Sampling of the Wiener Process for Remote Estimation Over a Channel With Unknown Delay Statistics

Haoyue Tang\textsuperscript{©}, Member, IEEE, Yin Sun\textsuperscript{©}, Senior Member, IEEE, Member, ACM, and Leandros Tassiulas, Fellow, IEEE, ACM

Abstract—In this paper, we study an online sampling problem of the Wiener process. The goal is to minimize the mean squared error (MSE) of the remote estimator under a sampling frequency constraint when the transmission delay distribution is unknown. The sampling problem is reformulated into an optional stopping problem, and we propose an online sampling algorithm that can adaptively learn the optimal stopping threshold through stochastic approximation. We prove that the cumulative MSE regret grows with rate $\Omega((\ln k)^{1\over k})$, where $k$ is the number of samples. Through Le Cam’s two point method, we show that the worst-case cumulative MSE regret of any online sampling algorithm is lower bounded by $\Omega(1\ln k)$. Hence, the proposed online sampling algorithm is minimax order-optimal. Finally, we validate the performance of the proposed algorithm via numerical simulations.

Index Terms—Age of Information, online learning, stochastic approximation.

I. INTRODUCTION

The omnipresence of the autonomous driving and the intelligent manufacturing systems involve tasks of sampling and remotely estimating fresh status information. For example, in autonomous driving systems, status information such as the position and the instant speed of cars keep changing, and the controller has to estimate the update-to-date status based on samples collected from the surrounding sensors. To ensure efficient control and system safety, it is important to estimate the fresh status information precisely under limited communication resources and random channel conditions.

To measure the freshness of the status update information, the Age of Information (AoI) metric has been proposed in [1]. By definition, AoI captures the difference between the current time and the time-stamp at which the freshest information available at the destination was generated. It is revealed that the AoI minimum sampling and transmission strategies behave differently from utility maximization and delay minimization [2]. Samples with fresher content should be delivered to the destination in a timely manner [3].

When the evolution of the dynamic source can be modeled by a random signal process, the mean square estimation error (MSE) based on the available information at the receiver can be used to capture freshness. Sampling to minimize the MSE of the random process in different communication networks are studied in [4], [5], [6], [7], [8], and [9]. Considering that the dynamic source is a Wiener process, the optimum sampling policy that minimizes the estimation MSE is shown to have a threshold structure, i.e., a new sample should be taken once the difference between the actual signal value and the estimate based on past samples exceed a certain threshold. Such thresholds also holds for the Ornstein-Uhlenbeck process [5], [10] and the Gaussian Markov source [9]. The optimum sampling thresholds can be obtained by the bisection search [7] or iterative thresholding [11] if the delay distribution and the statistics of the channel are known in advance.

When the statistics of the communication channel is unknown, the problem of sampling and transmissions for data freshness optimization can be formulated into a sequential decision making problem [12], [13], [14], [15], [16]. By using the AoI as the freshness metric, [12], [13], [14] design online link rate selection algorithms based on stochastic bandits. When the channels are time-varying and the transmitter has an average power constraint, [17], [18], [19], [20], [21] employ reinforcement learning algorithms to minimize the average AoI under unknown channel statistics. Notice that in applications such as the remote estimation, a linear AoI cannot fully capture the data freshness. To solve this problem, Tripathi et al. model the information freshness to be a time-varying function of the AoI [15], and a robust online learning algorithm is proposed. The above research tackles with unknown packet loss rate or utility functions, the problem of designing online algorithms under unknown delay statistics are not well studied.

The iterative thresholding algorithm proposed in [11] can be applied in the online setting when the delay statistics is unknown, the problem of sampling and transmissions for data freshness optimization can be formulated into a sequential decision making problem [12], [13], [14], [15], [16]. By using the AoI as the freshness metric, [12], [13], [14] design online link rate selection algorithms based on stochastic bandits. When the channels are time-varying and the transmitter has an average power constraint, [17], [18], [19], [20], [21] employ reinforcement learning algorithms to minimize the average AoI under unknown channel statistics. Notice that in applications such as the remote estimation, a linear AoI cannot fully capture the data freshness. To solve this problem, Tripathi et al. model the information freshness to be a time-varying function of the AoI [15], and a robust online learning algorithm is proposed. The above research tackles with unknown packet loss rate or utility functions, the problem of designing online algorithms under unknown delay statistics are not well studied.

The iterative thresholding algorithm proposed in [11] can be applied in the online setting when the delay statistics is unknown, the problem of sampling and transmissions for data freshness optimization can be formulated into a sequential decision making problem [12], [13], [14], [15], [16]. By using the AoI as the freshness metric, [12], [13], [14] design online link rate selection algorithms based on stochastic bandits. When the channels are time-varying and the transmitter has an average power constraint, [17], [18], [19], [20], [21] employ reinforcement learning algorithms to minimize the average AoI under unknown channel statistics. Notice that in applications such as the remote estimation, a linear AoI cannot fully capture the data freshness. To solve this problem, Tripathi et al. model the information freshness to be a time-varying function of the AoI [15], and a robust online learning algorithm is proposed. The above research tackles with unknown packet loss rate or utility functions, the problem of designing online algorithms under unknown delay statistics are not well studied.
unknown, whereas the convergence rate and the optimality of the derived online algorithm are not well understood.

In this paper, we consider an online sampling problem, where a sensor transmits status updates of the Wiener source to a destination through a channel with random delay. Our goal is to design a sampling policy that minimizes the estimation error when the delay distribution is unknown a priori. The main contributions of this paper are as follows:

- The design of the MSE minimum sampling policy is reformulated as an optimal stopping problem. By analyzing the sufficient conditions of the optimum threshold, we propose an online sampling policy that learns the optimum stopping threshold adaptively through stochastic approximation. Compared with [11], [22], and [23], the operation of the proposed algorithm does not require prior knowledge of an upper bound of the optimum threshold.

- We prove that the time averaged MSE of the proposed algorithm converges almost surely to the minimum MSE if the fourth order moment of the transmission delay is finite (Theorem 1). In addition, it is shown that the MSE regret, i.e., the sub-optimality gap between the expected cumulative MSE of the proposed algorithm and the optimum policy with distribution knowledge, grows at a speed of \(O(\sqrt{k})\), where \(k\) is the number of samples (Corollary 1). The perturbed ordinary differential equation (ODE) method is a popular tool for establishing the convergence rate of stochastic approximation algorithms [24]. However, this tool requires either the threshold being learned is in a bounded closed set, or the second moment of the updating directions are bounded. Because our algorithm does not require an upper bound on the optimum threshold, and the essential supremum of the transmission delay could be unbounded, we need to develop a new method for convergence rate analysis, which is based on the Lyapunov drift method for heavy traffic analysis.

- Using the Le Cam’s two point method, we show that for any causal algorithm that makes sampling decision based on historical information, under the worst case delay distribution, the estimation error of the optimum sampling threshold after \(k\) samples is lower bounded by \(\Omega(\frac{1}{k})\). Using similar ideas from the Le Cam’s method, we turn computation of the cumulative MSE regret into an optimization problem, and we show that the minimax MSE regret is lower bounded by \(\Omega(\ln k)\) (Theorem 4). By combining Theorem 1 and Theorem 4, we obtain that the proposed online sampling algorithm achieves the minimax order-optimal regret.

- We validate the performance of the proposed algorithm via numerical simulations. In contrast to [11], the proposed algorithm could meet an average sampling frequency constraint.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

As is depicted in Fig. 1, we revisit the status update system in [3], [7], and [25], where a sensor takes samples from a Wiener process and transmits the samples to a receiver through a network interface queue. The network interface serves the update packets in a non-preemptive fashion. In particular, the network interface must finish sending the current packet, before switching to serve another packet. An ACK is sent back to the sensor once an update packet is cleared at the interface. We assume that the transmission duration after passing the network interface is negligible.

Let \(X_t \in \mathbb{R}\) denote the value of the Wiener process at time \(t \in \mathbb{R}^+\). The sampling time-stamp of the \(k\)-th sample, denoted by \(S_k\), is determined by the sensor at will. Based on the FCFS principle, the network interface will start serving the \(k\)-th packet after the \((k-1)\)-th packet is cleared at the network interface and arrived at the receiver. We assume that the service time \(D_k\) are independent and identically distributed (i.i.d) with a probability distribution \(\mathbb{P}_D\), and \(D_k\) is independent of the wiener process \(X_t\). The reception time of the \(k\)-th packet, denoted by \(R_k\), satisfies the following recursive formula: \(R_k = \max\{S_k, R_{k-1}\} + D_k\) and we define \(R_0 = 0\) for simplicity. Let \(\overline{D} := \mathbb{E}_{D \sim \mathbb{P}_D}[D]\) denote the average transmission delay. We assume \(\overline{D}\) is lower bounded by \(\overline{D}_{th} > 0\).

B. MMSE Estimation

Let \(i(t) := \max_{k \in \mathbb{N}}\{k| R_k \leq t\}\) be the index of the latest sample received by the destination at time \(t\), and let \(M_t := (\{X_{S_k}, S_k, D_{k}^{(t)}\}_{k=1}^{i(t)}, t)\) denote the information stored at the receiver up to time \(t\). The minimum mean-square error (MMSE) estimator [26, Section 2.2] is:

\[
\hat{X}_t = \mathbb{E}[X_t|M_t] = X_{S_{i(t)}}.
\]

We use a sequence of sampling time instants \(\pi \triangleq \{S_k\}_{k=1}^{\infty}\) to represent a sampling policy. The expected time average mean square error (MSE) under \(\pi\) is denoted by \(\overline{\mathcal{E}}_\pi\), i.e.,

\[
\overline{\mathcal{E}}_\pi \triangleq \lim_{T \to \infty} \sup_{\pi} \mathbb{E} \left[ \frac{1}{T} \int_{t=0}^{T} (X_t - X_{S_{i(t)}})^2 dt \right].
\]

C. Problem Formulation

Our goal in this work is to design one sampling policy that can minimize the MSE for the estimator when the delay distribution \(\mathbb{P}_D\) is unknown. Specifically, we focus on the set of causal policies denoted by \(\Pi\), where each policy \(\pi \in \Pi\) selects the sampling time \(S_k\) of the \(k\)-th sample based on the transmission delay \(\{D_{k'}\}_{k'<k}\) and Wiener process evolution \(\{X_{t'}\}_{t' \leq S_k}\) from the past. The transmission delay and the evolution of the Wiener process in the future cannot be used to decide the sampling time. Due to the energy constraint, we require that the sampling frequency should below a certain
The optimal sampling problem is organized as follows:

**Problem 1 MMSE minimization:**

\[
\text{mse}_{\text{opt}} \triangleq \inf_{\pi \in \Pi} \lim_{T \to \infty} \sup \mathbb{E} \left[ \frac{1}{T} \int_{t=0}^{T} (\hat{X}_t - X_t)^2 \, dt \right],
\]

s.t. \( \lim_{T \to \infty} \sup \mathbb{E} \left[ \frac{i(T)}{T} \right] \leq f_{\text{max}}. \) (3b)

**III. Problem Solution**

In this section, the MSE minimization problem (i.e., Problem 1) is reformulated into an optimal stopping problem. Let \( \pi^* \) be an optimum policy whose average MSE achieves \( \text{mse}_{\text{opt}}. \) Sufficient conditions for \( \pi^* \) are provided in Subsection III-B. The online sampling algorithm \( \pi_{\text{online}} \) is provided in Subsection III-C and Subsection III-D characterizes the behaviors of the online sampling policy.

**A. Markov Decision Reformulation 1**

According to [7, Theorem 1], when the second moment of the transmission delay is bounded, i.e., \( \mathbb{E}[D^2] < \infty \), policy \( \pi^* \) should not take a new sample before the previous sample is delivered to the destination. As is depicted in Fig. 2, the waiting time between the delivery time of the \( k \)-th sample and the sampling time of the \( (k + 1) \)-th sample is denoted by \( W_k \geq 0. \) Define frame \( k \) as the time interval between the sampling time-stamp of the \( k \)-th and the \( (k + 1) \)-th sample. The following corollary enables us to reformulate Problem 1 into a Markov Decision Process.

**Lemma 1:** Let \( \mathcal{I}_k := (D_k, (X_{S_{k+1}} - X_{S_k}))_{t \geq 0} \) denote the recent information of the sampler in frame \( k. \) The set of sampling policies that determine the waiting time \( W_k \) only based on the recent information \( \mathcal{I}_k \) is denoted by \( \Pi_{\text{recent}}. \) Since for each frame \( k, \) the difference \( X_{S_{k+1}} - X_{S_k} \) evolves as a Wiener process that is independent of the past \( \{X_{S_{k+t}} - X_{S_k}\}_{t < k}, \) Problem 1 can be reformulated into the following Markov decision process:

**Problem 2 Markov Decision Process Reformulation:**

\[
\text{mse}_{\text{opt}} = \inf_{\pi \in \Pi_{\text{recent}}} \lim_{K \to \infty} \sup \mathbb{E} \left[ \sum_{k=1}^{K} \frac{\frac{1}{6}(X_{S_{k+1}} - X_{S_k})^4}{\sum_{k=1}^{K} \mathbb{E}[(S_{k+1} - S_k)]} + D \right],
\]

s.t. \( \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \mathbb{E}[(S_{k+1} - S_k)] \geq \frac{1}{f_{\text{max}}}. \) (4b)

The proof is provided in Appendix H of the supplementary material.

According to [7, Theorem 1], there exists a stationary policy \( \pi^* \) that selects the waiting time \( W_k \) using a conditional probability distribution given the recent \( \mathcal{I}_k \) that achieves \( \text{mse}_{\text{opt}}. \) Next, we will reveal the sufficient conditions of such policy for designing the online algorithm.

**B. Designing \( \pi^* \) With Known \( \mathbb{P}_D \)**

Let \( \Pi_{\text{cons}} \triangleq \{ \pi \in \Pi_{\text{recent}} | \lim_{T \to \infty} \mathbb{E} \left[ \frac{i(T)}{T} \right] \leq f_{\text{max}} \} \) denote the set of policies that satisfy the sampling frequency constraint. Since \( \pi^* \) achieves the minimum expected time-average MSE among \( \Pi_{\text{cons}}, \) we have:

\[
\lim_{K \to \infty} \sum_{k=1}^{K} \frac{\frac{1}{6}(X_{S_{k+1}} - X_{S_k})^4}{\sum_{k=1}^{K} \mathbb{E}[D_k + W_k]} \geq \mathcal{E}^\pi - D, \quad \pi \in \Pi_{\text{cons}}.
\]

(5)

For simplicity, denote \( \gamma^* := \mathcal{E}^\pi - D, \) which is the average cost of the MDP when the optimum policy \( \pi^* \) is used, i.e., \( \gamma^* = \lim_{K \to \infty} \frac{\sum_{k=1}^{K} \mathbb{E}[(X_{S_{k+1}} - X_{S_k})^4]}{\sum_{k=1}^{K} \mathbb{E}[D_k + W_k]} \). Because \( \frac{1}{K} \sum_{k=1}^{K} \mathbb{E}[D_k + W_k] > 0, \) for any policy \( \pi \in \Pi_{\text{cons}}, \) inequality (5) can be rewritten as:

\[
\theta_{\pi}(\gamma^*) := \liminf_{K \to \infty} \left( \frac{1}{K} \sum_{k=1}^{K} \mathbb{E} \left[ \frac{1}{6}(X_{S_{k+1}} - X_{S_k})^4 \right] - \gamma^* \cdot \frac{1}{K} \sum_{k=1}^{K} \mathbb{E}[D_k + W_k] \right) \geq 0.
\]

(6)

Inequality (6) takes the minimum value 0 if and only if policy \( \pi \) is optimum. Therefore, if the ratio \( \gamma^* \) is known, an optimum policy \( \pi^* \) can be obtained by solving the following functional optimization:

**Problem 3 Functional Optimization Problem:**

\[
\text{mse}_{\text{opt}} = \inf_{\pi \in \Pi_{\text{recent}}} \lim_{K \to \infty} \sup \mathbb{E} \left[ \frac{1}{K} \sum_{k=1}^{K} \mathbb{E} \left[ \frac{1}{6}(X_{S_{k+1}} - X_{S_k})^4 \right] \right]
\]

s.t. \( \liminf_{K \to \infty} \mathbb{E} \left[ \frac{1}{K} \sum_{k=1}^{K} (D_k + W_k) \right] \geq \frac{1}{f_{\text{max}}}. \) (7b)

To solve Problem 3, we can take the Lagrangian duality of the constraint (7b) with a dual variable \( \nu \) and obtain the Lagrange function \( \mathcal{L}(\pi, \gamma, \nu): \)

\[
\mathcal{L}(\pi, \gamma, \nu) \triangleq \limsup_{K \to \infty} \left( \frac{1}{K} \sum_{k=1}^{K} \mathbb{E} \left[ \frac{1}{6}(X_{S_{k+1}} - X_{S_k})^4 \right] \right)
\]

\[
- \left( \gamma + \nu \right) \cdot \liminf_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \mathbb{E} \left[ (S_{k+1} - S_k) \right] + \nu \cdot \frac{1}{f_{\text{max}}}. \]

(8)

We say that a stationary policy \( \pi \) has a threshold structure, if the waiting time \( W_k \) is determined by:

\[
W_k = \inf \{ w \geq 0 | X_{S_k + D_k + w} - X_{S_k} \geq \tau \}.
\]

(9)
Let $Z_t$ be a Wiener process starting from $t = 0$. Let $D$ be the random transmission delay following distribution $P_D$ and the value of the Wiener process at the random time $D$ is denoted by $Z_D$. Using the threshold policy (9), the expected frame-length $L_k := D_k + W_k$ and $1/E(X_{S_k+1} - X_{S_k})^4$ has the following properties:

Lemma 2: [7, Corollary 1 Restated]

$$
E[L_k] = E\left[\max\{\tau^2, Z_D^2\}\right],
$$

(10a)

$$
E\left[\frac{1}{6}(X_{S_k+1} - X_{S_k})^4\right] = \frac{1}{6}E\left[\max\{\tau^2, Z_D^2\}\right].
$$

(10b)

As is revealed by [7], the optimum policy $\pi^*$ has a threshold structure as in equation (9). To design an off-line algorithm that can learn the updating threshold $\tau^*$ of $\pi^*$, we then reveal the necessary conditions that $\tau^*$ should satisfy. With slightly abuse of notations, let $L(\tau, \gamma, \nu)$ denote the expected value of the Lagrange function $L(\pi, \gamma, \nu)$ when a stationary policy $\pi$ with threshold $\tau$ is used. According to Lemma 2, $L(\tau, \gamma, \nu)$ can be computed as follows:

$$
L(\tau, \gamma, \nu) = E\left[\frac{1}{6}\max\{\tau^2, Z_D^2\}\right] - (\gamma + \nu)E[\max\{\tau^2, Z_D^2\}] + \nu - \frac{1}{f_{max}}.
$$

(11)

Condition 1: [7, Theorem 6 Restated] Let $\tau(\gamma, \nu)$ be the optimum sampling threshold that minimizes function $L(\tau, \gamma, \nu)$, which can be computed as follows:

$$
\tau(\gamma, \nu) := \arg inf_{\tau \geq 0} L(\tau, \gamma, \nu) = \sqrt{3(\gamma + \nu)}.
$$

(12)

Recall that for any policy $\pi \in \Pi_{cons}$ with threshold $\tau$, inequality (5) implies

$$
\theta_\pi(\gamma^*) = \frac{1}{6}E[\max\{\tau^2, Z_D^2\}] - \gamma^*E[\max\{\tau^2, Z_D^2\}] \geq 0.
$$

(13)

According to (12), inequality (13) holds with equality if and only if $\pi^*$ with threshold $\tau^* = \sqrt{3(\gamma^* + \nu^*)}$ is used.

Condition 2: [7, Eq. (123, 125)]

$$
\nu^* \left( E\left[\max\{3(\gamma^* + \nu^*), Z_D^2\}\right] - \frac{1}{f_{max}} \right) = 0, \nu^* \geq 0.
$$

(14)

Adding the Complete Slackness (CS) condition (14) on both sides of (13), the necessary condition for $\gamma^*$ then becomes:

$$
\gamma^*_d(\gamma^*) = \theta_\pi(\gamma^*) = 0,
$$

(15)

where function $\gamma_d(\gamma) := E[g_d(\gamma; Z_D)]$ is the expectation of function $g_d(\gamma; Z_D)$ defined as follows:

$$
g_d(\gamma; Z_D) := \frac{1}{6}\max\{3(\gamma + \nu), Z_D^2\}^2 - \gamma \max\{3(\gamma + \nu), Z_D^2\}.
$$

(16)

As is shown by [7, Theorem 7], the duality gap between $E_{\pi^*}$ and $\sup_{\nu \geq 0} \inf_{\pi} L(\pi, \gamma^*, \nu)$ is zero, and (15) becomes a necessary and sufficient condition.

C. An Online Algorithm $\pi_{online}$

The goal is to learn the optimum sampling threshold that resolves Problem 3 while satisfying the sampling frequency constraint in the absence of the delay distribution $P_D$. We use a virtual queue $U_k$ to record the sampling frequency violation up to the $(k + 1)$-th sample. Then Problem 3 can be plugged into the Drift-Plus-Penalty framework [27] for network utility maximization. The virtual queue $U_{k+1} = (U_k + \frac{1}{f_{max}} - (D_k + W_k))^+$ evolves like a queueing system, where the arrivals $1/f_{max}$ is fixed as a constant, and $S_{k+1} - S_k$ is the departures. Let $J_k := \frac{1}{2}U_k$ denote the Lyapunov function, and the Lyapunov drift is denoted by $\Delta J_k := E[J_{k+1} - J_k|U_k]$. The stopping time $S_{k+1}$ is selected to minimize the average cost (7a) while keeping the virtual queue stable, and when $\gamma = \gamma_k$, the Drift-Plus-Penalty optimization goal can be written as:

$$
\Delta J_k + V \left( \frac{1}{6}E[(X_{S_k+1} - X_{S_k})^4] - \gamma_kE[(S_{k+1} - S_k)]\right) \\
\leq U_kE \left[ \frac{1}{f_{max}} - (S_{k+1} - S_k) \right] + V \left( \frac{1}{6}E[(X_{S_k+1} - X_{S_k})^4] - \gamma_kE[(S_{k+1} - S_k)] \right) \\
+ \frac{1}{f_{max}^2} + E[(S_{k+1} - S_k)^2].
$$

(17)

In the drift-plus-penalty optimization framework, $\frac{1}{f_{max}} + E[(S_{k+1} - S_k)^2]$ is usually upper bounded by the second order moment of arrival and departure rate upper bound. Then, scheduling decision are chosen to minimize the first two terms in (17), which is equivalent to minimizing:

$$
\frac{1}{6}E[(X_{S_k+1} - X_{S_k})^4] - (\gamma_k + \frac{1}{V}U_k)E[(X_{S_k+1} - X_{S_k})^2] \\
\geq \frac{1}{6}E\left[ (X_{S_k+1} - X_{S_k})^2 - 3(\gamma_k + \frac{1}{V}U_k)^2 \right],
$$

(18)

where equation (a) is obtained by Wald’s Lemma $E[S_{k+1} - S_k] = E[(X_{S_k+1} - X_{S_k})^2]$. The optimum solution that satisfies $S_{k+1} \geq S_k$ and minimize (18) is then

$$
S_{k+1} = \arg \min \{ t \geq S_k + D_k | X_{S_k+t} - X_{S_k} \} \\
\geq \sqrt{3(\gamma_k + \frac{1}{V}U_k)},
$$

(19)

Based on this observation, we then derive our sampling and learning policy as follows:

The algorithm is initialized by selecting $\gamma_1 = 0$ and $U_1 = 0$. In each frame $k$, the sampling and updating rules are as follows:

1. Sampling: The waiting time $W_{k+1}$ is selected to minimize the Lagrange function (8) following equation, and according to the statement after equation (19). Therefore, the waiting time $W_k$ is selected by:

$$
W_k = \inf\{ w \geq 0 | X_{S_k+D_k+w} - X_{S_k} \geq \sqrt{3(\gamma_k + \frac{1}{V}U_k)} \}.
$$

(20)
2. Update $\gamma_k$: To search for the root $\gamma > 0$ of equation $\overline{y}_{\nu_k}(\gamma) = 0$, we update $\gamma_k$ through the Robbins-Monro algorithm [28]. In each frame $k$, we are given an i.i.d sample $\delta X_k = X_{S_k + D_k} - X_{S_k} \sim Z_D$, and the Robbins-Monro algorithm operates by:

$$\gamma_{k+1} = (\gamma_k + \eta_k Y_k)^+, \quad (21)$$

where $Y_k = g_{\nu_k}(\gamma_k; \delta X_k)$ and function $g_{\nu}(.)$ is defined in (16). Recall that $\overline{D}_b$ is a non-zero lower bound of the average delay, the step-size $\{\eta_k\}$ is selected by:

$$\eta_k = \frac{1}{\overline{D}_b(2 + k^\alpha)} \alpha \in (0.5, 1]. \quad (22)$$

3. Update $U_k$: To guarantee that the sampling frequency constraint is not violated, we update the violation $U_k$ up to the end of frame $k$ by:

$$U_{k+1} = \left(U_k + \left(\frac{1}{f_{\max}} - (D_k + W_k)\right)\right)^+. \quad (23)$$

Remark 1: Another view to understand the usage of $U_k$ that the sequence $\nu_k := \frac{1}{U_k}$ is used to approximate $\nu^*$.

D. Theoretical Analysis

We analyze the convergence and optimality of algorithm $\pi_{\text{online}}$. We assume there is no sampling frequency constraint, i.e., $f_{\max} = \infty$ and make the following assumption on distribution $\mathbb{P}$:

Assumption 1: The fourth order moment of the transmission delay is upper bounded by $B$, i.e.,

$$\mathbb{E}[D^4] \leq B < \infty.$$  

The convergence behavior of the optimum threshold $3\gamma^*$ and the MSE performance are manifested in the following theorems:

Theorem 1: The proposed algorithm learns the optimum parameter $\gamma^*$ almost surely, i.e.,

$$\lim_{k \to \infty} \gamma_k = \gamma^*, \quad \text{w.p.}1. \quad (24)$$

The proof of Theorem 1 is obtained by the ODE method in [24, Chapter 5] and is provided in Appendix B.

Theorem 2: The second moment of $(\gamma_k - \gamma^*)$ satisfies:

$$\sup_k \frac{1}{\eta_k} \mathbb{E}[(\gamma_k - \gamma^*)^2] < \infty. \quad (25)$$

Specifically, if $\alpha = 1$ and $\eta_k = \frac{1}{\overline{D}_b(2 + k^\alpha)}$, then the mean square error decays with rate $\mathbb{E}[(\gamma_k - \gamma^*)^2] = \mathcal{O}(1/k)$. \n
One challenge in the proof of Theorem 2 is that $\gamma_k$ is unbounded and the second moment of $Y_k$ is unbounded. We notice that $Y_k$ could become very large when $\gamma_k$ is much larger than the true value $\gamma^*$, but the truncation of $(\gamma_k + \eta_k Y_k)^+ \to$ non-negative part actually prevents the actual update $(\gamma_k + \eta_k Y_k)^+ - \gamma_k$ from becoming too large. Based on this observation, we adopt a method from the heavy-traffic analysis by introducing the unused rate $\chi_k := (-(\gamma_k + \eta_k Y_k))^+$, then prove that the variance of the amount of the actual updating $\eta_k Y_k + \chi_k$ is finite. Detailed proofs are provided in Appendix C.

Theorem 3: The average MSE under policy $\pi_{\text{online}}$ converges to $\overline{\mathcal{E}}_{\pi^*}$, almost surely, i.e.,

$$\lim_{k \to \infty} \mathbb{E}\left[\frac{1}{S_{k+1}} \int_{t=0}^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] = \overline{\mathcal{E}}_{\pi^*}, \quad \text{w.p.}1. \quad (26)$$

With the mean-square convergence of $\gamma_k$, the proof of Theorem 3 is a direct application of the perturbed ODE method [24] and is provided Appendix D.

By using Theorem 2 and Theorem 3, we can upper bound the growth rate of the cumulative MSE optimality gap in the following corollary:

Corollary 1: If $\alpha = 1$, then the growth rate of the cumulative MSE optimality gap up to the k-th sample can be bounded as follows:

$$\left(\mathbb{E}\left[\int_{t=0}^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] - \overline{\mathcal{E}}_{\pi^*}\mathbb{E}[S_k]\right) = O \left(\ln k\right). \quad (27)$$

The proof of Corollary 1 is provided in Appendix I of the supplementary material.

Theorem 4: For any distribution $\mathbb{P}$, let $\pi^*(\mathbb{P})$ denote the MSE minimum sampling policy when the delay $D \sim \mathbb{P}$. The threshold obtained by solving equation (15) is denoted by $\gamma^*(\mathbb{P})$. After $k$-samples are taken, the minimax estimation error $\gamma^*(\mathbb{P})$ is lower bounded by:

$$\inf_{\gamma} \sup_{\mathbb{P}} \mathbb{E}[(\hat{\gamma} - \gamma^*(\mathbb{P}))^2] = \Omega(1/k). \quad (28)$$

Let $p_{w}(\mathbb{P}) := \Pr(Z_D^2 \leq 3\gamma^*(\mathbb{P})D \sim \mathbb{P})$ denote the probability of waiting by using policy $\pi^*(\mathbb{P})$. Specifically, let $p_{w, \text{uni}} := \Pr(Z_D^2 \leq 3\gamma^*_\text{uni} D \sim \text{Uni}(0, 1))$. Let $\Pi_k$ denote the set of policies which the sampling decision $S_k$ is made based on historical information $H_{k-1}$. We have the following result:

$$\inf_{\gamma} \sup_{\mathbb{P}} \mathbb{E}\left[\int_{t=0}^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] - \overline{\mathcal{E}}_{\pi^*(\mathbb{P})}\mathbb{E}[S_k] \right) \geq \frac{1}{4} \left(\frac{1}{24} - \delta\right) (\frac{1}{\overline{D}_b})^2 \sum_{k'=1}^{k} \frac{1}{k'} = \Omega \left(\ln k\right). \quad (29)$$

As the transmission delay $D$ considered in the paper does not belong to a specific family and could be quite general, obtaining a point-wise converse bound on $\mathbb{E}[(\hat{\gamma} - \gamma^*(\mathbb{P}))^2]$ for each distribution $\mathbb{P}$ is impossible. As an alternative, a minimax risk bound $\mathbb{E}[(\hat{\gamma} - \gamma^*(\mathbb{P}))^2]$ over a general distribution set $\mathcal{P}$ can be obtained using Le Cam’s two point method for non-parametric estimation [29]. The core idea is to construct two distributions $\mathcal{P}_1, \mathcal{P}_2$, whose $\ell_1$ distance $\|\mathbb{P}_1 - \mathbb{P}_2\|_1$ can be upper bounded by a constant, but $(\gamma^*(\mathbb{P}_1) - \gamma^*(\mathbb{P}_2))^2 \geq \Omega(1/k)$ is difficult to distinguish. Such a construction is still challenging because $\gamma^*(\mathbb{P})$ cannot be obtained in closed form even for the simplest distribution families such as the delta distribution or exponential distribution. Notice that the estimation error of $\gamma^*$ is closely related to the estimation error $\overline{\mathbb{P}}_{\nu}(\cdot)$ at a given point. Therefore, the construction of $\mathcal{P}_1$ and $\mathcal{P}_2$ for obtaining the converse bound of Hölder smooth functions [29, Chapter 2] are adopted. The proof of inequality (29) is a direct application of the minimax estimation error (28). Detailed proof of Theorem 4 is provided in Appendix E.
IV. SIMULATION RESULTS

In this section, we provide simulation results to verify the theoretic findings and illustrate the performance of our proposed algorithms. Recall that \( M_t = \{(S_j, D_j, X_{S_j})\}_{j=1}^{i(t)} \) is the sampling time-stamps, transmission delay and values of previous samples up to time \( t \) and \( i(t) \) is the index of the freshest sample received by time \( t \). In the following simulations, the estimation \( \hat{X}_t \) is obtained by the MMSE estimator

\[
\hat{X}_t = \mathbb{E}[X_t | M_t] = X_{S(i(t))}, \quad \text{i.e., equation (1)}
\]

regardless of the sampling policy.

We notice that the MSE minimization problem is closely related to the AoI minimization problem, where the AoI at time \( t \), denoted by \( A(t) = t - S_{i(t)} \), is the index of the freshest sample received by time \( t \). For signal-ignorant sampling policies (i.e., the sensor cannot always observe the time-varying process), according to the analysis in [3, Section IV-B], policies that minimize the average AoI achieves the minimum MSE. Therefore, we choose both offline and online AoI minimization policies (\( \pi^{\star}_{\text{AoI}} \) from [3], \( \pi_{i(t)} \) from [11]) for comparison. To show the convergence of online learning algorithm, we plotted the average MSE performance of the optimum off-line algorithm \( \pi^* \) from [7].

The transmission delay follows the log-normal distribution parameterized by \( \mu \) and \( \sigma \) such that the density function of the probability measure \( \mathbb{P}_D \) is:

\[
p(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right).
\]

In simulations, we set \( \mu = 0.8 \) and \( \sigma = 1.2 \), the expected time-averaged MSE is computed by taking the average of 20 runs. Fig. 3 depicts the time-average MSE performance up to the \( k \)-th frame of different sampling policies. The evolution of \( \{\gamma_k\} \) and the MSE regret

\[
\Delta_k := \mathbb{E}\left[\int_0^{S_{k+1}} (X_t - \hat{X}_t)^2 dt\right] - \mathbb{E}_\pi^* \mathbb{E}[S_{k+1}]
\]

are depicted in Fig. 4. From Fig. 3, with \( 5 \times 10^4 \) samples, the time averaged MSE is almost the same as using the optimum policy. From Fig. 5, the MSE regret is almost a...
logarithm function of frame $k$. The asymptotic MSE behaviour is consistent with the convergence results in Theorem 3 and Corollary 1.

When there is a sampling frequency constraint, the average MSE and the average sampling interval achieved by policy $\pi_{\text{online}}$ are depicted in Fig. 6 and Fig. 7, respectively. We set $f_{\text{max}} = \frac{1}{20}f_{\text{max}}$. From these figures, one can observe that the average MSE of $\pi_{\text{online}}$ is close to the optimum MSE $\pi_{\text{opt}}$ and the sampling frequency can be satisfied. In addition, by choosing a larger $V$, a smaller MSE performance can be achieved, whereas a larger number of iterations are needed to meet the sampling frequency constraint.

V. CONCLUSION

In this work, we studied the problem of sampling a Wiener process for remote estimation over a channel with unknown delay statistics. By reformulating the MSE minimization problem as a renewal-reward process, we proposed an online sampling algorithm that can adaptively learn the optimum algorithm as the number of samples grows. We showed that the average MSE obtained by the proposed algorithm converges to the minimum MSE almost surely, and the cumulative MSE has an order of $O(\ln k)$, where $k$ is the number of samples. We then prove that the cumulative MSE regret of any algorithm is at least $\Omega(\ln k)$. Numerical simulation results validate the convergence behaviors of the proposed algorithm. Extending the above signal-aware online learning method to error-prone and two-way delay channels [30], [31] will be our future work.

APPENDIX

We provide the proofs for our main results (Theorem 1-4) in Appendix B-F. The proof of the main results requires some additional lemmas and notations, which are organized in Appendix A. Due to page limitations, the proofs of those additional lemmas are provided in the supplementary materials.

APPENDIX A

NOTATIONS AND PRELIMINARY LEMMAS

In Table I, we summarize the notations used in the following proofs. Throughout the proofs, we use $N_1, N_2, \ldots$ to denote absolute constants and $C_1(\cdot), C_2(\cdot)$ to denote polynomials with finite order. For ease of exposition, the specific values and expressions of the constants and functions may vary across different contexts.

**Lemma 3:** Let $M := \mathbb{E}[D^2]$, the optimum ratio $\gamma^*$ is upper and lower bounded by:

\[ \frac{1}{6} \leq \gamma^* \leq \frac{1}{2} \frac{M + 2D \frac{1}{f_{\text{max}}} + \frac{1}{f_{\text{max}}}}{D + \frac{1}{f_{\text{max}}}}. \]  

(30)

The proof is provided in Appendix F of the supplementary material.

**Lemma 4:** For threshold $\gamma < \infty$, the first, second and fourth order moments of the stopping time $\tau_\gamma$ are bounded, i.e.,

\[ \mathbb{E}[t_\gamma] \leq 3\gamma + D, \]  

(31a)

\[ \mathbb{E}[t_\gamma^2] \leq \frac{10}{3} (3\gamma^2 + 3\sqrt{B}), \]  

(31b)

\[ \mathbb{E}[t_\gamma^3] < 4^3 (3\gamma^4 + 105B) < \infty. \]  

(31c)

The proof of Lemma 4 is provided in Appendix K of the supplementary material.

**Lemma 5:** Function $\gamma_0(\gamma) = q(\gamma) - \gamma l(\gamma)$ and has the following properties:

(i) $\gamma_0(\gamma)$ is concave and monotonically decreasing. The second order derivative $-3 \leq \gamma_0''(\gamma) \leq 0$.

(ii) $\gamma_0(\gamma^*) = 0$

(iii) For $\gamma \neq \gamma^*$, $(\gamma - \gamma^*)\gamma_0(\gamma) \leq -l(\gamma^*)(\gamma - \gamma^*)^2 \leq 0$.

The proof of Lemma 5 is provided in Appendix L of the supplementary material.

**Corollary 2:** For each $\gamma_k < \infty$, if the fourth order moment of the delay satisfies $\mathbb{E}[D^4] < B < \infty$, given historical transmission $\mathcal{H}_{k-1}$, the conditional second order moment of the cumulative error in frame $E_k = \int_{S_k}^X (X_t - \hat{X}_t)^2 dt$ can be bounded as follows:

\[ \mathbb{E}_k[E_k^2] = 3(X_{S_k} - X_{S_k-1})^2 \sqrt{B} \]
\[ + 12C_1(\gamma_k, B)(X_{S_k} - X_{S_k-1})^2 \]
\[ + 3C_2(\gamma_k, B) < \infty, \]  

(32)

where $C_1$ and $C_2$ are fourth order polynomials of $\gamma$.

The proof of Corollary 2 is provided in Appendix N of the supplementary material.

APPENDIX B

PROOF OF THEOREM 1

To show that $\gamma_k$ converges to $\gamma^*$ almost surely, we use the sufficient condition for stochastic approximation with projection [24, p. 170, Theorem 1.2]. Recall that the step-size $\eta_k = \frac{1}{k}$. Define $t_0 = 0$ and denote the sum of step-sizes up to frame $k$ by $t_k := \sum_{i=1}^k \eta_i$. For $t_0 \geq 0$, let $m(t)$ be the unique $k \in \mathbb{N}^*$ so that $t_k \leq t < t_{k+1}$. Without a sampling constraint, $\nu_k \equiv 0, \forall k$. Then the update rule for $\gamma_k$ from equation (21)
can be rewritten in the following recursive form:
\[
\gamma_{k+1} = (\gamma_k + \eta_k (g_0(\gamma_k) + \delta M_k)) \geq \gamma_k, \tag{33}
\]
where recall that \( Y_k = g_0(\gamma_k; \delta X_k)^{\frac{1}{6}} = \max\{3\gamma_k, \delta X_k^2\}^2 - \gamma_k \max\{3\gamma_k, \delta X_k^2\} \) and \( \delta M_k \) is the difference between realization and the conditional expectation \( E_k[Y_k] = g_0(\gamma_k) \).

Notice that the difference \( \delta M_k := Y_k - E_k[Y_k] \) depends only on the transmission delay and the Wiener process evolution \( (X_t - X_{S_k}) \) in frame \( k \) and \( \gamma_k \), which can be predictable given \( \gamma_{k-1} \) and is therefore a martingale sequence. We then show that \( \{Y_k\}, \{\delta M_k\} \) have the following properties:

(1.1) For each constant \( N < \infty \), \( \sup_k E[\|Y_k\|_{1(\gamma_k \leq N)}] \) is bounded, i.e.,
\[
\sup_k E[\|Y_k\|_{1(\gamma_k \leq N)}] = \sup_k E[\left(\frac{1}{6} \max\{3\gamma_k, \delta X_k^2\}^2 \cdot I_{(\gamma_k \leq N)}\right)]
\leq \sup_k E[\left(\frac{1}{6} \max\{3\gamma_k, \delta X_k^2\}^2 \cdot I_{(\gamma_k \leq N)}\right)]
\leq \frac{1}{6} (9N^2 + E[Z_{D_{k}}^2] + N \cdot (3N + E(Z_{D_{k}}^2)) \leq \infty. \tag{34}
\]

where inequality (a) is because \( E[Z_{D_{k}}^2] = 3E[D^2] \leq 3\sqrt{E[D^2]} < \infty \) and \( E[Z_{D_{k}}^2] = E[D] < \infty \).

(1.2) Function \( Y_k = g(\gamma_k; \delta x) \) is continuous in \( \gamma_k \) for each \( \delta x \).

(1.3) The martingale sequence \( \delta M_k \|_{1(\gamma_k \leq N)} \) can be bounded as follows:
\[
\text{Var}[\delta M_k \|_{1(\gamma_k \leq N)}] \leq E[Y_k^2 \|_{1(\gamma_k \leq N)}]
\leq E \left[ \left(\frac{1}{6} \max\{3\gamma_k, \delta X_k^2\}^2 - \gamma_k \max\{3\gamma_k, \delta X_k^2\} \right) \right] \leq \infty. \tag{35}
\]

where inequality (b) is because \( E[a-b] \leq E[a^2 + b^2] \); inequality (c) is because \( \delta X_t \sim Z_{D}\) is a Wiener process starting from \( t = 0 \) and therefore, \( E[Z_{D_k}^2] = 105E[D^2] \leq 105B \).

Since sequence \( \delta M_k \|_{1(\gamma_k \leq N)} \) has mean zero. Its value only depends on \( \gamma_k \) and the Wiener process evolution in frame \( k \). The correlation \( E[\delta M_k \|_{(\gamma_k \leq N)} \cdot \delta M_j \|_{(\gamma_j \leq N)}] = 0, \forall i \neq j \).

As the variance of \( \delta M_k \|_{1(\gamma_k \leq N)} \) is bounded in inequality (35), the stepsizes \( \eta_k \) satisfies \( \sum_{k=1}^{\infty} \frac{1}{2\eta_k^2} < 2\frac{1}{\alpha_0} \), according to [24, Chapter 5, Eq. (5.3.18)], for each \( \mu > 0 \) we have
\[
\lim \Pr \left( \sup_{k \geq 1} \sum_{i=m(T)}^{m(T+1)-1} \|e\delta M_k \|_{1(\gamma_k \leq N)} \geq \mu \right) = 0. \tag{36}
\]

Let \( \gamma_k(\omega) \) be the value of ratio \( \gamma \) on sample path \( \omega \). Recall that the stepsizes \( \{\eta_k\} \) selected in (22) satisfies \( \sum_{k=1}^{\infty} \frac{1}{\eta_k^2} < \infty \). According to [24, p.170, Theorem 1.2], with probability 1, the limit \( \lim_{k \to \infty} \theta_k(\omega) \) are trajectories of the following ordinary differential equation (ODE), i.e.,
\[
\dot{\gamma} = g_0(\gamma). \tag{37}
\]

The next step is to show the solution of the ODE in equation (37) converges to \( \gamma^* \) as time diverges. Equation (15) implies \( g_0(\gamma^*) = 0 \) and therefore, \( \gamma^* \) is an equilibrium point of ODE (37). To show that the ODE is stationary at \( \gamma = \gamma^* \), we use the Lyapunov approach by defining function \( V(\gamma) := \frac{1}{2} (\gamma - \gamma^*)^2 \), whose time derivative \( \dot{V} = \frac{d}{dt} V(\gamma(t)) \) can be computed by:
\[
\dot{V} = (\gamma - \gamma^*) \dot{\gamma} = (\gamma - \gamma^*) g_0(\gamma). \tag{38}
\]

According to Lemma 5-(iii), \( \dot{V} = (\gamma - \gamma^*) g_0(\gamma) \leq 0 \), the stability of \( \gamma^* \) is verified through Lyapunov theorem.

### APPENDIX C

#### PROOF OF THEOREM 2

The analysis of the convergence rate is obtained through Lyapunov analysis, where the Lyapunov function is denoted by \( V(\gamma) := \frac{1}{2} (\gamma - \gamma^*)^2 \). The proof is divided into two steps: first we will upper bound the Lyapunov drift for each \( \gamma_k \) by showing the following equation holds:
\[
E_k[V(\gamma_{k+1})] - V(\gamma_k) \leq -\eta_k D_{ib} V(\gamma_k) + O(\eta_k^2 N_1). \tag{39}
\]

Then, based on (39), we then compute \( E[V(\gamma_k)] \) directly.

**Step 1: Bounding the Lyapunov Drift:** The analysis is divided into two cases: For \( \gamma_k \leq 3\gamma^* \), inequality (39) can be verified easily (Case 1); For \( \gamma_k \geq 3\gamma^* \) we will first establish the relationship between \( E_k[V(\gamma_{k+1})] - V(\gamma_k) \) and \( Var[Y_k] \), then upper bound \( \text{Var}[Y_k] \) using the fact that \( Z_{D_k}^2 \) is sub-Gaussian when \( D \) is fourth order bounded (Case 2). Detailed proofs are as follows:

**Case 1:** If \( \gamma_k \leq 3\gamma^* \), we have:
\[
E_k[V(\gamma_{k+1})] - V(\gamma_k)
= E_k \left[ \frac{1}{2} ((\gamma_k + \eta_k Y_k)^+ - \gamma^*)^2 \right] - \frac{1}{2} (\gamma_k - \gamma^*)^2 \leq E_k \left[ \frac{1}{2} (\gamma_k - \gamma^* + \eta_k Y_k)^2 - \frac{1}{2} (\gamma_k - \gamma^*)^2 \right]
(a) \gamma_k - \gamma^* \text{ is bounded by inequality (35),}
+ \eta_k^2 E_k \left[ \frac{1}{6} \max\{3\gamma_k, \delta X_k^2\}^2 \right] \leq 0.
\]

where equality (a) is because \( E_k[Y_k] = E_k[g_0(\gamma_k; \delta X_k)] = g_0(\gamma_k) \); inequality (b) is obtained because according to Lemma 5-(iii), \( \gamma_k - \gamma^* \leq g_0(\gamma_k) \leq -l(\gamma^*)(\gamma_k - \gamma^*)^2 = -2l(\gamma^*)V(\gamma_k) \) and the assumption that \( \gamma_k \leq 3\gamma^* \).

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Case 2: If $\gamma_k \geq 3\gamma^*$, $\gamma_{k+1} = (\gamma_k + \eta_k Y_k)^+$ is truncated into the non-negative real part. We can view the evolution of $\gamma_k$ as a queueing system, where the queue $\gamma_k$ is non-negative, $\eta_k Y_k$ is the arrival rate minus the service rate. Therefore, it is natural to introduce the “unused rate” from [32], which is denoted by $\chi_k := (- (\gamma_k + \eta_k Y_k))^+$. If $\chi_k = 0$, $(\gamma_k + \eta_k Y_k)\chi_k = 0 = -\chi_k$ and if $\chi_k \geq 0$, $\gamma_k + \eta_k Y_k = -\chi_k$, therefore

$$
(\gamma_k + \eta_k Y_k)\chi_k = -\chi_k^2. \tag{41}
$$

Since $\gamma_k + \eta_k Y_k + \chi_k \geq 0$, we have:

$$
-\mathbb{E}_k[\gamma_k + \eta_k Y_k] \leq \mathbb{E}_k[\chi_k]. \tag{42}
$$

We can then upper bound $\mathbb{E}_k[V(\gamma_{k+1}) - V(\gamma_k)]$ by:

$$
\mathbb{E}_k[V(\gamma_{k+1}) - V(\gamma_k)] = \mathbb{E}_k\left[\frac{1}{2}(\gamma_k - \gamma^* + \eta_k Y_k + \chi_k)^2 - \frac{1}{2}(\gamma_k - \gamma^*)^2\right]
= \mathbb{E}_k\left[\frac{1}{2}(\gamma_k - \gamma^* + \eta_k Y_k)^2 - \frac{1}{2}(\gamma_k - \gamma^*)^2 + \frac{1}{2}\chi_k^2 + (\gamma_k + \eta_k Y_k)\chi_k - \gamma^* \gamma_k\right]
\leq \mathbb{E}_k\left[\frac{1}{2}(\gamma_k - \gamma^* + \eta_k Y_k)^2 - \frac{1}{2}(\gamma_k - \gamma^*)^2 + \frac{1}{2}\eta_k^2 \text{Var}[Y_k]ight], \tag{43}
$$

where equality (e) is because equation (41); inequality (d) is obtained because $\mathbb{E}_k[\chi_k^2] \geq \mathbb{E}_k[|\chi_k|^2] \geq 0$;

To upper bound (39), we then further divide the analysis into two cases:

Case 2(a): If $\mathbb{E}_k[\gamma_k + \eta_k Y_k] \leq \gamma^*$, we then have $\mathbb{E}_k[\gamma_k - \gamma^* + \eta_k Y_k] \leq 0$. According to (42), $| - \mathbb{E}_k[\chi_k] - \gamma^* | \geq | \gamma_k - \gamma^* + \eta_k \mathbb{E}_k[Y_k] |$. Therefore, inequality (43) can be upper bounded by:

$$
\mathbb{E}_k[V(\gamma_{k+1}) - V(\gamma_k)]
\leq -\frac{1}{2} (\gamma_k - \gamma^*)^2 + \frac{1}{2} \gamma^* + \frac{1}{2}\eta_k^2 \text{Var}[Y_k]
\leq -\frac{1}{4} (\gamma_k - \gamma^*)^2 + \frac{1}{2}\eta_k^2 \text{Var}[Y_k]
\leq -\eta_k D_h \mathbb{E}_k[V(\gamma_k)] + \frac{1}{2}\eta_k^2 \text{Var}[Y_k], \tag{44}
$$

where inequality (e) is obtained because $\frac{1}{4}(\gamma_k - \gamma^*)^2 \geq (\gamma^*)^2 \geq \frac{1}{2}(\gamma^*)^2$ because in Case 2 we have $\gamma_k \geq 3\gamma^*$; inequality (f) is obtained because $\eta_k D_h \leq \frac{1}{2}$ by the step-size selection rule in equation (22).

Case 2(b): If $\mathbb{E}_k[\gamma_k + \eta_k Y_k] \geq \gamma^*$, considering that $\mathbb{E}_k[Y_k] = \mathbb{E}_0[(\gamma_k) < 0 \text{ for } \gamma_k \geq \gamma^*$, we have $0 > \mathbb{E}_k[\eta_k Y_k] \geq 0$.

$$(\gamma_k - \gamma^*) \chi_k = -\chi_k^2.$$

Inequality (43) can be bounded by:

$$
\mathbb{E}_k[V(\gamma_{k+1}) - V(\gamma_k)]
\leq \frac{1}{2} (\gamma_k - \gamma^*) \chi_k - \gamma^* + \eta_k \mathbb{E}_k[Y_k] - \frac{1}{2} (\gamma_k - \gamma^*)^2 + \frac{1}{2}\eta_k^2 \text{Var}[Y_k]
\leq \frac{1}{2} \eta_k (\gamma_k - \gamma^*) \mathbb{E}_0(\gamma_k) + \frac{1}{2}\eta_k^2 \text{Var}[Y_k]
\leq -\frac{1}{2} \eta_k (\gamma_k - \gamma^*)^2 + \frac{1}{2}\eta_k^2 \text{Var}[Y_k]
\leq -\eta_k (\gamma_k - \gamma^*) \mathbb{E}_k[V(\gamma_k)] + \frac{1}{2}\eta_k^2 \text{Var}[Y_k], \tag{45}
$$

where equality (g) is because $-\mathbb{E}_k[\chi_k] - \gamma^* \geq (\gamma^*)^2$ and $(\gamma_k - \gamma^* + \eta_k \mathbb{E}_k[Y_k]) \leq (\gamma_k - \gamma^* + \eta_k \mathbb{E}_k[Y_k])/(\gamma_k - \gamma^*); inequality (h) is due to Lemma 5-(iii).

For proceed to show inequality (39) for $\gamma_k \geq 3\gamma^*$, we need to upper bound $\text{Var}[Y_k]$ in inequalities (44) and (45). First, we compute the expectation $\mathbb{E}[Y_k]$ as follows:

$$
\mathbb{E}_k[Y_k] = \mathbb{E}\left[\frac{1}{6}\max\{3\gamma_k, Z_D^2\} - 2\gamma_k \max\{3\gamma_k, Z_D^2\}\right]
\leq \frac{3}{2}\gamma_k^2 + \mathbb{E}\left[\frac{1}{6}(Z_D^2 - 3\gamma_k)^2\mathbb{I}(Z_D^2 \geq 3\gamma_k)\right]
\leq \frac{3}{2}\gamma_k^2 + \mathbb{E}\left[\frac{1}{6}(Z_D^2)^2\mathbb{I}(Z_D^2 \geq 3\gamma_k)\right]
\leq \frac{3}{2}\gamma_k^2 + \frac{1}{2}\mathbb{E}[D]\mathbb{I}(D \geq 3\gamma_k + 1/2). \tag{46}
$$

Given historical information $\mathcal{H}_{k-1}$, the variance of $Y_k$ can be computed by:

$$
\text{Var}[Y_k|\mathcal{H}_{k-1}]
= \mathbb{E}_k[(Y_k - \mathbb{E}_k[Y_k])^2]
= \mathbb{E}_k\left[\left(\frac{1}{6}Z_D^2 - \gamma_k Z_D^2 + \frac{3}{2}\gamma_k^2 - \frac{3}{2}\gamma_k - \mathbb{E}_k[Y_k]\right)^2\mathbb{I}(Z_D^2 \geq 3\gamma_k)\right]
+ \mathbb{E}_k\left[\left(\frac{3}{2}\gamma_k^2 - \mathbb{E}_k[Y_k]\right)^2\mathbb{I}(Z_D^2 \leq 3\gamma_k)\right]
\leq \frac{1}{4}B + 2\mathbb{E}_k\left[\left(\frac{1}{6}Z_D^2 - \gamma_k Z_D^2 + \frac{3}{2}\gamma_k^2 - \frac{3}{2}\gamma_k - \mathbb{E}_k[Y_k]\right)^2\mathbb{I}(Z_D^2 > 3\gamma_k)\right]
+ 2\mathbb{E}_k\left[\left(\frac{3}{2}\gamma_k^2 - \mathbb{E}_k[Y_k]\right)^2\mathbb{I}(Z_D^2 > 3\gamma_k)\right]
\leq \frac{3}{4}B + \frac{1}{4}\mathbb{E}_k\left[(Z_D^2 - 3\gamma_k)^4\mathbb{I}(Z_D^2 > 3\gamma_k)\right]
\leq \frac{3}{4}B + \frac{1}{4}\mathbb{E}[Z_D^4] \leq (35 + \frac{3}{4})B, \tag{47}
$$

where (i) is because $\mathbb{E}_k[Y_k] \leq -\frac{3}{2}\gamma_k^2 + \frac{1}{2}\sqrt{B}$ implies $(-\frac{3}{2}\gamma_k^2 - \mathbb{E}_k[Y_k])^2 \leq \frac{1}{4}B$ and $(a + b)^4 \leq 2(a^4 + b^4)$.

Denote $N_1 := \max\{(35 + \frac{3}{4})B, \frac{1}{4}((9\gamma^*)^4 + B) + (3\gamma^*)^2((9\gamma^*)^2 + 3\sqrt{B})\}$, inequalities (40), (44) and (45) then...
lead to:

$$E_k[V(\gamma_{k+1})] - V(\gamma_k) \leq -\eta_kD_{th}V(\gamma_k) + \eta_k^2N_1.$$  \hspace{1cm} (48)

**Step 2: Computing \(E[V(\gamma_k)]\) through iteration:** Taking the expectation with respect to \(H_{k-1}\) on both sides of (48), we have:

$$E[V(\gamma_{k+1})] \leq (1 - \eta_kD_{th})E[V(\gamma_k)] + \eta_k^2N_1.$$  \hspace{1cm} (49)

Multiplying inequality (49) from \(i = 1\) to \(k\) yields:

$$E[V(\gamma_{k+1})] \leq \prod_{i=1}^{k} (1 - \eta_iD_{th})V(\gamma_0) + \sum_{i=1}^{k} \eta_i^2N_1 \cdot \prod_{j=i+1}^{k} (1 - \eta_jD_{th}).$$  \hspace{1cm} (50)

Since the stepsize selected by (22) satisfies

$$\eta_k \to 0, \liminf_{k} \min_{n \geq t \geq m(t_k-T)} \eta_n = 1$$

according to [24, p. 343, Eq. (4.8)], term \(\prod_{i=1}^{k} (1 - \eta_iD_{th}) = O(\eta_k)\). Therefore,

$$\sup_{k} E \left[ \frac{(\gamma_k - \gamma^*)^2}{\eta_k} \right] = \sup_{k} E \left[ 2V(\theta_k)/\eta_k \right] = O(1).$$  \hspace{1cm} (51)

This finishes the proof of Theorem 2.

**APPENDIX D**

**PROOF OF THEOREM 3**

Notice that the waiting time \(W_k \geq 0, \forall k\), we have:

$$\liminf_{k \to \infty} \frac{1}{k} \sum_{k'=1}^{k} (D_{k'} + W_{k'}) \geq \liminf_{k \to \infty} \frac{1}{k} \sum_{k'=1}^{k} D_{k'} = \bar{D} > 0, \text{w.p.1.}$$  \hspace{1cm} (52)

Therefore, to show sequence \(\{\int_0^{S_{k+1}} (X_t - \hat{X}_t)^2 dt\}_{k=1}^{\infty} \) converges to \(\bar{E}_\pi\), with probability 1, it is sufficient to show that the following sequence

$$\theta_k := \frac{1}{k} \left( \int_0^{S_{k+1}} (X_t - \hat{X}_t)^2 dt - (\gamma^* + \bar{D})S_{k+1} \right)$$

$$= \frac{1}{k} \sum_{k'=1}^{k} \left( \int_{S_{k'}}^{S_{k'+1}} (X_t - \hat{X}_t)^2 dt - (\gamma^* + \bar{D})L_{k'} \right)$$  \hspace{1cm} (53)

converges to 0 with probability 1. Recall that \(E_{k'} = \int_{S_{k'}}^{S_{k'+1}} (X_t - \hat{X}_t)^2 dt\) is the cumulative error in frame \(k'\), we can rewrite \(\theta_k\) in the following recursive form:

$$\theta_k = \frac{1}{k} \left( (k-1)\theta_{k-1} + E_k - (\gamma^* + \bar{D})L_k \right)$$

$$= \theta_{k-1} + \frac{1}{k} \left( -\theta_{k-1} + E_k - (\gamma^* + \bar{D})L_k \right).$$  \hspace{1cm} (54)

For notational simplicity, denote \(G_k := (-\theta_{k-1} + E_k - (\gamma^* + \bar{D})L_k)\), which can be viewed as the descent direction and can be further decomposed into:

$$G_k = -\theta_{k-1} + \int_{S_{k}}^{S_{k+1}} (X_t - X_{S_{k-1}})^2 dt$$

$$+ \int_{S_{k}}^{S_{k+1}} (X_t - X_{S_{k}})^2 dt - (\gamma^* + \bar{D})L_k$$

$$= -\theta_{k-1} + \int_{S_{k}}^{S_{k+1}} (X_t - X_{S_{k}})^2 dt + \int_{S_{k}}^{S_{k+1}} (X_t - X_{S_{k}})^2 dt - (\gamma^* + \bar{D})L_k$$

$$= \theta_{k-1} + (X_{S_k} - X_{S_{k-1}})^2 D_k$$

$$+ 2 (X_{S_k} - X_{S_{k-1}}) \cdot \int_{S_{k}}^{S_{k+1}} (X_t - X_{S_k}) dt$$

$$+ \int_{S_{k}}^{S_{k+1}} (X_t - X_{S_k})^2 dt - (\gamma^* + \bar{D})L_k.$$  \hspace{1cm} (55)

Give historical transmissions \(H_{k-1}, \gamma_k\) can be predicted and \(X_{S_k} - X_{S_{k-1}}\) is fixed, \(X_t - X_{S_k}\) evolves like a Wiener process and is independent of \(X_{S_k} - X_{S_{k-1}}\). Therefore, the conditional mean of \(G_{k,1}, \ldots, G_{k,4}\) can be computed as follows:

$$E_k[G_{k,1}] = \bar{D}(X_{S_k} - X_{S_{k-1}})^2,$$  \hspace{1cm} (56a)

$$E_k[G_{k,2}] = 0,$$  \hspace{1cm} (56b)

$$E_k[G_{k,3}] = \frac{1}{6} E_k \left[ \max\{3\gamma_k, Z_2^2, 2D_2\} \right] = q(\gamma_k),$$  \hspace{1cm} (56c)

$$E_k[G_{k,4}] = (\gamma^* + \bar{D})E_k \left[ \max\{3\gamma_k, Z_2^2\} \right] = (\gamma^* + \bar{D})l(\gamma_k).$$  \hspace{1cm} (56d)

where equation (56a) is because \(D_k\) is independent of \(X_{S_k} - X_{S_{k-1}}\); equation (56b) is because \(X_t - X_{S_k}\) is independent of \(X_{S_k} - X_{S_{k-1}}\) and has mean 0 for all \(t \geq S_k\); equation (56c) and (56d) is because of Lemma 2. With equation (56a)-(56d), given historical transmissions \(H_{k-1}\), we can compute the conditional expectation of \(G_k\) as follows:

$$E_k[G_k] = E_k \left[ -\theta_{k-1} + G_{k,1} + 2G_{k,2} + G_{k,3} - G_{k,4} \right]$$

$$= -\theta_{k-1} + (X_{S_k} - X_{S_{k-1}})^2 \bar{D} + q(\gamma_k) - (\gamma^* + \bar{D})l(\gamma_k)$$

$$= \left[ -\theta_{k-1} - q(\gamma_k) - (\gamma_k l(\gamma_k) + (\bar{D}(l(\gamma_k-1) - l(\gamma_k))) \right]$$

$$+ \bar{D} \left[ (X_{S_k} - X_{S_{k-1}})^2 - l(\gamma_k-1) \right] + (\gamma_k - \gamma^*)l(\gamma_k).$$  \hspace{1cm} (57)

Denote function

$$f(\theta, \gamma) := -\theta + E \left[ \frac{1}{6} \max\{3\gamma, Z_2^2\}^2 \right] - \gamma \max\{3\gamma, Z_2^2\}.$$  \hspace{1cm} (58)
and let function $\overline{f}(\cdot)$ be:

$$
\overline{f}(\theta) := f(\theta, \gamma^*).
$$

(59)

In the following analysis, we will prove that sequence $\{\theta_k\}$ converges to the stationary point of an ODE induced by function $\overline{f}(\theta)$. Let $\delta M_i := G_i - E_k[G_i] \text{ and let } \delta M_{i,k} := G_{i,k} - E_k[G_{i,k}]$ be the difference between each term and their conditional mean. We view $\frac{1}{k} := \epsilon_k$ as the updating step-sizes, which satisfies:

$$
\sum_k \epsilon_k = \infty, \sum_k \epsilon_k^2 < \infty.
$$

(60)

With $\epsilon_k, \beta_{k,1}, \beta_{k,2}$ and $\delta M_k$, the recursive equation (55) can be rewritten as follows:

$$
\theta_k = \theta_{k-1} + \epsilon_k (f(\theta_{k-1}, \gamma_k) + \beta_{k,1} + \beta_{k,2} + \beta_{k,3} + \delta M_k).
$$

(61)

Similarly, denote $t_0 = 0$ and $t_k := \sum_{i=0}^{k-1} \epsilon_i$ to be the cumulative step-size sequences. Let $m(t)$ be the unique $k \in \mathbb{N}^+$ such that $m(t) \leq t < m(t) + 1$. We then state the following characteristics of $G_k$ and $\delta M_k$, detailed proofs are in Appendix G of the supplementary material.

Claim 1: Sequences $\{G_k\}$ and $\{\delta M_k\}$ have the following properties:

(2.1) For each constant $N$, $\sup_k E \left[ |G_k| I(\theta_k \leq N) \right] < \infty$.

(2.2) Function $f(e, \gamma)$ is continuous in $e$ for each $\gamma$.

(2.3) For any $T > 0$, the following limit hold for all $\theta$:

$$
\lim_{k \to \infty} \Pr \left( \sup_{j \geq k} \sum_{0 \leq t \leq T} \left| \sum_{i=m(jT)}^{m(jT+t-1)} \epsilon_i (f(\theta, \gamma_i) - \overline{f}(\theta)) \right| \geq \mu \right) = 0.
$$

(62)

(2.4) For any $T > 0$, the difference sequence satisfies:

$$
\lim_{k \to \infty} \Pr \left( \sup_{j \geq k} \sum_{0 \leq t \leq T} \left| \sum_{i=m(jT)}^{m(jT+t-1)} \epsilon_i (f(\theta, \gamma_i) - \overline{f}(\theta)) \right| \geq \mu \right) = 0.
$$

(63)

(2.5) The bias sequence satisfies:

$$
\lim_{k \to \infty} \Pr \left( \sup_{j \geq k} \sum_{0 \leq t \leq T} \epsilon_i (\beta_{i,1} + \beta_{i,2} + \beta_{i,3}) \right) \geq \mu = 0.
$$

(64)

(2.6) For each $\theta$, function $f$ can be bounded as follows:

$$
f(\theta, \gamma) = \overline{f}(\theta) + \rho(\gamma),
$$

(65)

where $\rho(\gamma) = -(q(\gamma) - \gamma(\gamma))$ and for any $\tau > 0$ we have the following inequality:

$$
\lim_{k \to \infty} \Pr \left( \sup_{j \geq k} \sum_{i=m(j\tau)}^{m(j\tau+\tau-1)} |\epsilon_i \rho(\gamma_k)| \right) = 0.
$$

(66)

(2.7) For each $\theta_1, \theta_2$, the difference

$$
|f(\theta_1, \gamma) - f(\theta_2, \gamma)| = |\theta_1 - \theta_2|.
$$

(67)

When $\theta_1 - \theta_2 \to 0$, the absolute difference $|\theta_1 - \theta_2| \to 0$.

Denote $\theta_k(\omega)$ as the time averaged MSE up to frame $k$ of sample path $\omega$. Then according to [24, p.166, Theorem 1.1], with probability 1, sequence $\{\theta_k(\omega)\}$ converges to some limit set of the ODE

$$
\dot{\theta} = \overline{f}(\theta) = -\theta.
$$

(68)

Because $\overline{f}(0) = 0$, the minimum error $\theta = 0$ is an equilibrum point of the ODE in equation (68). Moreover, as $\overline{f}(\gamma)$ is a monotonic decreasing function, it can be easily verified through Lyapunov stability criterion that 0 is a unique stability point of the ODE (68). Therefore, $\theta_k$ converges to 0 with probability 1, and the time averaged MSE converges to $\mathbb{E}_{\pi^*}$ with probability 1.

Appendix E
Proof of Theorem 4

1) Proof of Inequality (28): Let $P_1, P_2$ be two delay distributions and let $\gamma_1^*, \gamma_2^*$ be the solution to (15) when $D \sim P_1$ and $D \sim P_2$, respectively. Through Le Cam’s inequality [33], we have:

$$
\inf_{\gamma} \sup_{\rho} \mathbb{E}
\left[ (\gamma - \gamma^*(\rho))^2 \right] \geq (\gamma_1^* - \gamma_2^*)^2 \cdot (\mathbb{P}_1 \otimes \mathbb{P}_2^k \wedge \mathbb{P}_2 \otimes \mathbb{P}_1^k),
$$

(69)

where $\mathbb{P} \otimes \mathbb{Q} := \int_{\Omega} \min \{p(x), q(x)\} dx$ and $\mathbb{P} \otimes \mathbb{P}$ is the product of distribution of $k$ i.i.d random variables drawn from $\mathbb{P}$.

To use Le Cam’s inequality (69), we need to find two distributions $P_1$ and $P_2$, whose $\ell_1$ distance $\|P_1 \otimes \mathbb{P} \otimes \mathbb{P}_1 \|_1$ is bounded, and the difference $(\gamma_1^* - \gamma_2^*)^2$ is of order $1/k$. We consider $P_1$ to be a uniform distribution on $[0, 1]$ and let $\gamma_1^*$ be the optimum ratio of distribution $P_1$. Through Corollary 3, we can obtain a loose upper bound on $\gamma_1^*$ as follows:

$$
\gamma_1^* \leq \sqrt{\mathbb{E}[D^2]} = \frac{1}{3}.
$$

(70)

Let $c \leq \frac{1}{2}$ be a constant and we denote

$$
\delta := \min\{1 - 3\gamma_1^*, 1/3, P_{\text{w, unm}}/2\}.
$$

(71)

Let $P_2$ be a probability distribution with probability density function $p_2(x)$ defined as follows:

$$
p_2(x) = \begin{cases} 
1 - c \sqrt{1/k}, & x \leq \frac{1}{2} \\
1, & \frac{1}{2} < x \leq 1 - \frac{1}{2} \delta \\
1 + c \sqrt{1/k}, & x > 1 - \frac{1}{2} \delta \\
0, & \text{otherwise.}
\end{cases}
$$

(72)

We will first bound $(\gamma_1^* - \gamma_2^*)^2$ (in Step 1) and $\mathbb{P}_1 \otimes \mathbb{P}_2^k \wedge \mathbb{P}_2 \otimes \mathbb{P}_1^k$ (in Step 2) as follows:

Step 1: Lower bounding $\gamma_2^* - \gamma_1^*$: For notational simplicity, denote function $h_1(\gamma) := E_{D \sim P_1^k}[\max\{3\gamma, Z_2^D\}^2 - \gamma \max\{3\gamma, Z_2^D\}^2]$ and $h_2(\gamma) := E_{D \sim P_2^k}[\max\{3\gamma, Z_2^D\}^2 - \gamma \max\{3\gamma, Z_2^D\}^2]$. According to the definition of $P_2$ in (72), for each $\gamma$, the difference between $h_1(\gamma)$ and $h_2(\gamma)$ can be computed by:

$$
h_2(\gamma) - h_1(\gamma) = \int_0^{1/2} \frac{c}{\sqrt{k}} \left[ 1 - \frac{1}{6} \max\{3\gamma, Z_2^D\}^2 \right] d\gamma.
$$

(73)
where inequality (a) is obtained because

\[
\frac{1}{6} \max \{3 \gamma, Z^2_D \}^2 - \gamma \max \{3 \gamma, Z^2_D \} = -\frac{3}{2} \gamma^2 + \frac{1}{6} (Z_D^2 - 3 \gamma^2) \mathbb{I}_{\{Z_D^2 \geq 3 \gamma^2\}}. \tag{74}
\]

Since \( \gamma^*_1 \) is the optimum ratio for delay distribution \( \mathbb{P}_1 \), we have \( h_1(\gamma^*_1) = 0 \). According to equation (73), function \( h_2(\gamma^*_1) \) can be lower bounded by:

\[
h_2(\gamma^*_1) \geq \frac{c}{\sqrt{k}} \cdot \int_{1 - \delta/2}^{1} \mathbb{E} \left[ \frac{1}{6} (Z_D^2 - 3 \gamma^*_1)^2 \mathbb{I}_{\{Z_D^2 \geq 3 \gamma^*_1\}} \right] dx \quad (b)
\]

\[
- \int_{0}^{\delta/2} \frac{c}{\sqrt{k}} \frac{x^2}{2} dx \geq -\frac{c}{\sqrt{k}} \int_{1 - \delta/2}^{1} \mathbb{E} \left[ \frac{1}{6} (Z_D^2 - 3 \gamma^*_1)^2 \mathbb{I}_{\{Z_D^2 \geq 3 \gamma^*_1\}} \right] dx \quad (c)
\]

\[
- \frac{c}{\sqrt{k}} \left( \gamma^*_1 \right)^3. \tag{75}
\]

where inequality (b) is because \( \mathbb{E} \left[ \frac{1}{6} (Z_D^2 - 3 \gamma^*_1)^2 \mathbb{I}_{\{Z_D^2 \geq 3 \gamma^*_1\}} \right] = x \leq \mathbb{E} \left[ \frac{1}{6} Z_D^2 \right] = x = \frac{3}{2} x^2 \).

We then proceed to lower bound \( \mathbb{E} \left[ \frac{1}{6} (Z_D^2 - 3 \gamma^*_1)^2 \mathbb{I}_{\{Z_D^2 \geq 3 \gamma^*_1\}} \right] \) for each delay realization \( x \in [1 - \delta/2, 1] \) as follows:

\[
\mathbb{E} \left[ \frac{1}{6} (Z_D^2 - 3 \gamma^*_1)^2 \mathbb{I}_{\{Z_D^2 \geq 3 \gamma^*_1\}} \right] \geq \mathbb{E} \left[ \frac{1}{6} (Z_D^2 - 3 \gamma^*_1)^2 \mathbb{I}_{\{Z_D^2 \leq 3 \gamma^*_1\}} \right] + \mathbb{E} \left[ \frac{1}{6} (Z_D^2 - 3 \gamma^*_1)^2 \mathbb{I}_{\{Z_D^2 \geq 3 \gamma^*_1\}} \right] \geq \mathbb{E} \left[ \frac{1}{6} (Z_D^2 - 3 \gamma^*_1)^2 \mathbb{I}_{\{Z_D^2 \leq 3 \gamma^*_1\}} \right] + \frac{1}{6} (\text{Var}[Z^2_D] \mathbb{I}[D = x] - x^2 \Pr [Z^2_D \leq x | D = x]) \quad (d)
\]

\[
\geq \frac{1}{6} x^2 - \frac{1}{6} (1 - \delta/2)^2. \tag{76}
\]

where inequality (c) is because \( \delta \geq 1 - 3 \gamma^*_1 \) by equation (71), and for the conditional mean \( \mathbb{E}[Z^2_D | D = x] \geq x \geq 1 - \delta/2 \geq 1 - \delta \geq 3 \gamma^*_1 \); inequality (d) is because \( \text{Var}[Z^2_D | D = x] = 2 x^2 \) and \( x^2 \Pr [Z^2_D \leq x] \leq x^2 \) and \( \mathbb{E} \left[ \frac{1}{6} (Z_D^2 - 3 \gamma^*_1)^2 \mathbb{I}_{\{3 \gamma^*_1 \leq Z_D^2 \leq x\}} \right] \geq 0 \). Plugging inequality (76) into (75) and recall that \( \delta < 1 \) by definition, we have the lower bound of \( h_2(\gamma^*_1) \):

\[
h_2(\gamma^*_1) \geq \frac{c}{\sqrt{k}} \frac{1}{26} \left( 1 - \frac{\delta}{2} \right)^2 - \frac{\delta}{2} \frac{\delta}{2} = \frac{c}{\sqrt{k}} \left( 1 - \frac{\delta}{2} \right)^2 \geq \frac{c}{\sqrt{k}} \left( 1 - \frac{\delta}{2} \right) > 0. \tag{77}
\]

By Lemma 5(i), function \( h_2(\gamma) \) is monotonically decreasing. Since \( h_2(\gamma^*_1) > 0 \) and \( h_2(\gamma^*_2) = 0 \), we can conclude that \( \gamma^*_2 \geq \gamma^*_1 \). We then proceed to bound \( \gamma^*_2 - \gamma^*_1 \) through Taylor expansion at \( \gamma = \gamma^*_1 \):

\[
h_2(\gamma^*_2) = h_2(\gamma^*_1) + h'_2(\gamma^*_1)(\gamma^*_2 - \gamma^*_1), \tag{78}
\]

where \( \gamma \in [\gamma^*_1, \gamma^*_2] \). Therefore, \( \gamma^*_2 \) can be computed by:

\[
\gamma^*_2 - \gamma^*_1 = \frac{-h_2(\gamma^*_1)}{h'_2(\gamma^*_1)}. \tag{79}
\]

To lower bound \( \gamma^*_2 \), we will first find a loose upper bound of \( \gamma^*_2 \) using Lemma 3:

\[
\gamma^*_2 \leq \frac{1}{2} \mathbb{E}_{D \sim P_2} [D^2] \leq \frac{1}{2} \left( \frac{1}{3} + \delta \cdot c \sqrt{1/k} \right). \tag{80}
\]

Therefore, since \( \delta < 1/3 \), we have \( |h'_2(\gamma)| \leq |h'_2(\gamma^*_2)| = \mathbb{E} \left[ \max \{3 \gamma^*_2, Z^2_D \} \right] \leq D + 3 \gamma^*_2 \leq 1 + \frac{1}{2} + \frac{3}{2} \cdot \frac{1}{3} \leq 2. \)

Then by inequality (77), we have

\[
\gamma^*_2 - \gamma^*_1 \geq \frac{-h_2(\gamma^*_1)}{h'_2(\gamma^*_2)} \geq \frac{1}{24}(1 - \delta)c \sqrt{\frac{1}{k}}. \tag{81}
\]

**Step 2: Lower bounding \( \mathbb{P}^{\otimes k} \land \mathbb{P}^{\otimes k} \):** Let \( |P - Q| = \int_{0}^{1} |dP - dQ| \) be the \( \ell_1 \) distance between probability distribution \( \mathbb{P} \) and \( Q \). Then

\[
\mathbb{P}^{\otimes k} \land \mathbb{P}^{\otimes k} = \int \min \{ \mathbb{P}^{\otimes k}(dx), \mathbb{P}^{\otimes k}(dx) \} = \int \mathbb{P}^{\otimes k}(dx) \cdot \left( 1 - \left( \frac{\mathbb{P}^{\otimes k}(dx)}{\mathbb{P}^{\otimes k}(dx)} \right)^{+} \right) \]

\[
= 1 - \int \left( \mathbb{P}^{\otimes k}(dx) - \mathbb{P}^{\otimes k}(dx) \right)^{+} \]

\[
= 1 - \frac{1}{2} \left| \mathbb{P}^{\otimes k} - \mathbb{P}^{\otimes k} \right|. \tag{82}
\]

Equality (82) enables us to lower bound \( \mathbb{P}^{\otimes k} \land \mathbb{P}^{\otimes k} \) by upper bounding the \( \ell_1 \) distance \( \mathbb{P}^{\otimes k} - \mathbb{P}^{\otimes k} \), which can be obtained by the Pinsky’s inequality:

\[
\frac{1}{2} \left| \mathbb{P}^{\otimes k} - \mathbb{P}^{\otimes k} \right| \leq \left( \frac{1}{2} D_{KL}(\mathbb{P}^{\otimes k} || \mathbb{P}^{\otimes k}) \right)^{+} = \frac{1}{2} D_{KL}(\mathbb{P}^{\otimes k} || \mathbb{P}^{\otimes k}) \quad \text{c}
\]

\[
\leq \frac{1}{2} \int_{0}^{1} p_2(x) \ln p_2(x) dx \quad \text{f}
\]

\[
\leq \frac{1}{2} \int_{0}^{1} \left( p_2(x) - 1 + \frac{1}{\min(p_2(x), 1)}(p_2(x) - 1)^2 \right) dx \quad \text{g}
\]

\[
\leq \frac{1}{2} \inf_{0 \leq x \leq 1} p_2(x) \int_{0}^{1} (p_2(x) - 1)^2 dx \quad \text{h}
\]
where inequality (e) is because the density function \( p_1(x) = 1 \) for uniform distribution, therefore \( D_K(P_1||P_2) = \int_{-\infty}^{\infty} p_2(x) \ln \frac{p_2(x)}{p_1(x)} dx \); inequality (f) is because function \( g(t) := (t \ln t) \) is convex, its derivative \( g(t)' = 1/t \), therefore, through Taylor expansion we have \( g(t) = g(1) + (t - 1) + \frac{1}{2} \ln(t) \), \( (t - 1)^2 = (t - 1) + \frac{1}{2} \ln(t) \); inequality (g) is because \( \int_0^1 p_2(x) dx = 1 \).

By choosing \( c = 1/2 \) and recall that \( \delta < 1 \), inequality (83) can be upper bounded by:

\[
\frac{1}{2} |\mathbb{P}^{\otimes k}_1 - \mathbb{P}^{\otimes k}_2|_1 \leq \frac{1}{2}.
\]

Plugging (84) into (82) yields:

\[
\mathbb{P}^{\otimes k}_1 \wedge \mathbb{P}^{\otimes k}_2 \geq \frac{1}{2}.
\]

Finally, plugging (85) and (81) into the Le Cam’s inequality (69) finishes the proof of inequality (28):

\[
\inf_{\tau} \sup_p (\gamma - \gamma^*(\mathbb{P}))^2 \geq \frac{1}{2} \left( \frac{1}{24} (1 - \delta) \delta p_{\text{w},\text{med}}^* \right)^2 \cdot \frac{1}{k} =: N.
\]

(86)

2) Proof of Inequality (29): The proof is divided into three steps: First we decompose the cumulative MSE gap up to \( S_{k+1} \) into the cumulative MSE gap within each frame, and then lower bound the frame regret in each frame using the difference between frame-length \( L_k \) and the optimum frame-length \( l^*(\mathbb{P}) \); then we obtain the minimax lower bound of \( (\mathbb{E}[L_k] - l^*(\mathbb{P}))^2 \) and finish the proof.

Step 1: Cumulative MSE decomposition:

\[
\mathbb{E} \left[ \int_0^{S_{k+1}} (X_t - \hat{X}_t)^2 dt - (\gamma^* + D)S_{k+1} \right]
\]

\[
= \sum_{k'=1}^k \mathbb{E} \left[ \int_{S_{k'}}^{S_{k'+1}} (X_t - \hat{X}_t)^2 dt - (\gamma^* + D)L_{k'} \right]
\]

\[
= \sum_{k'=1}^k \left( \mathbb{E} \left[ \int_{S_{k'}}^{S_{k'+1}} (X_t - X_{S_{k'}})^2 dt - (\gamma^* + D)L_{k'} \right] \right)
\]

\[
+ \sum_{k'=1}^k \left( \mathbb{E} \left[ \int_{S_{k'}}^{S_{k'+1}} (X_t - X_{S_{k'}} + X_{S_{k'}} - X_{S_{k'-1}})^2 dt - (\gamma^* + D)L_{k'} \right] \right)
\]

\[
= \sum_{k'=1}^k \left( \mathbb{E} \left[ \int_{S_{k'}}^{S_{k'+1}} D_{k'} \right] \right)
\]

\[
= \sum_{k'=1}^k \left( \mathbb{E} \left[ \int_{S_{k'}}^{S_{k'+1}} (X_t - X_{S_{k'}})^2 dt + (X_{S_{k'}} - X_{S_{k'-1}})^2 D_{k'} \right] \right)
\]

\[
= \sum_{k'=1}^k \left( \mathbb{E} \left[ \int_{S_{k'}}^{S_{k'+1}} (X_t - X_{S_{k'}})^2 dt - \gamma^* L_{k'} \right] \right)
\]

where equation (a) is because for \( \mathbb{E}[(X_t - X_{S_{k'}})(X_{S_{k'}} - X_{S_{k'-1}})] = 0 \), equation (b) is because \( D_{k'} \) is independent of \( X_{S_{k'}} - X_{S_{k'-1}} \) and \( \mathbb{E}[D_{k'}] = D \), \( \mathbb{E}[(X_{S_{k'}} - X_{S_{k'-1}})^2] = \mathbb{E}[(S_{k'} - S_{k'-1})^2] = L_{k-1} \).

We then proceed to lower bound each item \( \Upsilon_k \) in equation (87) using the following Lemma:

Lemma 6: For each sample policy \( \pi \) with a random sampling interval \( \tau \), let \( l_{\pi} := \mathbb{E}[\tau] = \mathbb{E}[Z_D^2] \) denote the expected running length. Recall that \( \gamma^*(\mathbb{P}), l^*(\mathbb{P}) \) are the optimum ratio and optimum frame length when delay distribution \( D \sim \mathbb{P} \) and \( p_w(\mathbb{P}) := \Pr(Z_D^2 \leq 3\gamma^*(\mathbb{P})) \) be the probability of waiting, the following inequality holds:

\[
\mathbb{E} \left[ \int_{t=0}^{T} Z_D^2 dt \right] - \gamma^*(\mathbb{P}) \mathbb{E}[\tau] \geq \frac{1}{6} p_w(\mathbb{P}) (l_{\pi} - l^*(\mathbb{P}))^2,
\]

where \( l^*(\mathbb{P}) := \mathbb{E}_{D \sim \mathbb{P}}[\max(3\gamma^*(\mathbb{P}), Z_D^2)] \) is the average frame length when the optimum policy \( \pi^*(\mathbb{P}) \) is used.

Proof for Lemma 6 is provided in Appendix M of the supplementary material. Notice that \( X_t - X_{S_k} \) is a Wiener Process starting from time \( t = S_k \), \( k \rightarrow k-1 \) records the previous delay and Wiener process evolution at the beginning of frame \( k \). Since the sampling policy in frame \( k \) depends on \( H_{k-1} \) and \( \delta X_k = X_{S_k + H_k} - X_{S_k} \), we can lower bound the worst case regret of \( \Upsilon_k \) as follows:

\[
\inf_{\tau} \sup_{\mathbb{P}} \Upsilon_k
\]

\[
= \inf_{\tau} \sup_{\mathbb{P}} \frac{1}{6} p_w(\mathbb{P}) \mathbb{E}_{H_{k-1}} \left[ (\mathbb{E}[L_k|H_{k-1}] - l^*(\mathbb{P}))^2 \right]
\]

\[
\geq \inf_{\tau} \sup_{\mathbb{P}} \frac{1}{6} p_w(\mathbb{P}) \mathbb{E}_{H_{k-1}} \left[ (\mathbb{E}[L_k|H_{k-1}] - l^*(\mathbb{P}))^2 \right]
\]

\[
\geq \inf_{\mathbb{P}} \max_{P_1 \in \{P_1, P_2\}} \frac{1}{6} p_w(\mathbb{P}) \mathbb{E}_{H_{k-1}} \left[ (\mathbb{E}[L_k|H_{k-1}] - l^*(\mathbb{P}))^2 \right]
\]

\[
= \frac{1}{6} \min\left\{ p_w(\mathbb{P}_1), p_w(\mathbb{P}_2) \right\} \times \inf_{\mathbb{P}} \sup_{P_1 \in \{P_1, P_2\}} \mathbb{E}_{H_{k-1}} \left[ (\mathbb{E}[L_k|H_{k-1}] - l^*(\mathbb{P}))^2 \right].
\]

Inequality (89) works for any distribution \( P_1 \) and \( P_2 \). We select \( P_1 \) to be the uniform distribution over interval \([0, 1]\) and \( P_2 \) using equation (72). Then the first term \( H_1 \) in (89) can
be lower bounded by:

\[ H_1 = \min \{ \mu(P_1), \mu(P_2) \} \]

\[ \geq \min \left\{ \text{Pr} \left( Z_B^D \geq 3 \gamma_1^* | D \sim P_1 \right), \text{Pr} \left( Z_B^D \geq 3 \gamma_2^* | D \sim P_2 \right) \right\} \]

\[ \geq \min \left\{ \frac{\text{Pr}(Z_B^D \geq 3 \gamma_1^*)}{3 \times 3}, \frac{\text{Pr}(Z_B^D \geq 3 \gamma_2^*)}{3 \times 3} \right\} \]

\[ \geq \min \{1/2, 1/7\} = 1/2, \]  

(90)

where inequality (c) is by Markov inequality; inequality (d) is because \( \text{E}[Z_B^D] = \text{E}[D] \) by the optimal stopping theorem, \( \gamma_1^* \leq 1/4 \) from (70) and \( \gamma_2^* \leq 1/2 \left( \frac{1}{2} + \frac{1}{2} \cdot \sqrt{1/2} \right) < 7/24 \); inequality (e) is because \( \text{E}_P[Z_B^D] = 1/2 \) for uniform distribution \( P_1 \) and \( \text{E}_P[Z_B^D] = 1 \) due to the distribution of \( P_2 \) in equation (72). It then remains to prove that the second term \( H_2 \) in (89).

Step 2: Since \( L_k \) is made using \( k \) i.i.d samples \( \delta X^{\otimes k} = \{ \delta X_k = (X_{S_k+1} - X_{S_k}) \} \), where \( \delta X^{\otimes (k-1)} \) are from \( H_{k-1} \) and \( \delta X_k = X_{S_k+1} - X_{S_k} \), \( E[L_k | H_{k-1}] \) can be viewed as a deterministic estimator for the corresponding \( l^*(P) \). Let \( \hat{l} : \mathbb{R}^k \to \mathbb{R}^+ \) an arbitrary deterministic estimation function, term \( H_2 \) in equation (89) is equivalent to:

\[ H_2 = \inf \max_{\hat{l}} \left( \text{E}_P \left[ (\hat{l}(\delta X^{\otimes k}) - l^*(P))^2 \right] \right). \]

(91)

To obtain the lower bound of (91), we come up with the following optimization problem:

**Problem 4:**

\[ \epsilon^* := \min_{\epsilon, \hat{l}} \epsilon \]

\[ \text{s.t.} \quad \text{E}_P \left[ (\hat{l}(\delta X^{\otimes k}) - l^*(P))^2 \right] \leq \epsilon \]

\[ \text{E}_P \left[ (\hat{l}(\delta X^{\otimes k}) - l^*(P))^2 \right] \leq \epsilon. \]

The minimum \( \epsilon^* \) satisfies:

\[ \epsilon^* \geq \frac{1}{6} \left( \frac{1}{24} (1 - \delta) \delta p^{\text{uni}}_{\text{w}, \text{uni}} \right)^2 \frac{1}{k}. \]

(93)

Detailed proof is provided in Appendix F.

Step 3: Plugging (90) and (93) into (89), we have:

\[ \inf \sup \pi \gamma_k \geq \frac{1}{24} \left( \frac{1}{24} (1 - \delta) \delta p^{\text{uni}}_{\text{w}, \text{uni}} \right)^2 \frac{1}{k}. \]

(94)

Summing up \( \gamma_k \) from \( k' = 1, 2, \ldots, k \) and plugging (94) into (87), we have:

\[ \inf \pi \left[ \int_{S_{k+1}} (x_t - x_t)^2 dt - (\gamma^* + D) S_{k+1} \right] \]

\[ \geq \sum_{k'=1}^k \inf \sup \pi \gamma_k - \mathbb{E}[L_k] \]

\[ = \sum_{k'=1}^k \inf \sup \pi \gamma_k - \mathbb{D}[\pi][L_k] - l^*(\pi) - \mathbb{D}l^*(\pi) \]

\[ \geq \frac{1}{24} \left( \frac{1}{24} (1 - \delta) \delta p^{\text{uni}}_{\text{w}, \text{uni}} \right)^2 \frac{1}{k} = \Omega(\ln k). \]

(95)

### Appendix F

**Solution to Problem 4**

We use the Lagrange method for solving the optimization problem. Let \( \rho : \mathbb{R}^k \to \mathbb{R} \) and \( \lambda_1, \lambda_2 \geq 0 \) be Lagrange multipliers, the Lagrange function for solving Problem 4 is as follows:

\[ \mathcal{L}(\epsilon, \hat{l}, \lambda_1, \lambda_2) = \epsilon + \lambda_1 \left( \mathbb{E}_P \left[ (\hat{l}(\delta X^{\otimes k}) - l^*_1)^2 \right] - \epsilon \right) \]

\[ + \lambda_2 \mathbb{E}_P \left[ (\hat{l}(\delta X^{\otimes k}) - l^*_2)^2 \right] - \epsilon. \]

(96)

The Gâteaux derivative of the Lagrange \( \mathcal{L} \) in the direction of \( \rho : \mathbb{R}^k \to \mathbb{R} \) is defined as

\[ \delta \mathcal{L}(\hat{l}; \epsilon, \lambda_1, \lambda_2, \rho) \]

\[ := \lim_{\epsilon \to 0} \mathcal{L}(\epsilon, \hat{l} + \epsilon \rho, \lambda_1, \lambda_2) - \mathcal{L}(\hat{l}, \lambda_1, \lambda_2) \]

\[ = 2\rho \left( \delta X^{\otimes k} \right) \left( \lambda_1 p_1 \left( (\delta X^{\otimes k})^2 \right) (\hat{l}(\delta X^{\otimes k}) - l^*_1) \right) \]

\[ + \lambda_2 p_2 (\delta X^{\otimes k}) (\hat{l}(\delta X^{\otimes k}) - l^*_2). \]

(97)

Let \( (\hat{l}^*, \epsilon^*, \lambda_1^*, \lambda_2^*) \) be the dual optimizer. To satisfy the KKT condition, we require:

\[ \delta \mathcal{L}(\hat{l}; \epsilon^*, \lambda_1^*, \lambda_2^*, \rho) \bigg|_{\epsilon = \epsilon^*} = 0, \forall \rho, \]

(98a)

and the Complete Slackness (CS) condition require:

\[ \lambda_1^* \mathbb{E}_P \left[ (\hat{l}^*(\delta X^{\otimes k}) - l^*_1)^2 \right] - \epsilon^* = 0, \]

(98c)

\[ \lambda_2^* \mathbb{E}_P \left[ (\hat{l}^*(\delta X^{\otimes k}) - l^*_2)^2 \right] - \epsilon^* = 0. \]

(98d)

The KKT condition in equation (98a) implies the optimum estimator \( l^* \) is:

\[ \hat{l}^*(\delta X^{\otimes k}) = \lambda_1^* p_1 (\delta X^{\otimes k}) l^*_1 + \lambda_2^* p_2 (\delta X^{\otimes k}) l^*_2 \]

(98e)

and equation (98b) requires:

\[ \lambda_1^* + \lambda_2^* = 1. \]

(98f)

It can be verified that \( \lambda_1^* \neq 0 \) and \( \lambda_2^* \neq 0 \) because if \( \lambda_1^* = 0 \), to satisfy equation (98e), we have \( l^* \left( \delta X^{\otimes k} \right) \equiv l^*_1 \). Then \( \epsilon^* = (l^*_2 - l^*_1)^2 \) is clearly not the optimum value. For fixed \( \lambda_1, \lambda_2 \), by plugging function (98e) into (98c) and (98d), we have:

\[ \epsilon^* = \mathbb{E}_P \left[ (\hat{l}^*(\delta X^{\otimes k}) - l^*_1)^2 \right] \]

\[ = (l^*_2 - l^*_1)^2 \int \left( \frac{\lambda_1^* p_2 (\delta X^{\otimes k})}{\lambda_1^* p_1 (\delta X^{\otimes k}) + \lambda_2^* p_2 (\delta X^{\otimes k})} \right)^2 \text{d}\delta X^{\otimes k}, \]

(99)

\[ \epsilon^* = \mathbb{E}_P \left[ (\hat{l}^*(\delta X^{\otimes k}) - l^*_2)^2 \right] \]

\[ = (l^*_2 - l^*_1)^2 \int \left( \frac{\lambda_1^* p_2 (\delta X^{\otimes k})}{\lambda_1^* p_1 (\delta X^{\otimes k}) + \lambda_2^* p_2 (\delta X^{\otimes k})} \right)^2 \text{d}\delta X^{\otimes k}. \]

(100)
Since $\lambda_1^* + \lambda_2^* = 1$, (99) and (100) imply:

$$\varepsilon^* = \lambda_1^* \varepsilon_1^* + \lambda_2^* \varepsilon_2^* = \lambda_1^* E_p (i^*(\delta X^{\otimes k}) - l_1^*)^2 + \lambda_2^* E_p (i^*(\delta X^{\otimes k}) - l_2^*)^2 = (l_2^* - l_1^*)^2 \int \lambda_1^* p_1 (\delta X^{\otimes k}) \lambda_2^* p_2 (\delta X^{\otimes k}) d\delta X^{\otimes k}.$$ 

$$\geq \frac{1}{2} \min \{\lambda_1^*, \lambda_2^*\} \{p_1^{\otimes k} \land p_2^{\otimes k}\}.$$ 

(101)

where inequality (f) is because $\frac{a + b}{k} \geq \frac{1}{2} \min \{a, b\}$.

Next, we bound each term in (101) respectively.

**Term 1** The lower bound of $l_2^* - l_1^*$ is as follows:

$$l_2^* - l_1^* = \int_0^1 E[\max\{3\gamma_2^*, Z_D^2\}|D = x] dx + \int_{1 - \delta/2}^{1} c \sqrt{\frac{k}{\delta}} E[\max\{3\gamma_2^*, Z_D^2\}|D = x] dx - \int_{\delta/2}^{1} c \sqrt{\frac{k}{\delta}} E[\max\{3\gamma_1^*, Z_D^2\}|D = x] dx - \int_0^{1} E[\max\{3\gamma_1^*, Z_D^2\}|D = x] dx.$$ 

Notice that if $x_1 \geq x_2$,

$$E[\max\{3\gamma, Z_D^2\}|D = x_1] = E[\max\{3\gamma, Z_D^2\}|D = x_2] \geq 0.$$ 

(103)

Therefore, inequality (102) can be bounded by:

$$l_2^* - l_1^* \geq \int_0^1 E[\max\{3\gamma_2^*, Z_D^2\}|D = x] dx - \int_0^1 E[\max\{3\gamma_1^*, Z_D^2\}|D = x] dx \geq 3(\gamma_2^* - \gamma_1^*) E_D \cdot p_1(\sup\{Z_D^2 \leq 3\gamma_1^*\}) \geq \frac{1}{24}(1 - \delta)\delta \epsilon p^*_w, \text{\textsuperscript{uni}} \sqrt{\frac{1}{k}}.$$ 

(104)

where inequality (g) is obtained by equation (81).

**Term 2** To lower bound $\min\{\lambda_1, \lambda_2\}$, recall that equation (99) equals (100), we have:

$$(\lambda_2^*)^2 \int \frac{p_1 (\delta X^{\otimes k}) \times p_2 (\delta X^{\otimes k})}{\lambda_1^* p_1 (\delta X^{\otimes k}) + \lambda_2^* p_2 (\delta X^{\otimes k})} d\delta X^{\otimes k} = (\lambda_2^*)^2 \int \frac{p_1 (\delta X^{\otimes k}) \times p_2 (\delta X^{\otimes k})}{\lambda_1^* p_1 (\delta X^{\otimes k}) + \lambda_2^* p_2 (\delta X^{\otimes k})} p_1 (\delta X^{\otimes k}) d\delta X^{\otimes k}.$$ 

(105)

Equation (105) implies we can upper and lower bound $\lambda_1 / \lambda_2$ as follows:

$$\inf \sqrt{\frac{p_2 (\delta X^{\otimes k})}{p_1 (\delta X^{\otimes k})}} \leq \frac{\lambda_1}{\lambda_2} \leq \sup \sqrt{\frac{p_2 (\delta X^{\otimes k})}{p_1 (\delta X^{\otimes k})}}.$$ 

(106)

According to the density function defined in (72), we have:

$$- \frac{1}{\sqrt{k}} \leq \frac{\lambda_1}{\lambda_2} \leq \frac{1}{\sqrt{k}} + c \left[\frac{1}{\sqrt{k}}\right].$$ 

(107)

Since $\epsilon \leq 1/2$, we have $\sqrt{1/2} \leq \frac{\lambda_1}{\lambda_2} \leq \sqrt{3/2}$ and therefore

$$\min\{\lambda_1, \lambda_2\} \geq 1/3.$$ 

(108)

Finally, ($p_1^{\otimes k} \land p_2^{\otimes k}$) $\geq 1/2$ according to (85). Plugging (104) and (108) into (101), we have:

$$H_2 \geq \frac{1}{6} \left(\frac{1}{24} (1 - \delta)\delta \epsilon p^*_w, \text{\textsuperscript{uni}} \right)^2 (\lambda_2^*)^2.$$ 

(109)

REFERENCES

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

[2] R. D. Yates, “Lazy is timely: Status updates by an energy harvesting source,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Jun. 2015, pp. 3008–3012.

[3] 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.

[4] B. Hajek, K. Mitzel, and S. Yang, “Paging and registration in cellular networks: Jointly optimal policies and an iterative algorithm,” in Proc. 22nd Ann. Joint Conf. IEEE Comput. Commun. Societies (INFOCOM), San Francisco, CA, USA, vol. I, 2003, pp. 524–532, doi: 10.1109/INFOCOM.2003.1208703.

[5] T. Z. Ornee and Y. Sun, “Sampling and remote estimation for the Ornstein–Uhlenbeck process through queues: Age of information and beyond,” IEEE/ACM Trans. Netw., vol. 29, no. 5, pp. 1962–1975, Oct. 2021.

[6] A. Nayyar, T. Basar, D. Teneketzis, and V. V. Veeravalli, “Optimal strategies for communication and remote estimation with an energy harvesting sensor,” IEEE Trans. Autom. Control, vol. 58, no. 9, pp. 2246–2260, Sep. 2013.

[7] Y. Sun, Y. Polyaniskiy, and E. Uysal, “Sampling of the Wiener process for remote estimation over a channel with random delay,” IEEE Trans. Inf. Theory, vol. 66, no. 2, pp. 1118–1135, Feb. 2020.

[8] C.-H. Tsai and C.-C. Wang, “Unifying AoI minimization and remote estimation—Optimal sensor/controller coordination with random two-way delay,” IEEE/ACM Trans. Netw., vol. 30, no. 1, pp. 229–242, Feb. 2022.

[9] T. Z. Ornee and Y. Sun, “Performance bounds for sampling and remote estimation of Gauss–Markov processes over a noisy channel with random delay,” in Proc. IEEE 22nd Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC), Sep. 2021, pp. 1–5.

[10] Z. Wang, M.-A. Badiu, and J. P. Coon, “Relationship between age and value of information for a noisy Ornstein–Uhlenbeck process,” Entropy, vol. 23, no. 8, p. 940, Jul. 2021. [Online]. Available: https://www.mdpi.com/1099-4300/23/8/940.

[11] C.-H. Tsai and C.-C. Wang, “Age-of-information revisited: Two-way delay and distribution-oblivious online algorithm,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Jun. 2020, pp. 1782–1787.

[12] K. Bhandari, S. Fatale, U. Narula, S. Moharir, and M. K. Hanawal, “Age-of-information bandits,” in Proc. 18th Int. Symp. Model. Optim. Mobile, Ad Hoc, Wireless Netw. (WIoPT), Jun. 2020, pp. 1–8.

[13] E. U. Atay, I. Kadota, and E. Modiano, “Aging bandits: Regret analysis and order-optimal learning algorithm for wireless networks with stochastic arrivals,” 2020, arXiv:2012.00682.

[14] A. Banerjee, R. Bhattacharjee, and A. Sinha, “Fundamental limits of age-of-information in stationary and non-stationary environments,” in Proc. IEEE Int. Symp. Inf. Theory (ISIT), Jun. 2020, pp. 1741–1746.

[15] V. Tripathi and E. Modiano, “An online learning approach to optimizing time-varying costs of AoI,” in Proc. 22nd Int. Symp. Theory, Algorithmic Found., Protocol Design Mobile Netw. Mobile Comput. New York, NY, USA: Association for Computing Machinery, Jul. 2021, pp. 241–250.

[16] B. Li, “Efficient learning-based scheduling for information freshness in wireless networks,” in Proc. IEEE Conf. Comput. Commun. (INFOCOM), May 2021, pp. 1–10.

[17] E. T. Ceran, D. Gündüz, and A. György, “Reinforcement learning to minimize age of information with an energy harvesting sensor with HARQ and sensing cost,” in Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS), Apr. 2019, pp. 656–661.

[18] E. T. Ceran, D. Gündüz, and A. György, “A reinforcement learning approach to age of information in multi-user networks with HARQ,” IEEE J. Sel. Areas Commun., vol. 39, no. 5, pp. 1412–1426, May 2021.
[19] C. Kam, S. Kompella, and A. Ephremides, “Learning to sample a signal through an unknown system for minimum AoI,” in Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS), Apr. 2019, pp. 177–182.

[20] M. A. Abd-Elmagid, H. S. Dhillon, and N. Pappas, “A reinforcement learning framework for optimizing age of information in RF-powered communication systems,” IEEE Trans. Commun., vol. 68, no. 8, pp. 4747–4760, Aug. 2020.

[21] S. Leng and A. Yener, “Age of information minimization for wireless ad hoc networks: A deep reinforcement learning approach,” in Proc. IEEE Global Commun. Conf. (GLOBECOM), Dec. 2019, pp. 1–6.

[22] H. Tang, Y. Chen, J. Sun, J. Wang, and J. Song, “Sending timely status updates through channel with random delay via online learning,” in Proc. IEEE Conf. Comput. Commun. (INFOCOM), London, U.K., May 2022, pp. 1819–1827.

[23] H. Tang, Y. Chen, J. Wang, P. Yang, and L. Tassiulas, “Age optimal sampling under unknown delay statistics,” IEEE Trans. Inf. Theory, vol. 69, no. 2, pp. 1295–1314, Feb. 2023.

[24] H. J. Kushner and G. G. Yin, Stochastic Approximation and Recursive Algorithms and Applications. New York, NY, USA: Springer, 2003, doi: 10.1007/0-387-21769-X_5.

[25] A. Arafa, R. D. Yates, and H. V. Poor, “Timely cloud computing: Preemption and waiting,” in Proc. 57th Annu. Allerton Conf. Commun., Control, Comput. (Allerton), Sep. 2019, pp. 528–535.

[26] M. Rabi, G. V. Moustakides, and J. S. Baras, “Adaptive sampling for linear state estimation,” SIAM J. Control Optim., vol. 50, no. 2, pp. 672–702, Jan. 2012, doi: 10.1137/090757125.

[27] M. J. Neely, “Stochastic network optimization with application to communication and queueing systems,” Synth. Lectures Commun. Netw., vol. 3, no. 1, pp. 1–211, Jan. 2010.

[28] H. Robbins and S. Monro, “A stochastic approximation method,” Ann. Math. Statist., vol. 22, no. 3, pp. 400–407, Sep. 1951.

[29] A. B. Tsybakov, Introduction to Nonparametric Estimation, 1st ed. Cham, Switzerland: Springer, 2008.

[30] J. Pan, A. M. Bedewy, Y. Sun, and N. B. Shroff, “Optimal sampling for data freshness: Unreliable transmissions with random two-way delay,” IEEE/ACM Trans. Netw., vol. 31, no. 1, pp. 408–420, Feb. 2023.

[31] J. Pan, Y. Sun, and N. B. Shroff, “Sampling for remote estimation of the Wiener process over an unreliable channel,” 2023, arXiv:2310.11037.

[32] A. Eryilmaz and R. Srikant, “Asymptotically tight steady-state queue length bounds implied by drift conditions,” Queueing Syst., vol. 72, nos. 3–4, pp. 311–359, Dec. 2012.

[33] B. Yu, “Assouad, Fano, and Le Cam,” in Festschrift for Lucien Le Cam. New York, NY, USA: Springer, 1997, pp. 423–435.

Yin Sun (Senior Member, IEEE) received the B.Eng. and Ph.D. degrees in electronic engineering from Tsinghua University, Beijing, China, in 2006 and 2011, respectively. From 2011 to 2017, he was a Post-Doctoral Scholar and a Research Associate with The Ohio State University. From 2017 to 2023, he was an Assistant Professor with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL, USA, where he is currently a Godbold Associate Professor. He has coauthored a monograph of Age of Information: A New Metric for Information Freshness. His research interests include wireless networks, machine learning, semantic communications, age of information, information theory, and robotic control. He is also interested in applying AI and machine learning techniques in agricultural, food, and animal sciences. He is a member of the ACM. His articles received the Best Student Paper Award from the IEEE/IFIP WiOpt 2013, the Best Paper Award from the IEEE/IFIP WiOpt 2019, the Runner-Up for the Best Paper Award of ACM MobiHoc 2020, and the 2021 Journal of Communications and Networks (JCN) Best Paper Award. He also received the Auburn Author Award of 2020 and the National Science Foundation (NSF) CAREER Award in 2023. He was named a Ginn Faculty Achievement Fellow in 2023. He has served as an Associate Editor for IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, an Editor for Journal of Communications and Networks and IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, a guest editor for five special journal issues, and an organizing committee member for several conferences. He founded the Age of Information (AoI) Workshop in 2018 and the Modeling and Optimization in Semantic Communications (MOSC) Workshop in 2023.

Leandros Tassiulas (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Maryland, College Park, MD, USA, in 1991. He was a Faculty Member at Polytechnic University, New York, NY, USA; the University of Maryland; and the University of Thessaly, Greece. He is currently the John C. Malone Professor of electrical engineering with Yale University, New Haven, CT, USA. His most notable contributions include the max-weight scheduling algorithm and the back-pressure network control policy, opportunistic scheduling in wireless, the maximum lifetime approach for wireless network energy management, and the consideration of joint access control and antenna transmission management in multiple antenna wireless systems. His research interests include computer and communication networks, with an emphasis on fundamental mathematical models and algorithms of complex networks, architectures and protocols of wireless systems, sensor networks, novel internet architectures, and experimental platforms for network research. His research has been recognized by several awards, including the IEEE Koji Kobayashi Computer and Communications Award, the Inaugural INFOCOM 2007 Achievement Award for fundamental contributions to resource allocation in communication networks, the INFOCOM 1994 and 2017 Best Paper Awards, the National Science Foundation (NSF) Research Initiation Award in 1992, the NSF CAREER Award in 1995, the Office of Naval Research Young Investigator Award in 1997, and the Bodossaki Foundation Award in 1999.