteed.

In addition, although we focus on continuous-time systems in this paper, our results can be easily reduced to the discrete-time systems by removing the content of measure theory.

VI. Numerical Results

In this section, we compare our optimal sampling policy with the following sampling policies:

1. Zero-wait: Let the waiting time $Z_{i,j} = 0$, i.e., the source transmits a sample once it receives the feedback.

2. One-way (1-way): It falsely assumes that the backward delay $X_{i,j} = 0$ despite that $X_{i,j}$ may not be zero. In other words, it derives $Z_{i,j}$ by solving (15)—(17) and letting $X_{i,j} = 0$.

3. Two-way Error-free (2-wayEF) [18]: It assumes that the forward channel's probability of failure $\alpha = 0$ despite that α may not be zero. In other words, it derives $Z_{i,j}$ by solving (15)—(17) and letting $M_i = 1$.

4. One-way Error-free (1-wayEF) [17]: It assumes that $X_{i,j} = 0$ and $\alpha = 0$.

Note that all of the four policies have a stationary threshold structure. Therefore, similar to our optimal policy, each policy may have a nonzero waiting time only if the previous transmission was successful.

In this section, we consider linear age penalty $p(\delta) = 2\delta$ and lognormal distributions on both forward and backward delay with scale parameters σ_1, σ_2, respectively. Note that the lognormal random variable with scale parameter σ is expressed as $e^{\sigma R}$, where R is the standard normal random variable. The numerical results below show that our proposed policy always achieves the lowest average age.

Fig. 3 and Fig. 4 illustrate the relationship between age and σ_1, σ_2, respectively. In Fig. 3, we plot the evolution of average age in σ_1 given that $\sigma_2 = 1.5$ and $\alpha = 0.8$. As σ_1 increases, the lognormal distribution of the forward channel becomes more heavy tailed. We observe that Zero-wait policy evolves much quicker than other policies in σ_1. In addition, 2-wayEF and 1-wayEF policies grow faster than the optimal policy in σ_1. Note that our optimal policy differs from the other policies even if the average AoI is close. For example, when $\sigma_1 = 2.0$, the average AoI of 2-wayEF is close to that of our optimal policy. However, we observe that the waiting time of our optimal policy and 2-wayEF are 78 and 49, respectively, if the current age state and the last backward delay are 2 and 1, respectively. In Fig. 4, we fix $\sigma_1 = 1.5$ and plot the average age of the listed policies in σ_2. Unlike Fig. 3, 1-way and 1-wayEF policies perform poorly since they fail to take highly random backward delay into account.

Fig. 5 and Fig. 6 depict the evolution of average age in $1/(1-\alpha)$, where $(\sigma_1, \sigma_2) = (2.3, 1.5)$ and $(\sigma_1, \sigma_2) = (1.5, 2.3)$, respectively. Note that $1/(1-\alpha)$ is the average number of samples attempted for a successful transmission. In Fig. 5 and Fig. 6, when $1/(1-\alpha)$ increases, the gap between 2-wayEF policy and our optimal policy increases. In Fig. 6, since $\sigma_2 > \sigma_1$, the tail of backward delay is heavier than that of forward delay. Thus, 1-way and 1-wayEF, which neglect the knowledge of backward delay, fail to improve the age performance.

In summary, when either one of the channels is highly random, (i) Zero-wait policy is far from optimal, (ii) the age performance of 1-wayEF or 2-wayEF policy gets worse if the forward channel is more unreliable, (iii) 1-way and 1-wayEF polices are far from optimal if the backward channel is highly random.

VII. Conclusion

In this paper, we design a sampling policy to optimize data freshness, where the source generates the samples and sends to the remote destination via a fading forward channel, and the acknowledgements are sent back via a backward channel. We overcome the curse of dimensionality that arises from the time-varying forward channel conditions and the randomness of the channel delays in both directions. We reveal that the optimal sampling policy has a simple threshold based structure, and the optimal threshold is computed efficiently.

References

[1] J. Pan, A. M. Bedewy, Y. Sun, and N. B. Shroff, “Optimizing sampling for data freshness: Unreliable transmissions with random two-way delay,” *in Proc. IEEE INFOCOM*, 2022.

[2] A. M. Bedewy, Y. Sun, and N. B. Shroff, “The age of information in multihop networks,” *IEEE/ACM Transactions on Networking*, vol. 27, no. 3, pp. 1248–1257, 2019.

[3] A. M. Bedewy, Y. Sun, and N. B. Shroff, “Minimizing the age of information through queues,” *IEEE Transactions on Information Theory*, vol. 65, no. 8, pp. 5215–5232, 2019.

[4] Y. Sun, E. Uysal-Biyikoglu, and S. Kompella, “Age-optimal updates of multiple information flows,” in *IEEE INFOCOM WKSHPS*, pp. 136–141, IEEE, 2018.

[5] A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, “Optimal sampling and scheduling for timely status updates in multi-source networks,” *IEEE Transactions on Information Theory*, vol. 67, no. 6, pp. 4019–4034, 2021.

[6] A. M. Bedewy, Y. Sun, R. Singh, and N. B. Shroff, “Low-power status updates via sleep-wake scheduling,” *IEEE/ACM Transactions on Networking*, 2021.

[7] J. Pan, A. M. Bedewy, Y. Sun, and N. B. Shroff, “Minimizing age of information via scheduling over heterogeneous channels,” in *Proceedings of MobiHoc*, pp. 111–120, 2021.

[8] Y. Zou, K. T. Kim, X. Lin, and M. Chiang, “Minimizing age-of-information in heterogeneous multi-channel systems: A new partial-index approach,” in *Proceedings of MobiHoc*, pp. 11–20, 2021.

[9] Z. Qian, F. Wu, J. Pan, K. Srinivasan, and N. B. Shroff, “Minimizing age of information in multi-channel time-sensitive information update systems,” in *IEEE INFOCOM*, pp. 446–455, 2020.

[10] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, “Age of information: An introduction and survey,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1183–1210, 2021.

[11] 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, 2020.

[12] T. Z. Ornee and Y. Sun, “Sampling and remote estimation for the ornstein-uhlenbeck process through queues: Age of information and beyond,” *IEEE/ACM Transactions on Networking*, vol. 29, no. 5, pp. 1962–1975, 2021.

[13] C.-H. Tsai and C.-C. Wang, “Unifying AoI minimization and remote estimation—optimal sensor/controller coordination with random two-way delay,” in *IEEE INFOCOM*, pp. 466–475, 2020.

[14] K. Huang, W. Liu, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, “Real-time remote estimation with hybrid ARQ in wireless networked control,” *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3490–3504, 2020.

[15] R. D. Yates, “Lazy is timely: Status updates by an energy harvesting source,” in *2015 IEEE International Symposium on Information Theory (ISIT)*, pp. 3008–3012, 2015.

[16] 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 Transactions on Information Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.

[17] Y. Sun and B. Cyr, “Sampling for data freshness optimization: Non-linear age functions,” *Journal of Communications and Networks*, vol. 21, no. 3, pp. 204–219, 2019.

[18] C.-H. Tsai and C.-C. Wang, “Age-of-information revisited: Two-way delay and distribution-oblivious online algorithm,” in *2020 IEEE International Symposium on Information Theory (ISIT)*, pp. 1782–1787, 2020.

[19] C.-H. Tsai and C.-C. Wang, “Jointly minimizing AoI penalty and network cost among coexisting source-destination pairs,” in *2021 IEEE International Symposium on Information Theory (ISIT)*, pp. 3255–3260, 2021.

[20] A. Arafa, K. Banawan, K. G. Seddik, and H. V. Poor, “Sample, quantize, and encode: Timely estimation over noisy channels,” *IEEE Transactions on Communications*, vol. 69, no. 10, pp. 6485–6499, 2021.

[21] M. Klügel, M. H. Mamduhi, S. Hirche, and W. Kellerer, “AoI-penalty minimization for networked control systems with packet loss,” in *IEEE INFOCOM WKSHPS*, pp. 189–196, 2019.

[22] J. Pan, A. M. Bedewy, Y. Sun, and N. B. Shroff, “Age-optimal scheduling over hybrid channels,” *submitted to IEEE Transactions on Mobile Computing, eprint arXiv:2012.09403*, 2021.

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

[24] H. V. Poor, *An introduction to signal detection and estimation*. Springer Science & Business Media, 2013.

[25] H. C. Öttinger, *Stochastic processes in polymeric fluids: tools and examples for developing simulation algorithms*. Springer Science & Business Media, 2012.

[26] G. Peskir and A. Shiryaev, *Optimal stopping and free-boundary problems*. Springer, 2006.

[27] S. Finch, “Ornstein-Uhlenbeck process,” 2004.

[28] T. Park, W. Saad, and B. Zhou, “Centralized and distributed age of information minimization with nonlinear aging functions in the internet of things,” *IEEE Internet of Things Journal*, vol. 8, no. 10, pp. 8437–8455, 2020.

[29] D. P. Bertsekas and S. Shreve, *Stochastic optimal control: the discrete-time case*. 2004.

[30] P. J. Haas, *Stochastic petri nets: Modelling, stability, simulation*. Springer Science & Business Media, 2006.

[31] R. G. Gallager, *Stochastic processes: theory for applications*. Cambridge University Press, 2013.

[32] D. P. Bertsekas, *Dynamic programming and optimal control*, vol. 2. Athena scientific Belmont, MA, 1995.

[33] D. P. Bertsekas, *Abstract dynamic programming*. Athena Scientific Nashua, NH, USA, 2018.

[34] S. M. Ross, J. J. Kelly, R. J. Sullivan, W. J. Perry, D. Mercer, R. M. Davis, T. D. Washburn, E. V. Sager, J. B. Boyce, and V. L. Bristow, *Stochastic processes*, vol. 2. Wiley New York, 1996.

[35] W. Dinkelbach, "On nonlinear fractional programming," *Management science*, vol. 13, no. 7, pp. 492–498, 1967.

[36] D. Bertsekas, A. Nedic, and A. Ozdaglar, *Convex analysis and optimization*, vol. 1. Athena Scientific, 2003.

[37] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[38] D. P. Bertsekas, *Dynamic programming and optimal control*, vol. 1. Athena scientific Belmont, MA, 1995.

[39] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[40] L. I. Sennott, "A new condition for the existence of optimum stationary policies in average cost Markov decision processes-unbounded cost case," in *1986 25th IEEE Conference on Decision and Control*, pp. 1719–1721, 1986.

[41] 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.

[42] L. I. Sennott, "A new condition for the existence of optimal stationary policies in average cost Markov decision processes," *Operations research letters*, vol. 5, no. 1, pp. 17–23, 1986.

[43] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

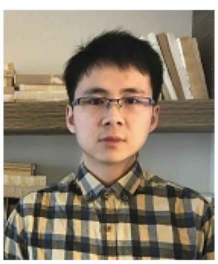
[44] M.-A. Poubelle, R. R. Bitmead, and M. R. Gevers, "Fake algebraic riccati techniques and stability," *IEEE Transactions on Automatic Control*, vol. 33, no. 4, pp. 379–381, 1988.

[45] S. Resnick, *A probability path*. Springer, 2019.

[46] D. P. Bertsekas, "Nonlinear programming," *Journal of the Operational Research Society*, vol. 48, no. 3, pp. 334–334, 1997.

[47] D. Butnariu and A. N. Iusem, *Totally convex functions for fixed points computation and infinite dimensional optimization*, vol. 40. Springer Science & Business Media, 2000.

[48] W. Rudin, *Principles of mathematical analysis*, vol. 3. McGraw-hill New York, 1976.



Jiayu Pan received the B.S. degree in Physics from Shanghai Jiao Tong University, Shanghai, China in 2017. He is currently pursuing the Ph.D. degree with the Electrical and Computer Engineering Department, at the Ohio State University, OH, USA. His research interests include wireless communication and age of information. He received the First Prize Scholarship from Shanghai Jiao Tong University in 2015.

Ahmed M. Bedewy received the B.S. and M.S. degrees in electrical and electronics engineering from Alexandria University, Alexandria, Egypt, in 2011 and 2015, respectively, and the Ph.D. degree in electrical and computer engineering from The Ohio State University, OH, USA, in 2021. His research interests include wireless communication, cognitive radios, resource allocation, communication networks, information freshness, optimization, and scheduling algorithms. He received the Awarded Certificate of Merit and First Class Honors for being one of the top ten undergraduate students, for the period 2006–2008, and the First, for the period 2008–2011, in electrical and electronics engineering. His article received the runner-up for the Best Paper Award of ACM MobiHoc 2020.

Yin Sun (Senior Member, IEEE) is currently an Assistant Professor with the Department of Electrical and Computer Engineering, Auburn University. He received his B.Eng. and Ph.D. degrees in electronic engineering from Tsinghua University in 2006 and 2011, respectively. From 2011 to 2017, he was a Post-Doctoral Scholar and a Research Associate with The Ohio State University. His research interests include wireless networks, age of information, machine learning, robotic control, and information theory. He co-authored a monograph *Age of Information: A New Metric for Information Freshness* (Morgan and Claypool Publishers, 2019). He has been a Guest Editor of the *IEEE Journal on Selected Areas in Communications* for the special issue on "Age of Information in Real-time Systems and Networks," a Guest Editor of *Entropy* for the special issue on "Age of Information: Concept, Metric and Tool for Network Control," a Guest Editor of *Frontiers in Communications and Networks* for the special issue on "Age of Information," an Associate Editor of the *IEEE Transactions on Network Science and Engineering*, and an Editor of the *Journal of Communications and Networks*. His articles received the Best Student Paper Award from the IEEE/IFIP WiOpt 2013, the Best Paper Award from the IEEE/IFIP WiOpt 2019, runner-up for the Best Paper Award of ACM MobiHoc 2020, and the 2021 *Journal of Communications and Networks (JCN)* Best Paper Award. He received the Auburn Author Award in 2020. He co-founded the Age of Information Workshop in 2018. He is a member of the ACM.

Ness B. Shroff (S'91–M'93–SM'01–F'07) received the Ph.D. degree in electrical engineering from Columbia University in 1994. He joined Purdue University immediately thereafter as an Assistant Professor with the School of Electrical and Computer Engineering. At Purdue, he became a Full Professor of ECE and the director of a university-wide center on wireless systems and applications in 2004. In 2007, he joined The Ohio State University, where he holds the Ohio Eminent Scholar Endowed Chair in networking and communications, in the departments of ECE and CSE. He holds or has held visiting (chaired) professor positions at Tsinghua University, Beijing, China, Shanghai Jiaotong University, Shanghai, China, and IIT Bombay, Mumbai, India. He has received numerous best paper awards for his research and is listed in Thomson Reuters' on The World's Most Influential Scientific Minds, and is noted as a Highly Cited Researcher by Thomson Reuters. He currently serves as the steering committee chair for ACM Mobihoc and the Editor in Chief of the IEEE/ACM Transactions on Networking. He received the IEEE INFOCOM Achievement Award for seminal contributions to scheduling and resource allocation in wireless networks.