

Sampling of the Wiener Process for Remote Estimation over a Channel with Unknown Delay Statistics

Haoyue Tang
Department of EE and YINS
Yale University
New Haven, CT, USA
haoyue.tang@yale.edu

Yin Sun
Department of ECE
Auburn University
Auburn, AL, USA
yzs0078@auburn.edu

Leandros Tassioulas
Department of EE and YINS
Yale University
New Haven, CT, USA
leandros.tassioulas@yale.edu

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 a renewal reward optimization problem, and we propose an online sampling algorithm that can adaptively learn the optimal sampling policy through stochastic approximation. We show that the cumulative MSE regret grows with rate $O(\ln k)$, 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 $\Omega(\ln k)$. Hence, the proposed online sampling algorithm is minimax order-optimal. Finally, we validate the performance of the proposed algorithm via numerical simulations.

CCS CONCEPTS

• **Networks** → **Network performance evaluation**; Network performance analysis.

KEYWORDS

Age of Information, Online Learning, Stochastic Approximation

ACM Reference Format:

Haoyue Tang, Yin Sun, and Leandros Tassioulas. 2022. Sampling of the Wiener Process for Remote Estimation over a Channel with Unknown Delay Statistics. In *MobiHoc '22: the 23rd International Symposium on Mobile Ad Hoc Networking and Computing*, Seoul, South Korea, 17-22, Oct. 2022. ACM, New York, NY, USA, 24 pages. <https://doi.org/XXX>

1 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 [12]. 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 [33]. Samples with fresher content should be delivered to the destination in a timely manner [25].

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 [10, 16, 18, 24, 31]. When the random process can be observed at the sampler, the optimum sampling policy 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. The optimum threshold can be computed by iterative thresholding [30] or the bi-section search [24] 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 [3–5, 15, 29]. By using the AoI as the freshness metric, [3–5] design online link rate selection algorithms based on stochastic bandits. When the channels are time-varying and the transmitter has an average power constraint, [1, 6, 7, 11, 14] 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 [29], 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 [30] can be applied in the online setting when the delay statistics is unknown, whereas the convergence rate and the optimality of the algorithm are not well understood.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MobiHoc '22, Oct 17–20, 2022, Seoul, South Korea

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00

<https://doi.org/XXX>

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 optional stopping problem. By analyzing the sufficient conditions of the optimum threshold, we propose an online sampling policy that learns the optimum threshold adaptively through stochastic approximation. Compared with [26, 27, 30], 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 bounded (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 offline policy, grows at a speed of $O(\ln k)$, 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 [13]. 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.
- Further by using the classic 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 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 [30], the proposed algorithm could meet an average sampling frequency constraint.

2 SYSTEM MODEL AND PROBLEM FORMULATION

2.1 System Model

As is depicted in Fig. 1, we revisit the status update system in [2, 24, 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 on the First-Come-First-Serve (FCFS) basis. 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

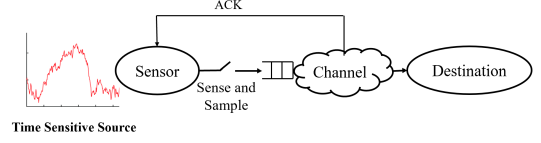


Figure 1: System model.

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 . The reception time of the k -th packet, denoted by R_k satisfies the following recursive formula: $R_k = \{S_k, R_{k-1}\} + D_k$ and we define $R_0 = 0$ for simplicity. We assume the average transmission delay $\bar{D} := \mathbb{E}_{D \sim \mathbb{P}_D}[D]$ is lower bounded by $\bar{D}_{lb} > 0$.

2.2 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 . The information available at the receiver at time t can be summarized as follows: (i). The sampling time-stamps, transmission delay and the values of previous samples $\mathcal{M}_t := \{(S_j, D_j, X_{S_j})\}_{j=1}^{i(t)}$; (ii). The fact that no packet was received during $(R_{i(t)}, t]$. Similar to [23, 25], we assume that the receiver estimates X_t only based on \mathcal{M}_t and neglects the second part of information. The minimum mean-square error (MMSE) estimator [19] in this case is:

$$\hat{X}_t = \mathbb{E}[X_t | \mathcal{M}_t] = X_{S_{i(t)}}. \quad (1)$$

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 π is denoted by $\bar{\mathcal{E}}_\pi$, i.e.,

$$\bar{\mathcal{E}}_\pi \triangleq \limsup_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \int_0^T (X_t - X_{S_{i(t)}})^2 dt \right]. \quad (2)$$

2.3 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 Π , 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 threshold. The optimal sampling problem is organized as follows:

PROBLEM 1 (MMSE MINIMIZATION).

$$\text{mse}_{\text{opt}} \triangleq \inf_{\pi \in \Pi} \limsup_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \int_0^T (\hat{X}_t - X_t)^2 dt \right], \quad (3a)$$

$$\text{s.t. } \limsup_{T \rightarrow \infty} \mathbb{E} \left[\frac{i(T)}{T} \right] \leq f_{\max}. \quad (3b)$$

3 PROBLEM SOLUTION

In this section, the MSE minimization problem (i.e., Problem 1) is reformulated into an optional stopping problem. Let π^* be an optimum policy whose average MSE achieves mse_{opt} . Sufficient conditions for π^* are provided in Subsection 3.2. The online sampling algorithm π_{online} is provided in Subsection 3.3 and Subsection 3.4 characterizes the behaviors of the online sampling policy.

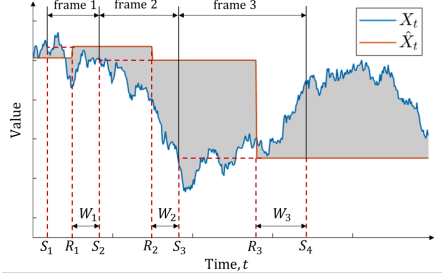


Figure 2: Illustration of the Wiener process and the estimation error. The sampling and reception time-stamp of the k -th sample are denoted by S_k and R_k , respectively. For MMSE estimator, $\hat{X}_t = X_{S_k}, \forall t \in [R_k, R_{k+1})$.

3.1 Markov Decision Reformulation 1

According to [24, Theorem 1], policy π^* 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 . 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 $\mathcal{I}_k := (D_k, (X_{S_k+t} - X_{S_k})_{t \geq 0})$ 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 Π_{recent} . Since for each frame k , the difference $X_{S_k+t} - X_{S_k}$ evolves as a Wiener process that is independent of the past $\{X_{S_{k'}+t} - X_{S_{k'}}\}_{k' < k}$, Problem 1 can be reformulated into the following Markov decision process:

PROBLEM 2 (MARKOV DECISION PROCESS REFORMULATION).

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

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

The proof is provided in Appendix E.

According to [24, Theorem 1], there exists a stationary policy π^* that selects the waiting time W_k using a conditional probability distribution given the recent \mathcal{I}_k . The average MSE $\bar{\mathcal{E}}_{\pi^*}$ achieves mse_{opt} . Therefore, we can restrict our search of π^* within the set of stationary policies. Next, we will reveal the sufficient conditions for the optimal stationary policy.

3.2 Designing π^* with Known \mathbb{P}_D

Let Π_{cons} denote the set of policies that satisfy the sampling frequency constraint. Since π^* achieves the minimum expected time-average MSE among Π_{cons} , we have:

$$\limsup_{K \rightarrow \infty} \frac{\sum_{k=1}^K \mathbb{E} \left[\frac{1}{6} (X_{S_{k+1}} - X_{S_k})^4 \right]}{\sum_{k=1}^K \mathbb{E} [D_k + W_k]} \geq \bar{\mathcal{E}}_{\pi^*} - \bar{D}, \pi \in \Pi_{\text{cons}}. \quad (5)$$

For simplicity, denote $\gamma^* := \bar{\mathcal{E}}_{\pi^*} - \bar{D}$, which is the average cost of the MDP when the optimum policy π^* is used, i.e., $\gamma^* = \limsup_{K \rightarrow \infty} \frac{\sum_{k=1}^K \mathbb{E} \left[\frac{1}{6} (X_{S_{k+1}} - X_{S_k})^4 \right]}{\sum_{k=1}^K \mathbb{E} [D_k + W_k]}$. 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:

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

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

PROBLEM 3 (FUNCTIONAL OPTIMIZATION PROBLEM).

$$\text{mse}_{\text{opt}} = \inf_{\pi \in \Pi} \limsup_{K \rightarrow \infty} \left(\frac{1}{K} \sum_{k=1}^K \mathbb{E} \left[\frac{1}{6} (X_{S_{k+1}} - X_{S_k})^4 \right] - \gamma^* \frac{1}{K} \sum_{k=1}^K \mathbb{E} [(D_k + W_k)] \right), \quad (7a)$$

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

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

$$\mathcal{L}(\pi, \gamma, \nu) \triangleq \limsup_{K \rightarrow \infty} \left(\frac{1}{K} \sum_{k=1}^K \mathbb{E} \left[\frac{1}{6} (X_{S_{k+1}} - X_{S_k})^4 \right] - (\gamma + \nu) \frac{1}{K} \sum_{k=1}^K \mathbb{E} [(S_{k+1} - S_k)] + \nu \frac{1}{f_{\text{max}}} \right). \quad (8)$$

We say that a stationary policy π has a threshold structure, if the waiting time W_k is determined by:

$$W_k = \inf \{w \geq 0 \mid |X_{S_k+D_k+w} - X_{S_k}| \geq \tau\}. \quad (9)$$

Let Z_t be a Wiener process starting from $t = 0$. Let D be the transmission delay following distribution \mathbb{P}_D and the value of the Wiener process at time D is denoted by Z_D . Using the threshold policy (9), the expected frame-length $L_k := D_k + W_k$ and $\frac{1}{6} (X_{S_{k+1}} - X_{S_k})^4$ has the following properties:

LEMMA 2. [24, Corollary 1 Restated]

$$\mathbb{E}[L_k] = \mathbb{E}[\max\{\tau^2, Z_D^2\}], \quad (10a)$$

$$\mathbb{E} \left[\frac{1}{6} (X_{S_{k+1}} - X_{S_k})^4 \right] = \frac{1}{6} \mathbb{E}[\max\{\tau^2, Z_D^2\}^2]. \quad (10b)$$

As is revealed by [24], the optimum policy π^* has a threshold structure as in equation (9). To design an off-line algorithm that can learn the updating threshold τ^* of π^* , we then reveal the necessary conditions that τ^* should satisfy. With slightly abuse of notations, let $\mathcal{L}(\tau, \gamma, \nu)$ denote the expected value of the Lagrange function $\mathcal{L}(\pi, \gamma, \nu)$ when a stationary policy π with threshold τ is used. According to Lemma 2, $\mathcal{L}(\tau, \gamma, \nu)$ can be computed as follows:

$$\mathcal{L}(\tau, \gamma, \nu) = \mathbb{E} \left[\frac{1}{6} \max\{\tau^2, Z_D^2\}^2 \right] - (\gamma + \nu) \mathbb{E}[\max\{\tau^2, Z_D^2\}] + \nu \frac{1}{f_{\max}}. \quad (11)$$

Condition 1: [24, Theorem 5 Restated] Let $\tau(\gamma, \nu)$ be the optimum sampling threshold that minimizes function $\mathcal{L}(\tau, \gamma, \nu)$, which can be computed as follows:

$$\tau(\gamma, \nu) := \arg \inf_{\tau \geq 0} \mathcal{L}(\tau, \gamma, \nu) = \sqrt{3(\gamma + \nu)}. \quad (12)$$

Condition 2: [24, Eq. (123, 125)]

$$\nu^* \left(\mathbb{E}[\max\{3(\gamma^* + \nu^*), Z_D^2\}] - \frac{1}{f_{\max}} \right) = 0, \nu^* \geq 0. \quad (13)$$

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

$$\theta_\pi(\gamma^*) = \frac{1}{6} \mathbb{E}[\max\{\tau^2, Z_D^2\}^2] - \gamma^* \mathbb{E}[\max\{\tau^2, Z_D^2\}] \geq 0. \quad (14)$$

According to (12), inequality (14) holds with equality if and only if π^* with threshold $\tau^* = \sqrt{3(\gamma^* + \nu^*)}$ is used. Adding the CS condition (13) on both sides of (14), the necessary condition for γ^* then becomes:

$$\bar{g}_\nu(\gamma^*) = \theta_{\pi^*}(\gamma^*) = 0, \quad (15)$$

where function $\bar{g}_\nu(\gamma) := \mathbb{E}[g_\nu(\gamma; Z_D)]$ is the expectation of function $g_\nu(\gamma; Z_D)$ defined as follows:

$$g_\nu(\gamma; Z_D) := \frac{1}{6} \max\{3(\gamma + \nu), Z_D^2\}^2 - \gamma \max\{3(\gamma + \nu), Z_D^2\}. \quad (16)$$

As is shown by [24, Theorem 7], the duality gap between $\bar{\mathcal{E}}_{\pi^*}$ and $\sup_{\nu \geq 0} \inf_{\pi} \mathcal{L}(\pi, \gamma^*, \nu)$ is zero, and (15) becomes a necessary and sufficient condition.

3.3 An Online Algorithm π_{online}

When \mathbb{P}_D is unknown but ν^* is known, we can approximate γ^* by solving equation (15) through stochastic approximation [13, 17, 21]. Notice that the role of ν^* is to satisfy the sampling frequency constraint. To achieve this goal, we approximate ν^* by maintaining a sequence $\{U_k\}$ that records the sampling constraint violations up to frame k .

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:** We treat $\nu_k := \frac{1}{V} U_k^+$ as the dual optimizer ν , where $V > 0$ is fixed as a constant. The waiting time W_{k+1} is selected to minimize the Lagrange function (8), and according to the statement after equation (14), W_k is selected by:

$$W_k = \inf\{w \geq 0 \mid |X_{S_k+D_k+w} - X_{S_k}| \geq \sqrt{3(\gamma_k + \nu_k)}\}. \quad (17)$$

2. **Update γ_k :** To search for the root $\gamma > 0$ of equation $\bar{g}_{\nu_k}(\gamma) = 0$, we update γ_k through the Robbins-Monro algorithm [21]. 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:

$$\gamma_{k+1} = (\gamma_k + \eta_k Y_k)^+, \quad (18)$$

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

$$\eta_k = \frac{1}{2\bar{D}_{\text{lb}}} k^{-\alpha}, \alpha \in (0.5, 1]. \quad (19)$$

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:

$$U_{k+1} = U_k + \left(\frac{1}{f_{\max}} - (D_k + W_k) \right). \quad (20)$$

3.4 Theoretical Analysis

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

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 γ^* almost surely, i.e.,

$$\lim_{k \rightarrow \infty} \gamma_k = \gamma^*, \quad \text{w.p.1.} \quad (21)$$

The proof of Theorem 1 is obtained by the ODE method in [13, Chapter 5] and is provided in Appendix A.

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

$$\sup_k \mathbb{E} \left[\frac{|\gamma_k - \gamma^*|^2}{\eta_k} \right] < \infty. \quad (22)$$

Specifically, if $\alpha = 1$ and $\eta_k = \frac{1}{2\bar{D}_{\text{lb}}k}$, then the mean square error decays with rate $\mathbb{E}[(\gamma_k - \gamma^*)^2] = O(1/k)$.

One challenge in the proof of Theorem 2 is that γ_k is unbounded and the second moment of Y_k is unbounded. We notice that Y_k could become very large when γ_k is much larger than the true value γ^* , 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 Section 5.2.

THEOREM 3. The average MSE under policy π_{online} converges to $\bar{\mathcal{E}}_{\pi^*}$ almost surely, i.e.,

$$\limsup_{k \rightarrow \infty} \frac{\int_{t=0}^{S_{k+1}} (X_t - \hat{X}_t)^2 dt}{S_{k+1}} = \bar{\mathcal{E}}_{\pi^*}, \quad \text{w.p.1.} \quad (23)$$

With the mean-square convergence of γ_k , the proof of Theorem 3 is a direct application of the perturbed ODE method [13] and is provided in 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:*

$$\left(\mathbb{E} \left[\int_0^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] - \bar{\mathcal{E}}_{\pi^*} \mathbb{E}[S_k] \right) = O(\ln k). \quad (24)$$

The proof of Corollary 1 is provided in Appendix E.

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:*

$$\inf_{\hat{\gamma}} \sup_{\mathbb{P}} \mathbb{E} [(\hat{\gamma} - \gamma^*(\mathbb{P}))^2] = \Omega(1/k). \quad (25)$$

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 $\pi^*(\mathbb{P})$ and let $\mathcal{P}_u(\mu) := \{\mathbb{P} | p_w(\mathbb{P}) \geq \mu\}$. Specifically, let $p_{w, \text{uni}}^* := \Pr(Z_D^2 \leq 3\gamma_{\text{uni}}^* | D \sim \text{Uni}([0, 1]))$. Let Π_h denote the set of policies which the sampling decision S_k is made based on historical information \mathcal{H}_{k-1} . We have the following result for $\mu \leq p_{w, \text{uni}}^*/2$:

$$\begin{aligned} \inf_{\pi \in \Pi_h} \sup_{\mathbb{P} \in \mathcal{P}_u(\mu)} \left(\mathbb{E} \left[\int_0^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] - \bar{\mathcal{E}}_{\pi^*(\mathbb{P})} \mathbb{E}[S_{k+1}] \right) \\ \geq \frac{1}{2} \mu \cdot \Omega(\ln k). \end{aligned} \quad (26)$$

As the transmission delay \mathbb{P}_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 [32]. The core idea is to construct two distributions $\mathbb{P}_1, \mathbb{P}_2$, whose ℓ_1 distance $|\mathbb{P}_1^{\otimes k} - \mathbb{P}_2^{\otimes k}|_1$ 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 γ^* is closely related to the estimation error $\bar{g}_v(\cdot)$ at a given point. Therefore, the construction of \mathbb{P}_1 and \mathbb{P}_2 for obtaining the converse bound of Hölder smooth functions [32, Chapter 2] are adopted. The proof of inequality (26) is a direct application of the minimax estimation error (25). Detailed proof of Theorem 4 is provided in Section 5.3.

4 SIMULATION RESULTS

In this section, we provide simulation results to verify the theoretic findings and illustrate the performance of our proposed algorithms. We notice that the MSE minimization problem is closely related to the AoI minimization problem, where the AoI at time t , denoted by $A(t) = t - S_{i(t)}$. For signal-ignorant sampling policies (i.e., the sensor cannot always observe the time-varying process), according

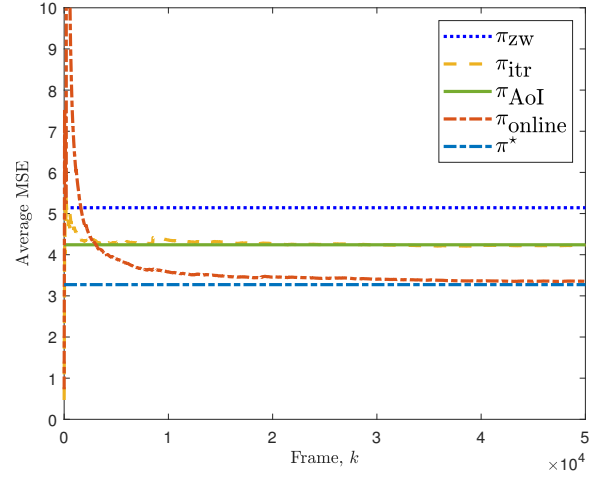


Figure 3: The time average MSE evolution as a function of frame k .

to the analysis in [25, Section IV-B], policies that minimize the average AoI achieves the minimum MSE. Therefore, we choose both offline and online AoI minimization policies (π_{AoI}^* from [25], π_{itr} from [30]) for comparison. To show the convergence of online learning algorithm, we plotted the average MSE performance of the optimum off-line algorithm π^* from [24].

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

$$p(x) := \frac{\mathbb{P}_D(dx)}{dx} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right).$$

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 $\mathbb{E} \left[\int_0^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] - \bar{\mathcal{E}}_{\pi^*} \mathbb{E}[S_{k+1}]$ are depicted in Fig. 4. From Fig. 3, with 5×10^4 samples, the time averaged MSE is almost the same as using the optimum policy. From Fig. 4, 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 π_{online} are depicted in Fig. 5 and Fig. 6, respectively. We set $f_{\text{max}} = \frac{1}{10D}$. From these figures, one can observe that the average MSE of π_{online} is close to the optimum MSE $\bar{\mathcal{E}}_{\pi^*}$ 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.

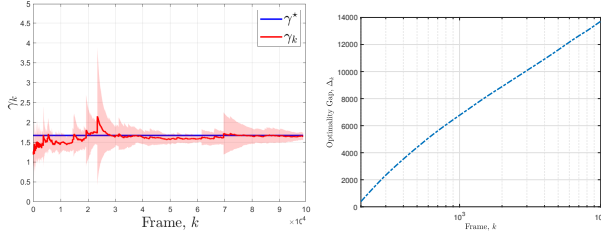


Figure 4: The evolution of the threshold estimate γ_k (left) and the MSE regret $\Delta_k := \mathbb{E} \left[\int_0^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] - \bar{\mathcal{E}}_{\pi^*} \mathbb{E}[S_{k+1}]$ (right).

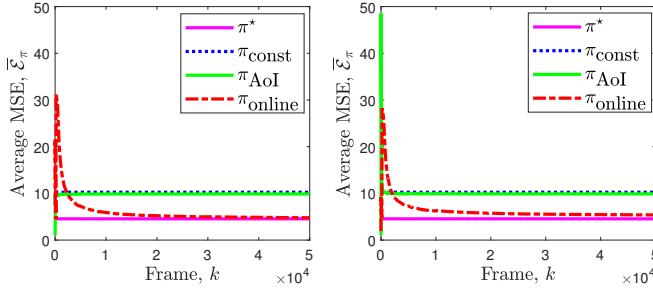


Figure 5: The time average MSE evolution as a function of frame k . (Left: $V = 10$, Right: $V = 1$.)

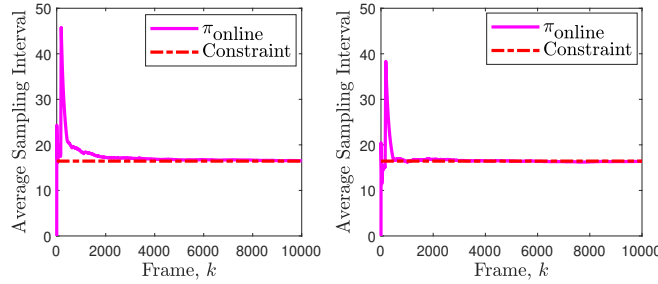


Figure 6: The average sampling interval under different constant V . (Left: $V = 10$, Right: $V = 1$.)

5 PROOFS OF MAIN RESULTS

5.1 Notations and Preliminary Lemmas

In Table 1, we summarize the notations used in the following proofs. Throughout the proofs, we use N_1, N_2, \dots 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 context.

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

$$\frac{1}{6}\bar{D} \leq \gamma^* \leq \frac{1}{2} \frac{M + 2\bar{D} \frac{1}{f_{\max}} + \frac{1}{f_{\max}^2}}{\bar{D} + \frac{1}{f_{\max}}}. \quad (27)$$

Table 1: Notations

Notation	Meaning
Z_t	a Wiener process starting from time 0
l_γ	length of running time using stopping rule $\tau_\gamma := \inf\{t \geq D \mid Z_t \geq \sqrt{3}\gamma\}$
δX_k	$\delta X_k := X_{S_k+D_k} - X_{S_k}$
Q_k	$Q_k := \frac{1}{6} (X_{S_k+D_k} - X_{S_k})^4$
L_k	$L_k := S_{k+1} - S_k = D_k + W_k$, frame length k
E_k	$E_k := \int_{S_k}^{S_{k+1}} (X_t - \hat{X}_t)^2 dt$, cumulative estimation error in frame k
$q(\gamma)$	$q(\gamma) := \frac{1}{6} \mathbb{E} [\max\{3\gamma, Z_D^2\}^2]$, the expectation of Q_k when $\gamma_k = \gamma$
$l(\gamma)$	$l(\gamma) := \mathbb{E}[\max\{3\gamma, Z_D^2\}]$, expected frame length L_k when $\gamma_k = \gamma$
\mathcal{I}_k	$(D_k, (X_t - X_{S_k})_{S_k \leq t \leq S_{k+1}})$, information in frame k
\mathcal{H}_k	$\mathcal{H}_k := \{\mathcal{I}_\kappa\}_{\kappa \leq k}$ historical information up to the end of frame k
$\mathbb{E}_k[\cdot]$	Conditional expectation $\mathbb{E}[\cdot \mathcal{H}_{k-1}]$
t_k	$t_k := \sum_{i=1}^k \eta_i$ or $t_k := \sum_{i=1}^k \epsilon_i$, the cumulative step-sizes depending on the context
$m(t)$	$m(t)$ is the unique k so that $t_k \leq t \leq t_{k+1}$

The proof is provided in Appendix F.

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

$$\mathbb{E}[l_\gamma] \leq 3\gamma + \bar{D}, \quad (28a)$$

$$\mathbb{E}[l_\gamma^2] \leq \frac{10}{3} \left((3\gamma)^2 + 3\sqrt{B} \right), \quad (28b)$$

$$\mathbb{E}[l_\gamma^4] < 4^3 \left((3\gamma)^4 + 105B \right) < \infty. \quad (28c)$$

The proof of Lemma 4 is provided in Appendix G.

LEMMA 5. Function $\bar{g}_0(\gamma) = q(\gamma) - \gamma l(\gamma)$ and has the following properties:

- (i) $\bar{g}_0(\gamma)$ is concave and monotonically decreasing. The second order derivative $-3 \leq \bar{g}_0''(\gamma) \leq 0$.
- (ii) $\bar{g}_0(\gamma^*) = 0$
- (iii) For $\gamma \neq \gamma^*$, $(\gamma - \gamma^*)\bar{g}_0(\gamma) \leq -l(\gamma^*)(\gamma - \gamma^*)^2 \leq 0$.

The proof of Lemma 5 is provided in Appendix H.

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 $E_k = \int_{S_k}^{S_{k+1}} (X_t - \hat{X}_t)^2 dt$ can be bounded as follows:

$$\mathbb{E}_k[E_k^2] = 3(X_{S_k} - X_{S_{k-1}})^2 \sqrt{B} + 12C_1(\gamma_k, B)(X_{S_k} - X_{S_{k-1}})^2 + 3C_2(\gamma_k, B) < \infty, \quad (29)$$

where C_1 and C_2 are fourth order polynomials of γ .

The proof of Corollary 2 is provided in Appendix I.

With these Lemmas we can proceed to prove main results in Section 3:

5.2 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 γ_k by showing the following equation holds:

$$\mathbb{E}_k[V(\gamma_{k+1})] - V(\gamma_k) \leq -\eta_k \bar{D}_{\text{lb}} V(\gamma_k) + \mathcal{O}(\eta_k^2 N_1). \quad (30)$$

Then, based on (30), we can then compute $\mathbb{E}[V(\gamma_k)]$ directly.

Step 1: Bounding the Lyapunov Drift: The analysis is divided into two cases, for $\gamma_k \leq 3\gamma^*$, inequality (30) can be verified easily; for $\gamma_k \geq 3\gamma^*$ we will first establish the relationship between $\mathbb{E}_k[V(\gamma_{k+1})] - V(\gamma_k)$ and $\text{Var}[Y_k]$, then upper bound $\text{Var}[Y_k]$ using the fact that Z_D^2 is sub-Gaussian when D is fourth order bounded.

Case 1: If $\gamma_k \leq 3\gamma^*$, we have:

$$\begin{aligned} & \mathbb{E}_k[V(\gamma_{k+1})] - V(\gamma_k) \\ &= \mathbb{E}_k \left[\frac{1}{2} ((\gamma_k + \eta_k Y_k)^+ - \gamma^*)^2 \right] - \frac{1}{2} (\gamma_k - \gamma^*)^2 \\ &\leq \mathbb{E}_k \left[\frac{1}{2} (\gamma_k - \gamma^* + \eta_k Y_k)^2 - \frac{1}{2} (\gamma_k - \gamma^*)^2 \right] \\ &\stackrel{(a)}{=} (\gamma_k - \gamma^*) \eta_k \bar{g}_0(\gamma_k) \\ &\quad + \frac{1}{2} \eta_k^2 \mathbb{E}_k \left[\left(\frac{1}{6} \max\{3\gamma_k, \delta X_k^2\}^2 - \gamma_k \max\{3\gamma_k, \delta X_k^2\} \right)^2 \right] \\ &\stackrel{(b)}{\leq} -2\eta_k l(\gamma^*) V(\gamma_k) \\ &\quad + \frac{1}{2} \eta_k^2 \left(\frac{1}{36} ((9\gamma^*)^4 + B) + (3\gamma^*)^2 ((9\gamma^*)^2 + 3\sqrt{B}) \right), \end{aligned} \quad (31)$$

where equality (a) is because $\mathbb{E}_k[Y_k] = \mathbb{E}_k[g_0(\gamma_k; \delta X_k)] = \bar{g}_0(\gamma_k)$; inequality (b) is obtained because according to Lemma 5-(iii), $(\gamma_k - \gamma^*) \bar{g}_0(\gamma_k) \leq -l(\gamma^*) (\gamma_k - \gamma^*)^2 = -2l(\gamma^*) V(\gamma_k)$ and the assumption that $\gamma_k \leq 3\gamma^*$.

Case 2: If $\gamma_k \geq 3\gamma^*$, $\gamma_{k+1} = (\gamma_k + \eta_k Y_k)^+$ is truncated into the non-negative real number. We can view the evolution of γ_k as a queueing system, where the queue γ_k is non-negative, and $\eta_k Y_k$ is the arrival rate minus the service rate. Therefore, it is natural to introduce the “unused rate” from [9], which is denoted by $\chi_k := -(\gamma_k + \eta_k Y_k)^+$. If $\chi_k = 0$, $(\gamma_k + \eta_k Y_k) \chi_k = 0 = -\chi_k^2$ and if $\chi_k \geq 0$, $\gamma_k + \eta_k Y_k = -\chi_k$, therefore

$$(\gamma_k + \eta_k Y_k) \chi_k = -\chi_k^2. \quad (32)$$

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

$$-\mathbb{E}_k[\gamma_k + \eta_k Y_k] \leq \mathbb{E}_k[\chi_k]. \quad (33)$$

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

$$\begin{aligned} & \mathbb{E}_k[V(\gamma_{k+1}) - V(\gamma_k)] \\ &= \mathbb{E}_k \left[\frac{1}{2} (\gamma_k - \gamma^* + \eta_k Y_k + \chi_k)^2 - \frac{1}{2} (\gamma_k - \gamma^*)^2 \right] \\ &= \mathbb{E}_k \left[\frac{1}{2} (\gamma_k - \gamma^* + \eta_k Y_k)^2 - \frac{1}{2} (\gamma_k - \gamma^*)^2 \right. \\ &\quad \left. + \frac{1}{2} \chi_k^2 + (\gamma_k + \eta_k Y_k) \chi_k - \gamma^* \chi_k \right] \\ &\stackrel{(c)}{=} \mathbb{E}_k \left[\frac{1}{2} (\gamma_k - \gamma^* + \eta_k Y_k)^2 - \frac{1}{2} (\gamma_k - \gamma^*)^2 - \frac{1}{2} \chi_k^2 - \gamma^* \chi_k \right] \end{aligned}$$

$$\begin{aligned} & \stackrel{(d)}{\leq} \frac{1}{2} (\gamma_k - \gamma^* + \eta_k \mathbb{E}_k[Y_k])^2 - \frac{1}{2} (\gamma_k - \gamma^*)^2 + \frac{1}{2} \eta_k^2 \text{Var}[Y_k] \\ &\quad - \frac{1}{2} \mathbb{E}_k[\chi_k]^2 - \gamma^* \mathbb{E}_k[\chi_k] \\ &= \frac{1}{2} (\gamma_k - \gamma^* + \eta_k \mathbb{E}_k[Y_k])^2 - \frac{1}{2} (\gamma_k - \gamma^*)^2 + \frac{1}{2} \eta_k^2 \text{Var}[Y_k] \\ &\quad - \frac{1}{2} (-\mathbb{E}_k[\chi_k] - \gamma^*)^2 + \frac{1}{2} (\gamma^*)^2, \end{aligned} \quad (34)$$

where equality (c) is because equation (32); inequality (d) is obtained because $\mathbb{E}_k[\chi_k^2] \geq \mathbb{E}_k[\chi_k]^2 \geq 0$;

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 (33), $|\mathbb{E}_k[\chi_k] - \gamma^*| \geq |\gamma_k - \gamma^* + \eta_k \mathbb{E}_k[Y_k]|$. Therefore, inequality (34) can be upper bounded by:

$$\begin{aligned} & \mathbb{E}_k[V(\gamma_{k+1}) - V(\gamma_k)] \\ &\leq -\frac{1}{2} (\gamma_k - \gamma^*)^2 + \frac{1}{2} (\gamma^*)^2 + \frac{1}{2} \eta_k^2 \text{Var}[Y_k] \\ &\stackrel{(e)}{\leq} -\frac{1}{4} (\gamma_k - \gamma^*)^2 + \frac{1}{2} \eta_k^2 \text{Var}[Y_k] \\ &\leq -2\eta_k \bar{D}_{\text{lb}} V(\gamma_k) + \frac{1}{2} \eta_k^2 \text{Var}[Y_k], \end{aligned} \quad (35)$$

where inequality (e) is obtained because $\frac{1}{4} (\gamma_k - \gamma^*)^2 \geq (\gamma^*)^2 \geq \frac{1}{2} (\gamma^*)^2$ by assumption that $\gamma_k \geq 3\gamma^*$ and the last inequality is obtained because $\eta_k \bar{D}_{\text{lb}} \leq \frac{1}{2}$ by the step-size selection rule in equation (19).

Case 2(b): If $\mathbb{E}_k[\gamma_k + \eta_k Y_k] \geq \gamma^*$, considering that $\mathbb{E}_k[Y_k] = \bar{g}_0(\gamma_k) < 0$ for $\gamma_k \geq \gamma^*$, we have $0 > \mathbb{E}_k[\eta_k Y_k] \geq -(\gamma_k - \gamma^*)$. Inequality (34) can be bounded by:

$$\begin{aligned} & \mathbb{E}_k[V(\gamma_{k+1}) - V(\gamma_k)] \\ &\stackrel{(f)}{\leq} \frac{1}{2} (\gamma_k - \gamma^*) (\gamma_k - \gamma^* + \eta_k \mathbb{E}_k[Y_k]) - \frac{1}{2} (\gamma_k - \gamma^*)^2 + \frac{1}{2} \eta_k^2 \text{Var}[Y_k] \\ &\leq \frac{1}{2} \eta_k (\gamma_k - \gamma^*) \bar{g}_0(\gamma_k) + \frac{1}{2} \eta_k^2 \text{Var}[Y_k] \\ &\stackrel{(g)}{\leq} -\frac{1}{2} \eta_k l(\gamma^*) (\gamma_k - \gamma^*)^2 + \frac{1}{2} \eta_k^2 \text{Var}[Y_k] \\ &= -\eta_k l(\gamma^*) V(\gamma_k) + \frac{1}{2} \eta_k^2 \text{Var}[Y_k], \end{aligned} \quad (36)$$

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

For proceed to show inequality (30) for $\gamma_k \geq 3\gamma^*$, we need to upper bound $\text{Var}[Y_k]$ in inequalities (35) and (36). First, we compute the expectation $\mathbb{E}[Y_k]$ as follows:

$$\begin{aligned} \mathbb{E}_k[Y_k] &= \mathbb{E} \left[\frac{1}{6} \max\{3\gamma_k, Z_D^2\}^2 - \gamma_k \max\{3\gamma_k, Z_D^2\} \right] \\ &= -\frac{3}{2} \gamma_k^2 + \mathbb{E} \left[\left(\frac{1}{6} Z_D^4 - \gamma_k Z_D^2 + \frac{3}{2} \gamma_k^2 \right) \mathbb{I}(Z_D^2 \geq 3\gamma_k) \right] \\ &= -\frac{3}{2} \gamma_k^2 + \mathbb{E} \left[\frac{1}{6} (Z_D^2 - 3\gamma_k)^2 \mathbb{I}(Z_D^2 \geq 3\gamma_k) \right] \\ &\leq -\frac{3}{2} \gamma_k^2 + \mathbb{E} \left[\frac{1}{6} (Z_D^2)^2 \right] \\ &\leq -\frac{3}{2} \gamma_k^2 + \frac{1}{2} \mathbb{E}[D^2] \leq -\frac{3}{2} \gamma_k^2 + \frac{1}{2} \sqrt{B}. \end{aligned} \quad (37)$$

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

$$\begin{aligned}
 & \text{Var}[Y_k | \mathcal{H}_{k-1}] \\
 &= \mathbb{E}_k[(Y_k - \mathbb{E}_k[Y_k])^2] \\
 &= \mathbb{E}_k \left[\left(-\frac{3}{2}Y_k^2 - \mathbb{E}_k[Y_k] \right)^2 \mathbb{I}_{(Z_D^2 \leq 3Y_k)} \right] \\
 &\quad + \mathbb{E}_k \left[\left(\frac{1}{6}Z_D^4 - Y_k Z_D^2 + \frac{3}{2}Y_k^2 + \left(-\frac{3}{2}Y_k^2 - \mathbb{E}_k[Y_k] \right) \right)^2 \mathbb{I}_{(Z_D^2 \geq 3Y_k)} \right] \\
 &\stackrel{(h)}{\leq} \frac{1}{4}B + 2\mathbb{E}_k \left[\left(\frac{1}{6}Z_D^4 - Y_k Z_D^2 + \frac{3}{2}Y_k^2 \right)^2 \mathbb{I}_{(Z_D^2 > 3Y_k)} \right] \\
 &\quad + 2\mathbb{E}_k \left[\left(-\frac{3}{2}Y_k^2 - \mathbb{E}_k[Y_k] \right)^2 \mathbb{I}_{(Z_D^2 > 3Y_k)} \right] \\
 &\leq \frac{3}{4}B + \frac{1}{3}\mathbb{E}_k[(Z_D^2 - 3Y_k)^4 \mathbb{I}_{(Z_D^2 \geq 3Y_k)}] \\
 &\leq \frac{3}{4}B + \frac{1}{3}\mathbb{E}[Z_D^8] \leq (35 + \frac{3}{4})B,
 \end{aligned} \tag{38}$$

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

Denote $N_1 := \max\{(35 + \frac{3}{4})B, \frac{1}{36}((9Y^\star)^4 + B) + (3Y^\star)^2((9Y^\star)^2 + 3\sqrt{B})\}$, inequalities (31), (35) and (36) then lead to:

$$\mathbb{E}_k[V(Y_{k+1})] - V(Y_k) \leq -\eta_k \bar{D}_{\text{lb}} V(Y_k) + \eta_k^2 N_1. \tag{39}$$

Step 2: Computing $\mathbb{E}[V(Y_k)]$ through iteration Taking the expectation on both sides of (39), we have:

$$\mathbb{E}[V(Y_{k+1})] \leq (1 - \eta_k \bar{D}_{\text{lb}}) \mathbb{E}[V(Y_k)] + \eta_k^2 N_1. \tag{40}$$

Multiplying inequality (40) from $i = 1$ to k yields:

$$\mathbb{E}[V(Y_{k+1})] \leq \prod_{i=1}^k (1 - \eta_i \bar{D}_{\text{lb}}) V(Y_0) + \sum_{i=1}^k \eta_i^2 N_1 \cdot \prod_{j=i+1}^k (1 - \eta_j \bar{D}_{\text{lb}}). \tag{41}$$

Since the stepsize selected by (19) satisfies

$$\eta_k \rightarrow 0, \liminf_k \min_{k \geq i \geq m(t_k - T)} \frac{\eta_k}{\eta_i} = 1$$

according to [13, p. 343, Eq. (4.8)], term $\prod_{i=1}^k (1 - \eta_i \bar{D}_{\text{lb}}) = O(\eta_k)$. Therefore,

$$\sup_k \mathbb{E} \left[\frac{(Y_k - Y^\star)^2}{\eta_k} \right] = \sup_k \mathbb{E} [2V(\theta_k)/\eta_k] = O(1). \tag{42}$$

This finishes the proof of Theorem 2.

5.3 Proof of Theorem 4

5.3.1 Proof of Inequality (25). Let $\mathbb{P}_1, \mathbb{P}_2$ be two delay distributions and let $\gamma_1^\star, \gamma_2^\star$ be the solution to (15) when $D \sim \mathbb{P}_1$ and $D \sim \mathbb{P}_2$, respectively. Through Le Cam's inequality [34], we have:

$$\inf_{\hat{\gamma}} \sup_{\mathbb{P}} \mathbb{E}[(\hat{\gamma} - \gamma^\star(\mathbb{P}))^2] \geq (\gamma_1^\star - \gamma_2^\star)^2 \cdot (\mathbb{P}_1^{\otimes k} \wedge \mathbb{P}_2^{\otimes k}), \tag{43}$$

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

To use Le Cam's inequality (65), we need to find two distributions \mathbb{P}_1 and \mathbb{P}_2 , whose ℓ_1 distance $|\mathbb{P}_1^{\otimes k} - \mathbb{P}_2^{\otimes k}|_1$ is bounded, and the

difference $(\gamma_1^\star - \gamma_2^\star)^2$ is of order $1/k$. We consider \mathbb{P}_1 to be a uniform distribution on $[0, 1]$ and let γ_1^\star be the optimum ratio of distribution \mathbb{P}_1 . Through Corollary 3, we can obtain a loose upper bound on γ_1^\star as follows:

$$\gamma_1^\star < \frac{1}{2} \frac{\mathbb{E}[D^2]}{\mathbb{E}[D]} = \frac{1}{3}. \tag{44}$$

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

$$\delta = \min\{1 - 3\gamma_1^\star, 1/3, p_{w, \text{uni}}^\star/2\}. \tag{45}$$

Let \mathbb{P}_2 be a probability distribution with probability density function $p_2(x)$ defined as follows:

$$p_2(x) = \begin{cases} 1 - c\sqrt{1/k}, & x \leq \frac{1}{2}\delta; \\ 1, & \frac{1}{2}\delta < x \leq 1 - \frac{1}{2}\delta; \\ 1 + c\sqrt{1/k}, & x > 1 - \frac{1}{2}\delta; \\ 0, & \text{otherwise.} \end{cases} \tag{46}$$

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

Step 1: Lower bounding $\gamma_2^\star - \gamma_1^\star$: For notational simplicity, denote function $h_1(\gamma) := \mathbb{E}_{D \sim \mathbb{P}_1}[\frac{1}{6} \max\{3\gamma, Z_D^2\}^2 - \gamma \max\{3\gamma, Z_D^2\}]$ and $h_2(\gamma) := \mathbb{E}_{D \sim \mathbb{P}_2}[\frac{1}{6} \max\{3\gamma, Z_D^2\}^2 - \gamma \max\{3\gamma, Z_D^2\}]$. According to the definition of \mathbb{P}_2 in (46), for each γ , the difference between $h_1(\gamma)$ and $h_2(\gamma)$ can be computed by:

$$\begin{aligned}
 & h_2(\gamma) - h_1(\gamma) \\
 &= \int_{1-\delta/2}^1 \frac{c}{\sqrt{k}} \mathbb{E} \left[\frac{1}{6} \max\{3\gamma, Z_D^2\}^2 - \gamma \max\{3\gamma, Z_D^2\} | D = x \right] dx \\
 &\quad - \int_0^{\delta/2} \frac{c}{\sqrt{k}} \mathbb{E} \left[\frac{1}{6} \max\{3\gamma, Z_D^2\}^2 - \gamma \max\{3\gamma, Z_D^2\} | D = x \right] dx \\
 &\stackrel{(a)}{=} \int_{1-\delta/2}^1 \frac{c}{\sqrt{k}} \mathbb{E} \left[\frac{1}{6} (Z_D^2 - 3\gamma)^2 \mathbb{I}_{(Z_D^2 \geq 3\gamma)} | D = x \right] dx \\
 &\quad - \int_0^{\delta/2} \frac{c}{\sqrt{k}} \mathbb{E} \left[\frac{1}{6} (Z_D^2 - 3\gamma)^2 \mathbb{I}_{(Z_D^2 \geq 3\gamma)} | D = x \right] dx,
 \end{aligned} \tag{47}$$

where inequality (a) is obtained because

$$\frac{1}{6} \max\{3\gamma, Z_D^2\}^2 - \gamma \max\{3\gamma, Z_D^2\} = -\frac{3}{2}\gamma^2 + \frac{1}{6}(Z_D^2 - 3\gamma)^2 \mathbb{I}_{(Z_D^2 \geq 3\gamma)}. \tag{48}$$

Since γ_1^\star is the optimum ratio for delay distribution \mathbb{P}_1 , we have $h_1(\gamma_1^\star) = 0$. According to equation (47), function $h_2(\gamma_1^\star)$ can be lower bounded by:

$$\begin{aligned}
 & h_2(\gamma_1^\star) \\
 &\stackrel{(b)}{\geq} \frac{c}{\sqrt{k}} \cdot \int_{1-\delta/2}^1 \mathbb{E} \left[\frac{1}{6} (Z_D^2 - 3\gamma_1^\star)^2 \mathbb{I}_{(Z_D^2 \geq 3\gamma_1^\star)} | D = x \right] dx \\
 &\quad - \int_0^{\delta/2} \frac{c}{\sqrt{k}} \frac{1}{2} x^2 dx \\
 &\geq \frac{c}{\sqrt{k}} \cdot \int_{1-\delta/2}^1 \mathbb{E} \left[\frac{1}{6} (Z_D^2 - 3\gamma_1^\star)^2 \mathbb{I}_{(Z_D^2 \geq 3\gamma_1^\star)} | D = x \right] dx - \frac{c}{\sqrt{k}} \frac{1}{6} \left(\frac{\delta}{2} \right)^3.
 \end{aligned} \tag{49}$$

where inequality (b) is because $\mathbb{E}[\frac{1}{6}(Z_D^2 - 3\gamma)^2 \mathbb{I}_{(Z_D^2 \geq 3\gamma)} | D = x] \leq \mathbb{E}[\frac{1}{6}Z_D^4 | D = x] = \frac{1}{2}x^2$.

We then proceed to lower bound $\mathbb{E} \left[\frac{1}{6} (Z_D^2 - 3\gamma_1^*)^2 \mathbb{I}_{(Z_D^2 \geq 3\gamma_1^*)} | D = x \right]$ for each delay realization $x \in [1 - \delta/2, 1]$ as follows:

$$\begin{aligned} & \mathbb{E} \left[\frac{1}{6} (Z_D^2 - 3\gamma_1^*)^2 \mathbb{I}_{(Z_D^2 \geq 3\gamma_1^*)} | D = x \right] \\ & \stackrel{(c)}{\geq} \mathbb{E} \left[\frac{1}{6} (Z_D^2 - 3\gamma_1^*)^2 \mathbb{I}_{(3\gamma_1^* \leq Z_D^2 \leq x)} + \frac{1}{6} (Z_D^2 - x)^2 \mathbb{I}_{(Z_D^2 \geq x)} | D = x \right] \\ & \geq \mathbb{E} \left[\frac{1}{6} (Z_D^2 - 3\gamma_1^*)^2 \mathbb{I}_{(3\gamma_1^* \leq Z_D^2 \leq x)} \right] \\ & \quad + \frac{1}{6} \left(\text{Var}[Z_D^2 | D = x] - x^2 \Pr(Z_D^2 \leq x | D = x) \right) \\ & \stackrel{(d)}{\geq} \frac{1}{6} x^2 \geq \frac{1}{6} (1 - \delta/2)^2, \end{aligned} \quad (50)$$

where inequality (c) is because $\delta \geq 1 - 3\gamma_1^*$ by equation (45), and for the conditional mean $\mathbb{E}[Z_D^2 | D = x] = x \geq 1 - \delta/2 \geq 1 - \delta \geq 3\gamma_1^*$; inequality (d) is because $\text{Var}[Z_D^2 | D = x] = 2x^2$ and $x^2 \Pr(Z_D^2 \leq x) \leq x^2$ 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 (50) into (49) and recall that $\delta < 1$ by definition, we have the lower bound of $h_2(\gamma_1^*)$:

$$h_2(\gamma_1^*) \geq \frac{c}{\sqrt{k}} \frac{\delta}{2} \frac{1}{6} \left(\left(1 - \frac{\delta}{2}\right)^2 - \left(\frac{\delta}{2}\right)^2 \right) \geq \frac{c}{\sqrt{k}} \frac{\delta}{12} (1 - \delta) > 0. \quad (51)$$

By Lemma 5-(i), function $h_2(\cdot)$ 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 $\gamma = \gamma_1^*$.

$$h_2(\gamma_2^*) = h_2(\gamma_1^*) + h_2'(\gamma)(\gamma_2^* - \gamma_1^*), \quad (52)$$

where $\gamma \in [\gamma_1^*, \gamma_2^*]$. Therefore, γ_2^* can be computed by:

$$\gamma_2^* - \gamma_1^* = -\frac{h_2(\gamma_1^*)}{h_2'(\gamma)}. \quad (53)$$

To lower bound γ_2^* , we will first find a loose upper bound of γ_2^* using Lemma 3:

$$\gamma_2^* \leq \frac{1}{2} \frac{\mathbb{E}_{D \sim \mathbb{P}_2}[D^2]}{\mathbb{E}_{D \sim \mathbb{P}_2}[D]} \leq \frac{1}{2} \left(\frac{1}{3} + \delta \cdot c\sqrt{1/k} \right), \quad (54)$$

Therefore, since $\delta < 1/3$, we have $|h_2'(\gamma)| \leq |h_2'(\gamma_2^*)| = \mathbb{E}[\max\{3\gamma_2^*, Z_D^2\}] \leq \bar{D} + 3\gamma_2^* \leq 1 + \frac{1}{2} + \frac{3}{2}c\sqrt{\frac{1}{k}}\delta \leq 2$. Then by inequality (51), we have

$$\gamma_2^* - \gamma_1^* \geq \frac{-h_2(\gamma_1^*)}{h_2'(\gamma_2^*)} \geq \frac{1}{24} (1 - \delta) \delta c \sqrt{\frac{1}{k}}. \quad (55)$$

Step 2: Lower bounding $\mathbb{P}_1^{\otimes k} \wedge \mathbb{P}_2^{\otimes k}$: Let $|\mathbb{P} - \mathbb{Q}| = \int_{\Omega} |\mathrm{d}\mathbb{P} - \mathrm{d}\mathbb{Q}|$ be the ℓ_1 distance between probability distribution \mathbb{P} and \mathbb{Q} . Then

$$\begin{aligned} \mathbb{P}_1^{\otimes k} \wedge \mathbb{P}_2^{\otimes k} &= \int \min\{\mathbb{P}_1^{\otimes k}(\mathrm{d}x), \mathbb{P}_2^{\otimes k}(\mathrm{d}x)\} \\ &= \int \mathbb{P}_1^{\otimes k}(\mathrm{d}x) \cdot \left(1 - \frac{(\mathbb{P}_2^{\otimes k}(\mathrm{d}x) - \mathbb{P}_1^{\otimes k}(\mathrm{d}x))^+}{\mathbb{P}_1^{\otimes k}(\mathrm{d}x)} \right) \\ &= 1 - \int (\mathbb{P}_2^{\otimes k}(\mathrm{d}x) - \mathbb{P}_1^{\otimes k}(\mathrm{d}x))^+ \\ &= 1 - \frac{1}{2} |\mathbb{P}_1^{\otimes k} - \mathbb{P}_2^{\otimes k}|_1. \end{aligned} \quad (56)$$

Equality (56) enables us to lower bound $\mathbb{P}_1^{\otimes k} \wedge \mathbb{P}_2^{\otimes k}$ by upper bounding the ℓ_1 distance $|\mathbb{P}_1^{\otimes k} - \mathbb{P}_2^{\otimes k}|_1$, which is done through the Pinsker's inequality:

$$\begin{aligned} & \frac{1}{2} |\mathbb{P}_1^{\otimes k} - \mathbb{P}_2^{\otimes k}|_1 \\ & \leq \sqrt{\frac{1}{2} D_{\text{KL}}(\mathbb{P}_2^{\otimes k} || \mathbb{P}_1^{\otimes k})} \\ & = \sqrt{\frac{1}{2} k D_{\text{KL}}(\mathbb{P}_2 || \mathbb{P}_1)} \\ & \stackrel{(e)}{\leq} \sqrt{\frac{1}{2} k \int_0^1 p_2(x) \ln p_2(x) \mathrm{d}x} \\ & \stackrel{(f)}{\leq} \sqrt{\frac{1}{2} k \int_0^1 \left(p_2(x) - 1 + \frac{1}{\min\{p_2(x), 1\}} (p_2(x) - 1)^2 \right) \mathrm{d}x} \\ & \stackrel{(g)}{\leq} \sqrt{\frac{1}{2} k \frac{1}{\inf_{0 \leq d \leq 1} p_2(d)} \int_0^1 (p_2(x) - 1)^2 \mathrm{d}x} \\ & \leq \sqrt{\frac{1}{2} k \frac{1}{1 - c\sqrt{1/k}} \delta \frac{c^2}{k}} \leq \sqrt{\delta c^2}, \end{aligned} \quad (57)$$

where inequality (e) is because the density function $p_1(x) = 1$ for uniform distribution, therefore $D_{\text{KL}}(\mathbb{P}_2 || \mathbb{P}_1) = \int_0^1 p_2(x) \ln p_2(x) \mathrm{d}x$; inequality (f) is because function $g(t) := (t \ln t)$ is convex, its derivative $g(t)'' = 1/t$, therefore, through Taylor expansion we have $g(t) \leq g(1) + (t-1) + \frac{1}{2} \frac{1}{\min\{t, 1\}} (t-1)^2 = (t-1) + \frac{1}{2} \frac{1}{\min\{t, 1\}} (t-1)^2$; inequality (g) is because $\int_0^1 p_2(x) \mathrm{d}x = 1$.

By choosing $c = 1/2$ and recall that $\delta < 1$, inequality (57) can be upper bounded by:

$$\frac{1}{2} |\mathbb{P}_1^{\otimes k} - \mathbb{P}_2^{\otimes k}|_1 \leq \frac{1}{2}. \quad (58)$$

Plugging (58) into (56) yields:

$$\mathbb{P}_1^{\otimes k} \wedge \mathbb{P}_2^{\otimes k} \geq \frac{1}{2}. \quad (59)$$

Finally, plugging (59) and (55) into the Le Cam's inequality (43) finishes the proof of inequality (25):

$$\inf_{\hat{\gamma}} \sup_{\mathbb{P}} (\hat{\gamma} - \gamma^*(\mathbb{P}))^2 \geq \frac{1}{2} \left(\frac{1}{24} (1 - \delta) \delta p_{w, \text{uni}}^* \right)^2 \cdot \frac{1}{k}. \quad (60)$$

5.3.2 Proof of Inequality (26). The proof is divided into three steps: consider a delay distribution $\mathbb{P} \in \mathcal{P}_u(\mu)$, first we will show for each sample policy π with a random sampling interval τ , let $l_\pi := \mathbb{E}[\tau] = \mathbb{E}[Z_\tau^2]$ denote the expected running length, the following inequality holds:

$$\mathbb{E} \left[\int_{t=0}^{\tau} Z_t^2 \mathrm{d}t \right] - \gamma^*(\mathbb{P}) \mathbb{E}[\tau] \geq \frac{1}{6} p_w(\mathbb{P}) (l_\pi - l^*(\mathbb{P}))^2, \quad (61)$$

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. Next, We will show that given k samples $\delta X_1, \dots, \delta X_k \stackrel{\text{i.i.d.}}{\sim} Z_D$, where $D \sim \mathbb{P}$, the minimax estimation error satisfies:

$$\inf_{\hat{l}} \sup_{\mathbb{P} \in \mathcal{P}_u(\mu)} \mathbb{E} \left[\left(\hat{l} - l^*(\mathbb{P}) \right)^2 \right] \geq N \cdot \frac{1}{k}, \quad (62)$$

where N is a constant independent of k and μ with expressions provided in equation (60).

Finally, notice that:

$$\begin{aligned} & \mathbb{E} \left[\int_0^{S_k} (X_t - \hat{X}_t)^2 dt - (\gamma^\star + \bar{D}) S_k \right] \\ &= \sum_{k'=1}^k \mathbb{E} \left[(X_{S_{k'+1}} - X_{S_{k'}})^2 D_k + \frac{1}{6} (X_{S_{k'+1}} - X_{S_{k'}})^4 \right] \\ & \quad - (\gamma^\star + \bar{D}) \sum_{k'=1}^k \mathbb{E}[S_{k'+1} - S_{k'}] \\ & \geq \frac{1}{6} p_w(\mathbb{P}) \sum_{k'=1}^k (\mathbb{E}[L_k] - l^\star)^2. \end{aligned} \quad (63)$$

Take $\inf_{\pi} \sup_{\mathbb{P} \in \mathcal{P}_w(\mu)}$ on both sides of inequality (63), we have:

$$\begin{aligned} & \min_{\pi} \max_{\mathbb{P} \in \mathcal{P}_u(\mu)} \mathbb{E} \left[\int_0^{S_k} (X_t - \hat{X}_t)^2 dt - (\gamma^\star + \bar{D}) S_k \right] \\ & \geq \frac{1}{6} \mu \inf_{\hat{l}} \sup_{\mathbb{P} \in \mathcal{P}_u(\mu)} \mathbb{E}[(\hat{l} - l^\star(\mathbb{P}))^2] \geq \frac{1}{6} \mu N \sum_{k'=1}^k \frac{1}{k'} \geq \Omega(\ln k). \end{aligned} \quad (64)$$

Proof of inequality (61) follows similar ideas as [26, Lemma 4]. Details are provided in Appendix K due to space limitations. The proof of the minimax risk bound (62) is based on Le Cam's two point method as follows:

5.3.3 Proof of inequality (62). Let $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{P}_w(\mu)$, through Le Cam's inequality [34], we have:

$$\inf_{\hat{l}} \sup_{\mathbb{P} \in \mathcal{P}_u(\mu)} \mathbb{E} \left[(\hat{l} - l^\star(\mathbb{P}))^2 \right] \geq (l_1^\star - l_2^\star)^2 \cdot \left(\mathbb{P}_1^{\otimes k} \wedge \mathbb{P}_2^{\otimes k} \right). \quad (65)$$

Similarly to the proof of bounding $(\hat{\gamma} - \gamma^\star(\mathbb{P}))$, let \mathbb{P}_1 be a uniform distribution and \mathbb{P}_2 through the density function in equation (46). For $\mu \leq p_{w, \text{uni}}^\star/2$, it is easy to show that $p_w(\mathbb{P}_2) \in \mathcal{P}_u(\mu)$ as follows:

$$\begin{aligned} & \Pr(Z_D^2 \leq 3\gamma_2^\star | D \sim \mathbb{P}_2) \\ &= \int_0^\infty \Pr(Z_D^2 \leq 3\gamma_2^\star | D = x) p_2(x) dx \\ &= \int_0^\infty \Pr(Z_D^2 \leq 3\gamma_2^\star | D = x) p_1(x) dx \\ & \quad - \int_0^{\delta/2} \Pr(Z_D^2 \leq 3\gamma_2^\star | D = x) \frac{c}{\sqrt{k}} dx \\ & \quad + \int_{1-\delta/2}^1 \Pr(Z_D^2 \leq 3\gamma_2^\star | D = x) \frac{c}{\sqrt{k}} dx \\ & \stackrel{(i)}{\geq} \int_0^\infty \Pr(Z_D^2 \leq 3\gamma_1^\star | D = x) p_1(x) dx - c/\sqrt{k}\delta \\ & \geq p_{w, \text{uni}}^\star - c/\sqrt{k}\delta \\ & \stackrel{(j)}{\geq} p_{w, \text{uni}}^\star/2. \end{aligned} \quad (66)$$

where inequality (i) holds because $\gamma_1^\star \leq \gamma_2^\star$ from inequality (55); inequality (j) holds because $\delta < p_{w, \text{uni}}^\star/2$ by definition from equation

(45). The difference between $l_2^\star - l_1^\star$ can be computed by:

$$\begin{aligned} l_2^\star - l_1^\star &= \int_0^1 \mathbb{E} [\max\{3\gamma_2^\star, Z_D^2\} | D = x] dx \\ & \quad + \int_{1-\delta/2}^1 \frac{c}{\sqrt{k}} \mathbb{E} [\max\{3\gamma_2^\star, Z_D^2\} | D = x] dx \\ & \quad - \int_0^{\delta/2} \frac{c}{\sqrt{k}} \mathbb{E} [\max\{3\gamma_2^\star, Z_D^2\} | D = x] dx \\ & \quad - \int_0^1 \mathbb{E} [\max\{3\gamma_1^\star, Z_D^2\} | D = x] dx. \end{aligned} \quad (67)$$

Notice that if $x_1 \geq x_2$,

$$\mathbb{E}[\max\{3\gamma, Z_D^2\} | D = x_1] - \mathbb{E}[\max\{3\gamma, Z_D^2\} | D = x_2] \geq 0. \quad (68)$$

Therefore, inequality (67) can be bounded by:

$$\begin{aligned} l_2^\star - l_1^\star &\geq \int_0^1 \mathbb{E} [\max\{3\gamma_2^\star, Z_D^2\} | D = x] dx \\ & \quad - \int_0^1 \mathbb{E} [\max\{3\gamma_1^\star, Z_D^2\} | D = x] dx \\ & \geq 3(\gamma_2^\star - \gamma_1^\star) \mathbb{E}_{D \sim \mathbb{P}_1} [\Pr(Z_D^2 \leq 3\gamma_1^\star)] \\ & \stackrel{(g)}{\geq} \frac{1}{24} (1 - \delta) \delta c p_{w, \text{uni}}^\star \sqrt{\frac{1}{k}}, \end{aligned} \quad (69)$$

where inequality (g) is obtained by equation (55).

Finally, plugging (58) and (69) into Le Cam's inequality (65) finishes the proof of inequality (62):

$$\inf_{\hat{l}} \sup_{\mathbb{P} \in \mathcal{P}_w(\mu)} (\hat{l} - l^\star(\mathbb{P}))^2 \geq \frac{1}{2} \left(\frac{1}{48} (1 - \delta) \delta c p_{w, \text{uni}}^\star \right)^2 \cdot \frac{1}{k}. \quad (70)$$

6 CONCLUSIONS

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 best $\Omega(\ln k)$. Numerical simulation results validate the convergence behaviors of the proposed algorithm.

ACKNOWLEDGMENTS

The authors are grateful to Prof. Pengkun Yang and Prof. Yuhao Wang for helpful discussions on the proof of Theorem 4. The work of H. Tang and L. Tassiulas was supported by the NSF CNS-2112562 AI Institute for Edge Computing Leveraging Next Generation Networks (Athena) and the ONR N00014-19-1-2566. The work of Y. Sun was supported by the NSF grant CCF-1813078 and the ARO grant W911NF-21-1-0244.

REFERENCES

- [1] Mohamed A. Abd-Elmagid, Harpreet S. Dhillon, and Nikolaos Pappas. 2020. A Reinforcement Learning Framework for Optimizing Age of Information in RF-Powered Communication Systems. *IEEE Transactions on Communications* 68, 8 (2020), 4747–4760. <https://doi.org/10.1109/TCOMM.2020.2991992>

- [2] Ahmed Arafa, Roy D. Yates, and H. Vincent Poor. 2019. Timely Cloud Computing: Preemption and Waiting. In *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. 528–535. <https://doi.org/10.1109/ALLERTON.2019.8919891>
- [3] Eray Unsal Atay, Igor Kadota, and Eytan Modiano. 2020. Aging Bandits: Regret Analysis and Order-Optimal Learning Algorithm for Wireless Networks with Stochastic Arrivals. *arXiv:2012.08682 [eess.SY]*
- [4] Subhankar Banerjee, Rajarshi Bhattacharjee, and Abhishek Sinha. 2020. Fundamental Limits of Age-of-Information in Stationary and Non-stationary Environments. In *2020 IEEE International Symposium on Information Theory (ISIT)*. 1741–1746. <https://doi.org/10.1109/ISIT44484.2020.9174054>
- [5] Kavya Bhandari, Santosh Fatale, Urvidh Narula, Sharayu Moharir, and Manjesh Kumar Hanawal. 2020. Age-of-Information Bandits. In *2020 18th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT)*. 1–8.
- [6] Elif Tuğçe Ceran, Deniz Gündüz, and András György. 2019. Reinforcement Learning to Minimize Age of Information with an Energy Harvesting Sensor with HARQ and Sensing Cost. In *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. 656–661. <https://doi.org/10.1109/INFOCOMW.2019.8845182>
- [7] Elif Tuğçe Ceran, Deniz Gündüz, and András György. 2021. A Reinforcement Learning Approach to Age of Information in Multi-User Networks With HARQ. *IEEE Journal on Selected Areas in Communications* 39, 5 (2021), 1412–1426. <https://doi.org/10.1109/JSAC.2021.3065057>
- [8] Rick Durrett. [n.d.]. *Probability: Theory and Examples*.
- [9] Atilla Eryilmaz and Rayadurgam Srikant. 2012. Asymptotically tight steady-state queue length bounds implied by drift conditions. *Queueing Systems* 72, 3 (2012), 311–359.
- [10] B. Hajek, K. Mitzel, and S. Yang. 2003. Paging and registration in cellular networks: jointly optimal policies and an iterative algorithm. In *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No.03CH37428)*, Vol. 1. 524–532 vol.1. <https://doi.org/10.1109/INFOCOM.2003.1208703>
- [11] Clement Kam, Sastry Kompella, and Anthony Ephremides. 2019. Learning to Sample a Signal through an Unknown System for Minimum AoI. In *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. 177–182. <https://doi.org/10.1109/INFOCOMW.2019.8845311>
- [12] Sanjit Kaul, Roy Yates, and Marco Gruteser. 2012. Real-time status: How often should one update?. In *2012 Proceedings IEEE INFOCOM*. 2731–2735. <https://doi.org/10.1109/INFOCOM.2012.6195689>
- [13] Harold J. Kushner and G. George Yin. 2003. *Stochastic Approximation and Recursive Algorithms and Applications*. Springer New York, New York, NY. https://doi.org/10.1007/0-387-21769-X_5
- [14] Shiyang Leng and Aylin Yener. 2019. Age of Information Minimization for Wireless Ad Hoc Networks: A Deep Reinforcement Learning Approach. In *2019 IEEE Global Communications Conference (GLOBECOM)*. 1–6. <https://doi.org/10.1109/GLOBECOM38437.2019.9013454>
- [15] Bin Li. 2021. Efficient Learning-based Scheduling for Information Freshness in Wireless Networks. In *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*. *arXiv:2101.00257 [cs.PF]*
- [16] Ashutosh Nayyar, Tamer Başar, Demosthenis Teneketzis, and Venugopal V. Veeravalli. 2013. Optimal Strategies for Communication and Remote Estimation With an Energy Harvesting Sensor. *IEEE Trans. Automat. Control* 58, 9 (2013), 2246–2260. <https://doi.org/10.1109/TAC.2013.2254615>
- [17] Michael J. Neely. 2021. Fast Learning for Renewal Optimization in Online Task Scheduling. , 44 pages. <http://jmlr.org/papers/v22/20-813.html>
- [18] Tasmeen Zaman Ornee and Yin Sun. 2021. Sampling and Remote Estimation for the Ornstein-Uhlenbeck Process Through Queues: Age of Information and Beyond. *IEEE/ACM Transactions on Networking* 29, 5 (2021), 1962–1975. <https://doi.org/10.1109/TNET.2021.3078137>
- [19] H Vincent Poor. 2013. *An introduction to signal detection and estimation*. Springer Science & Business Media.
- [20] Daniel Revuz and Marc Yor. [n.d.]. *Continuous Martingales and Brownian Motion*.
- [21] Herbert Robbins and Sutton Monro. 1951. A Stochastic Approximation Method. *The Annals of Mathematical Statistics* 22, 3 (1951), 400 – 407.
- [22] Sheldon M Ross. 2013. *Applied probability models with optimization applications*. Courier Corporation.
- [23] Touraj Soleymani, Sandra Hirche, and John S. Baras. 2016. Optimal Information Control in Cyber-Physical Systems. *IFAC-PapersOnLine* 49, 22 (2016), 1–6. 6th IFAC Workshop on Distributed Estimation and Control in Networked Systems NECSYS 2016.
- [24] Yin Sun, Yury Polyanskiy, and Elif Uysal. 2020. Sampling of the Wiener Process for Remote Estimation Over a Channel With Random Delay. *IEEE Transactions on Information Theory* 66, 2 (2020), 1118–1135. <https://doi.org/10.1109/TIT.2019.2937336>
- [25] Yin Sun, Elif Uysal-Biyikoglu, Roy D. Yates, C. Emre Koksall, and Ness B. Shroff. 2017. Update or Wait: How to Keep Your Data Fresh. *IEEE Transactions on Information Theory* 63, 11 (2017), 7492–7508. <https://doi.org/10.1109/TIT.2017.2735804>
- [26] Haoyue Tang, Yuchao Chen, Jintao Wang, Jingzhou Sun, and Jian Song. 2022. Sending Timely Status Updates through Channel with Random Delay via Online Learning. In *IEEE INFOCOM 2022 - IEEE Conference on Computer Communications (INFOCOM 2022)*. London, United Kingdom (Great Britain).
- [27] Haoyue Tang, Yuchao Chen, Jintao Wang, Pengkun Yang, and Leandros Tassiulas. 2022. Age Optimal Sampling Under Unknown Delay Statistics. *arXiv preprint arXiv:2202.13367* (2022).
- [28] Haoyue Tang, Yin Sun, and Leandros Tassiulas. 2022. Sampling of the Wiener Process for Remote Estimation over a Channel with Unknown Delay Statistics. *Technical Report* (2022). <https://www.dropbox.com/s/a00sxn4we42tnvl/wiener-mobihoc.pdf?dl=0>
- [29] Vishrant Tripathi and Eytan Modiano. 2021. An Online Learning Approach to Optimizing Time-Varying Costs of AoI. In *Proceedings of the Twenty-Second International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing (Shanghai, China) (MobiHoc '21)*. Association for Computing Machinery, New York, NY, USA, 241–250. <https://doi.org/10.1145/3466772.3467053>
- [30] Cho-Hsin Tsai and Chih-Chun Wang. 2020. Age-of-Information Revisited: Two-way Delay and Distribution-oblivious Online Algorithm. In *2020 IEEE International Symposium on Information Theory (ISIT)*. 1782–1787. <https://doi.org/10.1109/ISIT44484.2020.9174306>
- [31] Cho-Hsin Tsai and Chih-Chun Wang. 2022. Unifying AoI Minimization and Remote Estimation—Optimal Sensor/Controller Coordination With Random Two-Way Delay. *IEEE/ACM Transactions on Networking* 30, 1 (2022), 229–242. <https://doi.org/10.1109/TNET.2021.3111495>
- [32] Alexandre B. Tsybakov. 2008. *Introduction to Nonparametric Estimation* (1st ed.). Springer Publishing Company, Incorporated.
- [33] Roy D. Yates. 2015. Lazy is timely: Status updates by an energy harvesting source. In *2015 IEEE International Symposium on Information Theory (ISIT)*. 3008–3012. <https://doi.org/10.1109/ISIT.2015.7283009>
- [34] Bin Yu. 1997. Assouad, Fano, and Le Cam. In *Festschrift for Lucien Le Cam*. Springer, 423–435.

A PROOF OF THEOREM 1

To show that γ_k converges to γ^* almost surely, we use the sufficient condition from [13, p.190, Theorem 7.1]. 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 $t \geq 0$, let $m(t)$ be the unique $k \in \mathbb{N}^+$ so that $t_k \leq t < t_{k+1}$. Without a sampling constraint, $v_k \equiv 0, \forall k$. Then the update rule for γ_k from equation (18) can be rewritten in the following recursive form:

$$\gamma_{k+1} = \gamma_k + \eta_k \underbrace{(\bar{g}_0(\gamma_k) + \delta M_k)}_{Y_k}, \quad (71)$$

where we recall that $Y_k = \frac{1}{6} \max\{3\gamma_k, \delta X_k^2\}^2 - \gamma_k \max\{3\gamma_k, \delta X_k^2\}$ and δM_k is the difference between realization and the conditional expectation $\mathbb{E}_k[Y_k] = \bar{g}_0(\gamma_k)$. Notice that the difference $\delta M_k := Y_k - \mathbb{E}_k[Y_k]$ depends only on the transmission delay and the Wiener process evolution $(X_t - X_{S_k})$ in frame k and γ_k , which can be predictable given \mathcal{H}_{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 $N < \infty$, $\sup_k \mathbb{E}[|Y_k| \mathbb{I}_{(|\gamma_k| \leq N)}]$ is bounded, i.e.,

$$\begin{aligned} & \sup_k \mathbb{E}[|Y_k| \mathbb{I}_{(|\gamma_k| \leq N)}] \\ & \leq \sup_k \mathbb{E}\left[\frac{1}{6} \max\{3\gamma_k, \delta X_k^2\}^2 \cdot \mathbb{I}_{(|\gamma_k| \leq N)}\right] + \sup_k \mathbb{E}[\gamma_k \max\{3\gamma_k, \delta X_k^2\} \mathbb{I}_{(|\gamma_k| \leq N)}] \\ & < \frac{1}{6} (9N^2 + \mathbb{E}[Z_D^4]) + N \cdot (3N + \mathbb{E}[Z_D^2]) \stackrel{(a)}{\leq} \infty. \end{aligned} \quad (72)$$

where inequality (a) is because $\mathbb{E}[Z_D^4] = 3\mathbb{E}[D^2] \leq 3\sqrt{\mathbb{E}[D^4]} < \infty$ and $\mathbb{E}[Z_D^2] = \mathbb{E}[D] < \infty$.

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

(1.3) The martingale sequence $\delta M_k \mathbb{I}_{(|\gamma_k| \leq N)}$ can be bounded as follows:

$$\begin{aligned} & \text{Var}[\delta M_k \mathbb{I}_{(|\gamma_k| \leq N)}] \leq \mathbb{E}[Y_k^2 \mathbb{I}_{(|\gamma_k| \leq N)}] \\ & \leq \mathbb{E}\left[\left(\frac{1}{6} \max\{3\gamma_k, \delta X_k^2\}^2 - \gamma_k \max\{3\gamma_k, \delta X_k^2\}\right)^2 \mathbb{I}_{(|\gamma_k| \leq N)}\right] \\ & \stackrel{(b)}{\leq} \mathbb{E}\left[2 \left(\left(\frac{1}{6} \max\{3\gamma_k, \delta X_k^2\}^2\right)^2 + \gamma_k^2 \left(\max\{3\gamma_k, \delta X_k^2\}\right)^2\right) \mathbb{I}_{(|\gamma_k| \leq N)}\right] \\ & \stackrel{(c)}{\leq} 2 \left(\frac{1}{36} (3N)^4 + 105B + N^2(9N^2 + 3\sqrt{B})\right) \leq N_1. \end{aligned} \quad (73)$$

where inequality (b) is because $\mathbb{E}[(a-b)^2] \leq \mathbb{E}[2(a^2 + b^2)]$; inequality (c) is because $\delta X_t \sim Z_D$ is a Wiener process starting from $t = 0$ and therefore, $\mathbb{E}[Z_D^8] = 105\mathbb{E}[D^4] \leq 105B$.

Since sequence $\delta M_k \mathbb{I}_{(|\gamma_k| \leq N)}$ has mean zero. Its value only depends on γ_k and the Wiener process evolution in frame k . The correlation $\mathbb{E}[\delta M_i \mathbb{I}_{(|\gamma_i| \leq N)} \cdot \delta M_j \mathbb{I}_{(|\gamma_j| \leq N)}] = 0, \forall i \neq j$. As the variance of $\delta M_k \mathbb{I}_{(|\gamma_k| \leq N)}$ is bounded in inequality (73), the stepsizes η_k satisfies $\sum_{k=1}^{\infty} \frac{1}{2D_{lb}} k^{-2\alpha} = \frac{1}{2lb} \left(1 + \frac{1}{2\alpha-1}\right)$, according to [13, Chapter 5, Eq. (5.3.18)], for each $\mu > 0$ we have

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \max_{0 \leq t \leq T} \left| \sum_{i=m(jT)}^{m(jT+t)-1} \epsilon \delta M_i \mathbb{I}_{(|\gamma_i| \leq N)} \right| \geq \mu \right) = 0. \quad (74)$$

Let $\gamma_k(\omega)$ be the value of ratio γ on sample path ω . Recall that the stepsizes $\{\eta_k\}$ selected in (19) satisfies $\sum_{k=1}^{\infty} \eta_k = \infty, \sum_{k=1}^{\infty} \eta_k^2 \leq \infty$. According to [13, p.170, Theorem 1.2], with probability 1, the limit $\lim_{k \rightarrow \infty} \theta_k(\omega)$ are trajectories of the following ordinary differential equation (ODE), i.e.,

$$\dot{\gamma} = \bar{g}_0(\gamma). \quad (75)$$

The next step is to show the solution of the ODE in equation (75) converges to γ^* as time diverges. Equation (15) implies $\bar{g}_0(\gamma^*) = 0$ and therefore, γ^* is an equilibrium point of ODE (75). 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;

$$\dot{V} = (\gamma - \gamma^*) \dot{\gamma} = (\gamma - \gamma^*) \bar{g}_0(\gamma). \quad (76)$$

According to Lemma 5-(iii), $\dot{V} = (\gamma - \gamma^*) \bar{g}_0(\gamma) < 0$, the stability of γ^* is verified through Lyapunov theorem.

B PROOF OF THEOREM 3

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

$$\liminf_{k \rightarrow \infty} \frac{1}{k} \sum_{k'=1}^k (D_{k'} + W_{k'}) \geq \liminf_{k \rightarrow \infty} \frac{1}{k} \sum_{k'=1}^k D_{k'} = \bar{D} > 0, \text{ w.p.1.} \quad (77)$$

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

$$\begin{aligned} \theta_k &:= \frac{1}{k} \int_0^{S_{k+1}} (X_t - \hat{X}_t)^2 dt - (\gamma^* + \bar{D})S_{k+1} \\ &= \frac{1}{k} \sum_{k'=1}^k \left(\int_{S_{k'}}^{S_{k'+1}} (X_t - \hat{X}_t)^2 dt - (\gamma^* + \bar{D})L_{k'} \right) \end{aligned} \quad (78)$$

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 θ_k in the following recursive form:

$$\begin{aligned} \theta_k &= \frac{1}{k} \left((k-1)\theta_{k-1} + E_k - (\gamma^* + \bar{D})L_k \right) \\ &= \theta_{k-1} + \frac{1}{k} \left(-\theta_{k-1} + E_k - (\gamma^* + \bar{D})L_k \right). \end{aligned} \quad (79)$$

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:

$$\begin{aligned} G_k &= -\theta_{k-1} + \int_{S_k}^{S_k+D_k} (X_t - X_{S_{k-1}})^2 dt + \int_{S_k+D_k}^{S_k+D_k+W_k} (X_t - X_{S_k})^2 dt - (\gamma^* + \bar{D})L_k \\ &= -\theta_{k-1} + \int_{S_k}^{S_k+D_k} (X_t - X_{S_k} + X_{S_k} - X_{S_{k-1}})^2 dt + \int_{S_k}^{S_{k+1}} (X_t - X_{S_k})^2 dt - (\gamma^* + \bar{D})L_k \\ &= -\theta_{k-1} + \underbrace{(X_{S_k} - X_{S_{k-1}})^2 D_k}_{=:G_{k,1}} + \underbrace{2(X_{S_k} - X_{S_{k-1}}) \cdot \int_{S_k}^{S_k+D_k} (X_t - X_{S_k}) dt}_{=:G_{k,2}} + \underbrace{\int_{S_k}^{S_{k+1}} (X_t - X_{S_k})^2 dt}_{=:G_{k,3}} - \underbrace{(\gamma^* + \bar{D})L_k}_{=:G_{k,4}}. \end{aligned} \quad (80)$$

Give historical transmissions \mathcal{H}_{k-1} , γ_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 $G_{k,1}, \dots, G_{k,4}$ can be computed as follows:

$$\mathbb{E}_k [G_{k,1}] = \bar{D}(X_{S_k} - X_{S_{k-1}})^2, \quad (81a)$$

$$\mathbb{E}_k [G_{k,2}] = 0, \quad (81b)$$

$$\mathbb{E}_k [G_{k,3}] = \frac{1}{6} \mathbb{E}_k [\max\{3\gamma_k, Z_D^2\}^2] = q(\gamma_k), \quad (81c)$$

$$\mathbb{E}_k [G_{k,4}] = (\gamma^* + \bar{D}) \mathbb{E}_k [\max\{3\gamma_k, Z_D^2\}] = (\gamma^* + \bar{D})l(\gamma_k). \quad (81d)$$

where equation (81a) is because D_k is independent of $X_{S_k} - X_{S_{k-1}}$; equation (81b) 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 (81c) and (81d) is because of Lemma 2. With equation (81a)-(81d), given historical transmissions \mathcal{H}_{k-1} , we can compute the conditional expectation of G_k as follows:

$$\begin{aligned} &\mathbb{E}_k [G_k] \\ &= \mathbb{E}_k [-\theta_{k-1} + G_{k,1} + 2G_{k,2} + G_{k,3} - G_{k,4}] \\ &= -\theta_{k-1} + (X_{S_k} - X_{S_{k-1}})^2 \bar{D} + q(\gamma_k) - (\gamma^* + \bar{D})l(\gamma_k) \\ &= -\theta_{k-1} + q(\gamma_k) - \gamma_k l(\gamma_k) + \underbrace{\bar{D}(l(\gamma_{k-1}) - l(\gamma_k))}_{=: \beta_{k,1}} + \underbrace{\bar{D}((X_{S_k} - X_{S_{k-1}})^2 - l(\gamma_{k-1}))}_{=: \beta_{k,2}} + \underbrace{(\gamma_k - \gamma^*)l(\gamma_k)}_{=: \beta_{k,3}}. \end{aligned} \quad (82)$$

Denote function

$$f(\theta, \gamma) := -\theta + \mathbb{E} \left[\frac{1}{6} \max\{3\gamma, Z_D^2\}^2 - \gamma \max\{3\gamma, Z_D^2\} \right], \quad (83)$$

and let function $\bar{f}(\cdot)$ be:

$$\bar{f}(\theta) := f(\theta, \gamma^*). \quad (84)$$

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

$$\sum_k \epsilon_k = \infty, \sum_k \epsilon_k^2 < \infty. \quad (85)$$

With $\epsilon_k, \beta_{k,1}, \beta_{k,2}$ and δM_k , the recursive equation (80) can be rewritten as follows:

$$\theta_k = \theta_{k-1} + \epsilon_k (f(\theta_{k-1}, \gamma_k) + \beta_{k,1} + \beta_{k,2} + \beta_{k,3} + \delta M_k). \quad (86)$$

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 $t_{m(t)} \leq t < t_{m(t)+1}$. We then state the following characteristics of G_k and δM_k , detailed proofs are in Appendix C:

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

(2.1) For each constant N , $\sup_k \mathbb{E} [|G_k| \mathbb{I}_{(|\theta_k| \leq N)}] < \infty$.

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

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

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \max_{0 \leq t \leq T} \left| \sum_{i=m(jT)}^{m(jT+t)-1} \epsilon_i (f(\theta, \gamma_i) - \bar{f}(\theta)) \right| \geq \mu \right) = 0. \quad (87)$$

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

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \max_{0 \leq t \leq T} \left| \sum_{i=k}^j \epsilon_i \delta M_i \right| \geq \mu \right) = 0. \quad (88)$$

(2.5) The bias sequence satisfies:

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \max_{0 \leq t \leq T} \left| \sum_{i=m(jT)}^{m(jT+t)-1} \epsilon_i (\beta_{i,1} + \beta_{i,2} + \beta_{i,3}) \right| \geq \mu \right) = 0. \quad (89)$$

(2.6) For each θ , function f can be bounded as follows:

$$f(\theta, \gamma) = \bar{f}(\theta) + \rho(\gamma), \quad (90)$$

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

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq n} \sum_{i=m(j\tau)}^{m(j\tau+\tau)-1} |\epsilon_i \rho(\gamma_k)| \right) = 0. \quad (91)$$

(2.7) For each θ_1, θ_2 , the difference

$$|f(\theta_1, \gamma) - f(\theta_2, \gamma)| = |\theta_1 - \theta_2|. \quad (92)$$

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

Denote $\theta_k(\omega)$ as the time averaged MSE up to frame k of sample path ω . Then according to [13, p.166, Theorem 1.1], with probability 1, sequence $\{\theta_k(\omega)\}$ converges to some limit set of the ODE

$$\dot{\theta} = \bar{f}(\theta) = -\theta. \quad (93)$$

Because $\bar{f}(0) = 0$, the minimum error $\theta = 0$ is an equilibrium point of the ODE in equation (93). Moreover, as $\bar{f}(\cdot)$ is a monotonic decreasing function, it can be easily verified through Lyapunov stability criterion that 0 is a unique stability point of the ODE (93). Therefore, θ_k converges to 0 with probability 1, and the time averaged MSE converges to $\bar{\mathcal{E}}_{\pi^*}$ with probability 1.

C PROOF OF CLAIM 1

Before we starts to prove each condition in Claim 1, we provide the following corollary from Theorem 2:

COROLLARY 3. There exists a $\Gamma < \infty$ so that $\mathbb{E}[(\gamma_k - \gamma^*)^2 / \eta_k] < \Gamma, \forall k$. Recall that the step-sizes is selected to be $\eta_k = \frac{1}{4D_{lb}k^\alpha}$, where $\alpha \in (0.5, 1]$, we then have:

$$\mathbb{E}[(\gamma_k - \gamma^*)^2] \leq \frac{D_{lb}\Gamma}{2k^\alpha} < \infty, \mathbb{E}[\gamma_k^2] \leq 2(\mathbb{E}[(\gamma_k - \gamma^*)^2] + (\gamma^*)^2) < \infty. \quad (94)$$

Through Cauchy-Schwarz inequality, we have:

$$\mathbb{E}[|\gamma_k - \gamma^*|] \leq \sqrt{\mathbb{E}[(\gamma_k - \gamma^*)^2]} \leq \sqrt{\frac{D_{lb}\Gamma}{2k^\alpha}}. \quad (95)$$

(2.1): According to the definition of G_k from equation (79), the expectation $\mathbb{E} [|G_k| \mathbb{I}_{(|\theta_k| \leq N)}]$ can be upper bound as follows:

$$\mathbb{E} [|G_k| \mathbb{I}_{(|\theta_k| \leq N)}] \leq \mathbb{E} [|\theta_k| \mathbb{I}_{(|\theta_k| \leq N)}] + \mathbb{E} \left[\int_{t=S_k}^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] + \mathbb{E} [\gamma_k L_k]. \quad (96)$$

The first term on the RHS of inequality (96) satisfies

$$\sup_k \mathbb{E} [|\theta_k| \mathbb{I}_{(|\theta_k| \leq N)}] \leq N < \infty. \quad (97)$$

The expectation of the second term can be computed as follows,

$$\begin{aligned} & \mathbb{E} \left[\int_{t=S_k}^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] \\ &= \mathbb{E} \left[\int_{t=S_k}^{S_k+D_k} (X_t - X_{S_k} + X_{S_k} - X_{S_{k-1}})^2 dt + \int_{S_k+D_k}^{S_k+D_k+W_k} (X_t - X_{S_k})^2 dt \right] \\ &= \mathbb{E} [(X_{S_k} - X_{S_{k-1}})^2 D_k] + 2\mathbb{E} \left[(X_{S_k} - X_{S_{k-1}}) \int_{t=S_k}^{S_k+D_k} (X_t - X_{S_k}) dt \right] + \mathbb{E} \left[\int_{t=S_k}^{S_{k+1}} (X_t - X_{S_k})^2 dt \right] \\ &= \mathbb{E} [\max\{3\gamma_{k-1}, Z_D^2\}] \bar{D} + \frac{1}{6} \mathbb{E} [\max\{3\gamma_k, Z_D^2\}^2] \\ &\leq \mathbb{E} [3\gamma_{k-1}] \bar{D} + \frac{1}{6} \mathbb{E} [(3\gamma_k)^2] + \bar{D} B^{1/4} + \frac{1}{2} \sqrt{B}. \end{aligned} \quad (98)$$

Inequality (94) and (95) implies inequality (98) is bounded for all k . Therefore, the second term on the RHS of inequality (96) can be upper bounded as follows:

$$\sup_k \mathbb{E} \left[\int_{t=S_k}^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] = \sup_k \left(\mathbb{E} [\max\{3\gamma_{k-1}, Z_D^2\}] \bar{D} + \frac{1}{6} \mathbb{E} [\max\{3\gamma_k, Z_D^2\}^2] \right) < \infty. \quad (99)$$

Similarly, since $\mathbb{E}[\gamma_k^2] < \infty$ is bounded by (94) is bounded, we can upper bound the third term on the RHS of inequality (96) as follows:

$$\sup_k \mathbb{E} [\gamma_k L_k] = \sup_k \mathbb{E} [\gamma_k \max\{3\gamma_k, Z_D^2\}] < \infty. \quad (100)$$

Taking the supremum of inequality (96) and then plugging equality (97)-(99) into the inequality verifies Claim (2.1).

Notice that statement (2.3)-(2.7) has similar forms,

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i \psi_i \right| \geq \mu \right) = 0. \quad (101)$$

where ψ_k can be the bias term $\beta_{k,i}$, the martingale sequence δM_k or the difference $f(\theta, \gamma_k) - \bar{f}(\theta)$ and $\rho(\gamma_k)$. We then provide the following lemma: (101) to hold true:

LEMMA 6. If one of the following condition holds, then (101) holds:

(S.1) ψ_k is a martingale sequence and $\sup_k \mathbb{E}[\psi_k^2] < \infty$. The correlation satisfies $\mathbb{E}[\psi_i \psi_j] = 0, \forall i \neq j$.

(S.2) $\mathbb{E}[|\psi_k|] = O(k^{-\zeta}), \zeta > 0$.

PROOF. If condition (S.1) holds, since $\epsilon_k = \frac{1}{k}$ satisfies $\sum_k \epsilon_k^2 < \infty$, equality (101) holds because of [13, p. 172, example 3].

If condition (S.2) holds, there exists a Ψ so that $\mathbb{E}[\psi_k] = \Psi k^{-\zeta}$. For each $\mu > 0$, we first upper bound $\Pr \left(\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i \psi_i \right| \geq \mu \right)$ for each k as follows:

$$\Pr \left(\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i \psi_i \right| \geq \mu \right) \leq \Pr \left(\sum_{i=k}^{\infty} \epsilon_i |\psi_i| \geq \mu \right) \stackrel{(a)}{\leq} \frac{1}{\mu} \mathbb{E} \left[\sum_{i=k}^{\infty} i^{-1} |\psi_i| \right] \stackrel{(b)}{\leq} \frac{\Psi}{\mu} \left(\sum_{i=k}^{\infty} i^{-1-\zeta} \right) = \frac{\Psi}{\mu \zeta} (k-1)^{-\zeta}. \quad (102)$$

where inequality (a) is from the Markov inequality; inequality (b) is from statement (S.2). Finally, taking the limit of (102) yields (101). \square

(2.2): Since function $f(\theta, \gamma; \delta X) := -\theta + \frac{1}{6} \max\{3\gamma, \delta X^2\}^2 - \gamma \max\{3\gamma, \delta X^2\}$ is continuous for each δX , the expectation $\bar{f}(\theta) = \mathbb{E}[f(\theta, \gamma^*; \delta X^2)]$ is continuous for θ .

(2.3): Recall the definition of $f(\theta, \gamma)$ and $\bar{f}(\gamma)$ from equation (83), (84). The absolute difference between $f(\theta, \gamma)$ and $\bar{f}(\theta)$ can be upper bounded by:

$$|f(\theta, \gamma) - \bar{f}(\theta)| = |\bar{g}_0(\gamma) - \bar{g}_0(\gamma^*)| \stackrel{(a)}{\leq} 3(\gamma - \gamma^*)^2 + 3|\gamma - \gamma^*| \bar{D}, \quad (103)$$

where inequality (a) is because function $\bar{g}_0(\gamma)$ is concave and $|\bar{g}_0''(\gamma)| < 3$ according to Lemma 5-(i). Therefore $\mathbb{E}[|f(\theta, \gamma) - \bar{f}(\theta)|] = O(k^{-\alpha/2})$, which satisfies statement (S.2) in Lemma 6. This verifies inequality (87).

(2.4): The difference $\delta M_k = \delta M_{k,1} + 2\delta M_{k,2} + \delta M_{k,3} - \delta M_{k,4}$ consists of four parts. Through the union bound, the probability that $\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i \delta M_i \right| \geq \mu$ can be upper bounded by:

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i \delta M_i \right| \geq \mu \right) \leq \sum_{a=1}^4 \lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i \delta M_{i,a} \right| \geq \mu/5 \right). \quad (104)$$

We will then show that each item on the RHS of inequality (104) has limit 0. The first term $\delta M_{k,1} = (X_{S_k} - X_{S_{k-1}})^2 (D_k - \bar{D})$. Since $D_k - \bar{D}$ depends only on the delay in frame k and has mean zero, term $\mathbb{E}[\delta M_{k,1} \delta M_{k+i,1}] = 0, \forall i > 0$. The second moment of $\delta M_{k,1}$ can be upper bounded as follows:

$$\mathbb{E} \left[(X_{S_k} - X_{S_{k-1}})^4 (D_k - \bar{D})^2 \right] = \mathbb{E} \left[(X_{S_k} - X_{S_{k-1}})^4 \right] \text{Var}[D^2] \leq \mathbb{E} \left[\max\{3\gamma_k, Z_D^2\}^2 \right] \mathbb{E}[D^2]. \quad (105)$$

By Theorem 2, the expectation $\mathbb{E}[(\gamma_k - \gamma^\star)^2 / \eta_k] = O(1)$. Since $\eta_k \rightarrow 0$, $\mathbb{E}[\gamma_k^2]$ is bounded. Therefore, $\sup_k \mathbb{E}[\delta M_{k,1}^2] \leq \infty$. Then according to [13, p.142, Eq. (5.3.18)]

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i \delta M_{i,1} \right| \geq \mu/5 \right) = 0. \quad (106)$$

Similarly, recall that $\delta M_{k,2} = (X_{S_k} - X_{S_{k-1}}) \cdot \left(\int_{S_k}^{S_k+D_k} (X_t - X_{S_k}) dt \right)$. Sequence $\delta M_{k,2}$ is a martingale sequence with mean zero. Moreover, $\mathbb{E}[M_{k,2} M_{k+i,2}] = 0, \forall i \geq 1$. The variance $\text{Var}[\delta M_{k,2}]$ can be bounded as follows:

$$\text{Var}[\delta M_{k,2}] = \mathbb{E}[\delta M_{k,2}^2] = \mathbb{E} \left[(X_{S_k} - X_{S_{k-1}})^2 \right] \cdot \mathbb{E} \left[\left(\int_{t=S_k}^{S_k+D_k} (X_t - X_{S_k}) dt \right)^2 \right] = \mathbb{E} \left[\max\{3\gamma_{k-1}, Z_D^2\} \right] \cdot \mathbb{E}[D^2]. \quad (107)$$

Inequality (95) upper bounds $\mathbb{E}[\max\{3\gamma_{k-1}, Z_D^2\}]$ and verifies (S.1) in Lemma 6. Therefore, we have:

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i \delta M_{i,2} \right| \geq \mu/5 \right) = 0. \quad (108)$$

It can be verified that the sequence $\delta M_{k,3} = \int_{S_k}^{S_{k+1}} (X_t - X_{S_k})^2 dt - \mathbb{E}_k \left[\int_{S_k}^{S_{k+1}} (X_t - X_{S_k})^2 dt \right]$ is a martingale sequence. It then remains to upper bound its variance, which is as follows:

$$\begin{aligned} \text{Var}[\delta M_{k,3}] &= \mathbb{E} \left[\left(\int_{S_k}^{S_k+D_k} (X_t - X_{S_k})^2 dt - \int_{S_k+D_k}^{S_{k+1}} (X_t - X_{S_k})^2 dt \right)^2 \right] \\ &\stackrel{(c)}{\leq} 2\mathbb{E} \left[\left(\int_{S_k}^{S_k+D_k} (X_t - X_{S_k})^2 dt \right)^2 \right] + 2\mathbb{E} \left[\left(\int_{S_k+D_k}^{S_{k+1}} (X_t - X_{S_k})^2 dt \right)^2 \right] \\ &\stackrel{(d)}{\leq} 2\mathbb{E} \left[\left(\int_{S_k}^{S_k+D_k} (X_t - X_{S_k})^2 dt \right)^2 \right] + 2\mathbb{E} \left[\Pr(Z_D^2 \geq 3\gamma_k) (3\gamma_k)^2 \mathbb{E}_k[L_k^2] \right] \\ &\stackrel{(e)}{\leq} N_1 + \mathbb{E} \left[\mathbb{E}[Z_D^4] \left(\frac{10}{3} (3\gamma_k)^2 + 3\sqrt{B} \right) \right], \end{aligned} \quad (109)$$

where inequality (c) is because $\mathbb{E}[(a-b)^2] \leq 2\mathbb{E}[a^2 + b^2]$; inequality (d) is because if $L_k \geq D_k$, then $(X_t - X_{S_k})^2 \leq 3\gamma_k$ for $t \in [S_k + D_k, S_{k+1}]$; inequality (e) is because through Markov inequality $\Pr(Z_D^2 \geq 3\gamma_k) \leq \mathbb{E}[Z_D^4] / (3\gamma_k)^2$ and $\mathbb{E}_k[L_k^2] \leq (\frac{10}{3} (3\gamma_k)^2 + 3\sqrt{B})$ from Lemma 4. Since $\mathbb{E}[Z_D^4] \leq 3\mathbb{E}[D^2] < 3\sqrt{B}$ and $\mathbb{E}[\gamma_k^2]$ is bounded according to inequality (94), $\text{Var}[\delta M_{k,3}]$ is bounded according to inequality (109). Condition (S.1) in Lemma 6 is satisfied and we have

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i \delta M_{i,3} \right| \geq \mu/5 \right) = 0. \quad (110)$$

Following similar approaches, the second order expansion of the fourth term is bounded, i.e.,

$$\text{Var}[\delta M_{k,4}] \leq \mathbb{E}_k[G_{k,4}] \leq (\gamma^\star + \bar{D})^2 \cdot \frac{10}{3} (3\gamma_k^2 + 3\sqrt{B}). \quad (111)$$

Again using Lemma 6 condition (S.1), we have:

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i \delta M_{i,4} \right| \geq \mu/5 \right) = 0. \quad (112)$$

Plugging inequalities (106), (108), (110) and (112) into (104) completes the proof of (89).

(2.5): Through the union bound we have:

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i (\beta_{i,1} + \beta_{i,2} + \beta_{i,3}) \right| \geq \mu \right) \leq \sum_{a=1}^3 \lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i \beta_{i,a} \right| \geq \mu/3 \right). \quad (113)$$

From the definition of $\beta_{k,1}$ in equation (82), $\beta_{k,1} = \bar{D} \left(\mathbb{E}[3\gamma_{k-1}, Z_D^2] - \mathbb{E}[3\gamma_k, Z_D^2] \right) = \bar{D} \left(\mathbb{E}[3\gamma_{k-1}, Z_D^2] - \mathbb{E}[3\gamma^*, Z_D^2] + \mathbb{E}[3\gamma^*, Z_D^2] - \mathbb{E}[3\gamma_k, Z_D^2] \right)$. Then we can upper bound the expectation of $|\beta_{k,1}|$ as follows:

$$\mathbb{E}[|\beta_{k,1}|] \leq 3\bar{D}\mathbb{E}[|\gamma_{k-1} - \gamma^*| + |\gamma_k - \gamma^*|] \stackrel{(f)}{=} O(k^{-\alpha}),$$

where equality (f) is from inequality (95). Since $\alpha \in (0.5, 1]$, which satisfies condition (S.2) in Lemma 6. We have:

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i \beta_{i,1} \right| \geq \mu/3 \right) = 0. \quad (114)$$

Next we upper bound $\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i \beta_{i,2} \right| \geq \mu/3 \right) = 0$. First, $\beta_{k,2}$ has mean zero because $\mathbb{E}_k[(X_{S_{k-1}+D_{k-1}} - X_{S_{k-1}})^2] = \max\{3\gamma_{k-1}, Z_D^2\}$, and its variance is upper bounded by

$$\text{Var} \left[\bar{D} \left((X_{S_k} - X_{S_{k-1}})^2 - l(\gamma_{k-1}) \right) \right] = \bar{D}^2 \mathbb{E}[\max\{3\gamma_{k-1}, Z_D^2\}^2] \stackrel{(g)}{<} \infty, \quad (115)$$

where inequality (g) is due to (94). Using condition (S.1) Lemma 6, we have:

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \left| \sum_{i=k}^j \epsilon_i \beta_{i,2} \right| \geq \mu/3 \right) = 0. \quad (116)$$

Finally the third bias term satisfies $\mathbb{E}[|\beta_{k,3}|] = \mathbb{E}[l(\gamma_k) \cdot |\gamma_k - \gamma^*|] \leq \sqrt{\mathbb{E}[\max\{3\gamma_k, Z_D^2\}]^2 \mathbb{E}[(\gamma_k - \gamma^*)^2]}$. Since $\mathbb{E}[\max\{3\gamma_k, Z_D^2\}] \leq \mathbb{E}[3\gamma_k] + \bar{D}$ is bounded according to (95) and $\mathbb{E}[(\gamma_k - \gamma^*)^2] = O(k^{-\alpha})$, $\mathbb{E}[|\beta_{k,3}|] = O(k^{-\alpha})$. Condition (S.2) in Lemma 6 is verified and we have

$$\lim_{k \rightarrow \infty} \Pr \left(\sup_{j \geq k} \sum_{i=k}^j \epsilon_i \beta_{i,3} \geq \mu/3 \right) = 0. \quad (117)$$

Plugging (114), (116) and (117) into (113) verifies statement (2.5).

(2.6): Function $\rho(\gamma) \leq l(\gamma^*)|\gamma - \gamma^*| + \frac{3}{2}(\gamma - \gamma^*)^2$. Since $\mathbb{E}[|\gamma - \gamma^*|^2] = O(k^{-\alpha})$ and $\mathbb{E}[|\gamma - \gamma^*|] = O(k^{-\alpha/2})$ satisfies condition (S.2) in Lemma 6, statement (2.6) is verified.

D PROOF OF COROLLARY 1

The cumulative MSE up to the beginning of frame $k+1$ can be decomposed into:

$$\begin{aligned} & \mathbb{E} \left[\int_0^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] \\ & \stackrel{(a)}{=} \sum_{k'=1}^k \mathbb{E} \left[\int_{S_{k'}}^{S_{k'}+D_{k'}} (X_t - X_{S_{k'-1}})^2 dt + \int_{S_{k'}+D_{k'}}^{S_{k'+1}} (X_t - X_{S_{k'}})^2 dt \right] \\ & = \sum_{k'=1}^k \mathbb{E} \left[\int_{S_{k'}}^{S_{k'}+D_{k'}} (X_t - X_{S_{k'}} + X_{S_{k'}} - X_{S_{k'-1}})^2 dt + \int_{S_{k'}+D_{k'}}^{S_{k'+1}} (X_t - X_{S_{k'}})^2 dt \right] \\ & = \sum_{k'=1}^k \mathbb{E} \left[\int_{S_{k'}}^{S_{k'}+D_{k'}} \left((X_t - X_{S_{k'}})^2 + 2(X_t - X_{S_{k'}})(X_{S_{k'}} - X_{S_{k'-1}}) + (X_{S_{k'}} - X_{S_{k'-1}})^2 \right) dt + \int_{S_{k'}+D_{k'}}^{S_{k'+1}} (X_t - X_{S_{k'}})^2 dt \right] \\ & \stackrel{(b)}{=} \sum_{k'=1}^k \mathbb{E} \left[(X_{S_{k'}} - X_{S_{k'-1}})^2 D_{k'} + \int_{S_{k'}}^{S_{k'+1}} (X_t - X_{S_{k'}})^2 dt \right] \\ & \stackrel{(c)}{=} \sum_{k'=1}^k \mathbb{E}[(X_{S_{k'}} - X_{S_{k'-1}})^2] \bar{D} + \sum_{k'=1}^k \frac{1}{6} \mathbb{E}[(X_{S_{k'+1}} - X_{S_{k'}})^4], \end{aligned} \quad (118)$$

where equality (a) is because $\hat{X}_t = X_{S_{k-1}}, \forall t \in [S_k + D_k]$ and $\hat{X}_t = X_{S_k}, \forall t \in [S_k + D_k, S_{k+1})$; equality (b) is because $\mathbb{E}[X_t - X_{S_{k'}}] = 0$ and because $X_{S_{k'}} - X_{S_{k'-1}}$ is independent of $D_{k'}$; equality (c) is because $D_{k'}$ is independent of $(X_{S_{k'}} - X_{S_{k'-1}})$.

With equation (118), we proceed to bound the difference $\mathbb{E} \left[\int_0^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] - (\gamma^* + \bar{D}) \mathbb{E} \left[\sum_{k'=1}^k L_{k'} \right]$ as follows:

$$\begin{aligned}
 & \mathbb{E} \left[\int_0^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] - (\gamma^* + \bar{D}) \mathbb{E} \left[\sum_{k'=1}^k L_{k'} \right] \\
 & \stackrel{(a)}{=} \sum_{k'=1}^k \mathbb{E}[(X_{S_{k'+1}} - X_{S_{k'}})^2] \bar{D} + \sum_{k'=1}^k \frac{1}{6} \mathbb{E}[(X_{S_{k'+1}} - X_{S_{k'}})^4] - \gamma^* \mathbb{E}[(X_{S_{k'+1}} - X_{S_{k'}})^2] - \bar{D} \sum_{k'=1}^k \mathbb{E}[(X_{S_{k'+1}} - X_{S_{k'}})^2] \\
 & = \sum_{k'=1}^k (\mathbb{E}[q(\gamma_{k'}) - \gamma^* l(\gamma_{k'})]) \\
 & = \sum_{k'=1}^k \mathbb{E} [q(\gamma_{k'}) - \gamma_{k'} l(\gamma_{k'}) + (\gamma_{k'} - \gamma^*) l(\gamma_{k'})] \\
 & \stackrel{(b)}{\leq} \sum_{k'=1}^k \mathbb{E} [q(\gamma^*) - \gamma_{k'} l(\gamma^*) + (\gamma_{k'} - \gamma^*) l(\gamma_{k'})] \\
 & \stackrel{(c)}{=} \sum_{k'=1}^k \mathbb{E}[(\gamma^* - \gamma_{k'})(l(\gamma^*) - l(\gamma_{k'}))] \\
 & \stackrel{(d)}{\leq} 3 \sum_{k'=1}^k \mathbb{E}[(\gamma_{k'} - \gamma^*)^2], \tag{119}
 \end{aligned}$$

where equality (a) is because of equation (118) and by martingale stopping theorem, $\mathbb{E}[(X_{S_{k'+1}} - X_{S_{k'}})^2] = \mathbb{E}[L_{k'}]$; inequality (b) is because choosing threshold $\gamma_{k'}$ minimizes function $q(\gamma) - \gamma l(\gamma)$; inequality (c) is obtained because $q(\gamma^*) - \gamma_{k'} l(\gamma^*) = q(\gamma^*) - \gamma^* l(\gamma^*) + (\gamma^* - \gamma_{k'}) l(\gamma^*)$ and $q(\gamma^*) - \gamma^* l(\gamma^*) = 0$ according to Lemma 5-(ii); inequality (d) is obtained because $|l(\gamma^*) - l(\gamma_{k'})| = |\mathbb{E}[\max\{3\gamma^*, Z_D^2\}] - \mathbb{E}[\max\{3\gamma_{k'}, Z_D^2\}]| \leq 3|\gamma_{k'} - \gamma^*|$.

E PROOF OF LEMMA 1

PROOF. First we will turn the time-averaged MSE computation into frame-level computation. For stationary policy π that decides sampling time S_{k+1} only on information \mathcal{I}_k , tuple $\{(I_k, (S_{k+1} - S_k))\}$ is a generative sequence. Recall that $E_k = \int_{S_k}^{S_{k+1}} (X_t - X_{S_k})^2 dt$ and $L_k := S_{k+1} - S_k$ are the cumulative estimation error and length of frame k , which are both generative because policy π is stationary. Therefore, sequence $\{\frac{1}{K} \mathbb{E} [\sum_{k=1}^K E_k]\}$ and $\{\frac{1}{K} \mathbb{E} [\sum_{k=1}^K L_k]\}$ have limits. Then according to the renewal reward theory [22], the time averaged MSE can be computed by:

$$\begin{aligned}
 & \limsup_{T \rightarrow \infty} \mathbb{E} \left[\int_{t=0}^T (X_t - \hat{X}_t)^2 dt \right] \\
 & = \limsup_{K \rightarrow \infty} \frac{\mathbb{E} \left[\sum_{k=1}^K \int_{S_k}^{S_{k+1}} (X_t - X_{S_k})^2 dt \right]}{\mathbb{E} \left[\sum_{k=1}^K (S_{k+1} - S_k) \right]} \\
 & = \limsup_{K \rightarrow \infty} \frac{\sum_{k=1}^K \mathbb{E} \left[\int_{S_k}^{S_{k+1}} (X_t - X_{S_k})^2 dt \right]}{\sum_{k=1}^K \mathbb{E} [(S_{k+1} - S_k)]}. \tag{120}
 \end{aligned}$$

To simplify the computation of equation 120, we will first introduce the following lemma:

LEMMA 7 (LEMMA 6, [24]RESTATED). *Let Z_t be a Wiener process starting from time zero, let τ be a stopping time of Z_t , we have:*

$$\frac{1}{6} \mathbb{E} [Z_\tau^4] = \mathbb{E} \left[\int_0^\tau Z_t^2 dt \right] \tag{121}$$

Using Lemma 7, we can then compute the expected cumulative estimation error during interval $[S_k, R_k]$. Notice that during the interval, the k -th sample has not been received. Therefore, the estimation error $X_t - \hat{X}_t = X_t - X_{S_{k-1}}$ can be viewed as a Wiener process starting from time S_{k-1} . We can decouple and compute the cumulative estimation error as follows:

$$\mathbb{E} \left[\int_{S_k}^{R_k} (X_t - \hat{X}_t)^2 dt \right]$$

$$\begin{aligned}
&= \mathbb{E} \left[\int_{S_{k-1}}^{R_k} (X_t - X_{S_{k-1}})^2 dt \right] - \mathbb{E} \left[\int_{S_{k-1}}^{S_k} (X_t - X_{S_{k-1}})^2 dt \right] \\
&= \frac{1}{6} \mathbb{E} [(X_{R_k} - X_{S_{k-1}})^4] - \frac{1}{6} \mathbb{E} [(X_{S_k} - X_{S_{k-1}})^4].
\end{aligned} \tag{122}$$

Similarly, we can then obtain that:

$$\mathbb{E} \left[\int_{R_k}^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] = \frac{1}{6} \mathbb{E} [(X_{S_{k+1}} - X_{S_k})^4] - \frac{1}{6} \mathbb{E} [(X_{R_k} - X_{S_k})^4]. \tag{123}$$

Notice that the transmission delay D_k is i.i.d across all slots, since we only focus on stationary policies whose waiting time W_k relies only on recent information $\mathcal{I}_k := \{Y_k, (W_{S_k+t} - W_{S_k}), \forall t \geq 0\}$, we have:

$$\mathbb{E}[(X_{S_k} - X_{S_{k-1}})^4] = \mathbb{E}[(X_{S_{k+1}} - X_{S_k})^4]$$

Therefore, by summing up (122) and (123), for any policy π that makes decisions only on \mathcal{I}_k , the expected cumulative estimation error in frame k can be computed by:

$$\begin{aligned}
\mathbb{E}[E_k] &= \mathbb{E} \left[\int_{S_k}^{S_{k+1}} (X_t - \hat{X}_t)^2 dt \right] \\
&= \frac{1}{6} \mathbb{E} [(X_{R_k} - X_{S_{k-1}})^4] - \frac{1}{6} \mathbb{E} [(X_{R_k} - X_{S_k})^4] \\
&= \frac{1}{6} \mathbb{E} [(X_{R_k} - X_{S_k}) + (X_{S_k} - X_{S_{k-1}})]^4 - \frac{1}{6} \mathbb{E} [(X_{R_k} - X_{S_k})^4] \\
&= \frac{1}{6} \mathbb{E} [(X_{R_k} - X_{S_k})^4] + \frac{2}{3} \mathbb{E} [(X_{R_k} - X_{S_k})^3 (X_{S_k} - X_{S_{k-1}})] + \mathbb{E} [(X_{R_k} - X_{S_k})^2 (X_{S_k} - X_{S_{k-1}})^2] \\
&\quad + \frac{2}{3} \mathbb{E} [(X_{R_k} - X_{S_k}) (X_{S_k} - X_{S_{k-1}})^3] + \frac{1}{6} \mathbb{E} [(X_{S_k} - X_{S_{k-1}})^4] - \frac{1}{6} \mathbb{E} [(X_{R_k} - X_{S_k})^4] \\
&\stackrel{(a)}{=} \mathbb{E} [(X_{R_k} - X_{S_k})^2 (X_{S_k} - X_{S_{k-1}})^2] + \frac{1}{6} \mathbb{E} [(X_{R_k} - X_{S_k})^4] \\
&\stackrel{(b)}{=} \mathbb{E} [(X_{R_k} - X_{S_k})^2] \cdot \mathbb{E} [(X_{S_k} - X_{S_{k-1}})^2] + \frac{1}{6} \mathbb{E} [(X_{S_k} - X_{S_{k-1}})^4] \\
&\stackrel{(c)}{=} \mathbb{E}[R_k - S_k] \cdot \mathbb{E}[S_{k+1} - S_k] + \frac{1}{6} \mathbb{E} [(X_{S_k} - X_{S_{k-1}})^4] \\
&= \mathbb{E}[L_k] \cdot \bar{D} + \frac{1}{6} \mathbb{E} [(X_{S_k} - X_{S_{k-1}})^4],
\end{aligned} \tag{124}$$

where equality (a) is obtained because $(X_{R_k} - X_{S_k})$ is independent of $(X_{S_k} - X_{S_{k-1}})$ and $\mathbb{E}[X_{R_k} - X_{S_k}] = \mathbb{E}[X_{S_k} - X_{S_{k-1}}] = 0$ due to Wiener process evolution; equality (b) is obtained because $(X_{R_k} - X_{S_k})$ and $X_{S_k} - X_{S_{k-1}}$ are independent; and equality (c) is because the Wald's Lemma.

Therefore, for any stationary policy π that makes sampling decision only on \mathcal{I}_k , with probability 1, the objective function in the MSE minimization problem 1 can be rewritten as:

$$\begin{aligned}
&\limsup_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \int_{t=0}^T (X_t - \hat{X}_t)^2 dt \right] \\
&= \limsup_{K \rightarrow \infty} \frac{\sum_{k=1}^K \left(\frac{1}{6} \mathbb{E} [(X_{S_k} - X_{S_{k-1}})^4] + \mathbb{E}[L_k] \bar{D} \right)}{\sum_{k=1}^K \mathbb{E}[L_k]} \\
&= \limsup_{K \rightarrow \infty} \frac{\sum_{k=1}^K \mathbb{E} \left[\frac{1}{6} (X_{S_k} - X_{S_{k-1}})^4 \right]}{\sum_{k=1}^K \mathbb{E}[L_k]} + \bar{D}.
\end{aligned} \tag{125}$$

□

F PROOF OF LEMMA 3

PROOF. Consider a constant wait policy π_{const} that chooses $W_k \equiv \frac{1}{f_{\text{max}}}$ regardless of the transmission delay and estimation error $\mathcal{I}_k = (D_k, (X_{S_k} + t - X_{S_k}))$. Let Z_t be a Wiener process starting from time 0. Then according to equation (4a) from Corollary 1, the time-average MSE by using policy π_{const} can be computed by:

$$\bar{\mathcal{E}}_{\pi_{\text{const}}} = \limsup_{K \rightarrow \infty} \frac{\sum_{k=1}^K \mathbb{E} \left[\frac{1}{6} (X_{S_k} - X_{S_{k-1}})^4 \right]}{\sum_{k=1}^K \mathbb{E}[L_k]} + \bar{D}$$

$$\begin{aligned}
& \stackrel{(a)}{=} \limsup_{K \rightarrow \infty} \frac{\frac{1}{6} \sum_{k=1}^K \mathbb{E} \left[Z_{D_k + \frac{1}{f_{\max}}}^4 \right]}{\frac{1}{K} \sum_{k=1}^K \mathbb{E}[L_k]} + \bar{D} \\
& \stackrel{(b)}{=} \limsup_{K \rightarrow \infty} \frac{\frac{1}{2} \sum_{k=1}^K \mathbb{E} \left[\left(D_k + \frac{1}{f_{\max}} \right)^2 \right]}{\frac{1}{K} \sum_{k=1}^K \mathbb{E}[L_k]} + \bar{D} \\
& \stackrel{(c)}{=} \frac{1}{2} \frac{M + 2\bar{D} \frac{1}{f_{\max}} + \left(\frac{1}{f_{\max}} \right)^2}{\bar{D} + \frac{1}{f_{\max}}} + \bar{D},
\end{aligned} \tag{126}$$

where equality (a) is because by using policy π_{const} , given transmission delay D_k , the difference $X_{S_{k-1}+t} - X_{S_{k-1}}$ evolves like Wiener process Z_t starting from $t = 0$ and $S_{k+1} - S_k = D_k + \frac{1}{f_{\max}}$ due to the constant wait policy; equality (b) is obtained because given delay D_k , $Z_{D_k + \frac{1}{f_{\max}}} \sim \mathcal{N}(0, D_k + \frac{1}{f_{\max}})$ is a zero-mean Gaussian distribution with variance $D_k + \frac{1}{f_{\max}}$, and therefore the fourth order moment $\mathbb{E} \left[Z_{D_k + \frac{1}{f_{\max}}}^4 | D_k \right] = 3 \left(D_k + \frac{1}{f_{\max}} \right)^2$; equality (c) is obtained because D_k is i.i.d following distribution \mathbb{P}_D , by definition $M = \mathbb{E}_{\mathbb{P}_D} [D^2]$ and $\bar{D} = \mathbb{E}_{\mathbb{P}_D} [D]$. Since π_{const} may not be the MSE minimum sampling policy, we have $\bar{\mathcal{E}}_{\pi^*} \leq \bar{\mathcal{E}}_{\pi_{\text{const}}}$. Therefore, recall the definition $\gamma^* = \bar{\mathcal{E}}_{\pi^*} - \bar{D}$ from Subsection 3.2, we have:

$$\gamma^* \leq \bar{\mathcal{E}}_{\pi_{\text{const}}} - \bar{D} = \frac{1}{2} \frac{M + 2\bar{D} \frac{1}{f_{\max}} + \left(\frac{1}{f_{\max}} \right)^2}{\bar{D} + \frac{1}{f_{\max}}} =: \gamma_{\text{ub}}. \tag{127}$$

We then derive the lower bound of $\bar{\mathcal{E}}_{\pi^*}$. Recall that the optimum decision rule of policy π^* is given in equation (9), according to [24, Theorem 1], $\bar{\mathcal{E}}_{\pi^*}$ can be computed by

$$\bar{\mathcal{E}}_{\pi^*} = \frac{1}{6} \frac{\mathbb{E} \left[\max\{3(\gamma^* + \nu^*), Z_D^2\}^2 \right]}{\mathbb{E} \left[\max\{3(\gamma^* + \nu^*), Z_D^2\} \right]} + \bar{D} \stackrel{(d)}{\geq} \frac{1}{6} \mathbb{E}[\max\{3(\gamma^* + \nu^*), Z_D^2\}] + \bar{D} \geq \frac{1}{6} \mathbb{E}[Z_D^2] + \bar{D} = \frac{7}{6} \bar{D}, \tag{128}$$

where inequality (d) is from Cauchy-Schwartz inequality.

Finally, γ^* can be lower bonded by:

$$\gamma^* = \bar{\mathcal{E}}_{\pi^*} - \bar{D} \geq \frac{1}{6} \bar{D}. \tag{129}$$

□

G PROOF OF LEMMA 4

For any time t , the value Z_t of the Wiener process $Z_t \sim \mathcal{N}(0, t)$, according to [8, Theorem 7.5.6], $\forall \theta \in \mathbb{R}$, sequence $M_t(\theta) := \exp \left(\theta Z_t - \frac{\theta^2}{2} t \right)$ is a martingale with initial value $M_0(\theta) = 1, \forall \theta$.

Let $T \geq 0$ fixed as a constant, then $\tau_Y \wedge T$ is a stopping time, where $a \wedge b = \min\{a, b\}$. Then according to the optional stopping theorem [8, Theorem 7.5.1], denote $\phi_T(\theta)$ to be the expected value of $M_{\tau_Y \wedge T}(\theta)$, we have:

$$\phi_T(\theta) := \mathbb{E} \left[M_{\tau_Y \wedge T}(\theta) \right] = \mathbb{E} [M_0(\theta)] = 1, \forall \theta. \tag{130}$$

Therefore, the n -th order derivative of function $\phi_T(\theta)$, denoted by $\phi_T^{(n)}(\theta)$ can be computed by:

$$\phi_T^{(n)}(\theta) = \frac{\partial^n \mathbb{E} \left[M_{\tau_Y \wedge T}(\theta) \right]}{\partial \theta^n} = 0. \tag{131}$$

For each sample path ω , the absolute value $|Z_{l_Y \wedge T}| \leq \sqrt{3}\gamma$. Therefore the derivative $\left| \frac{\partial^n M_{\tau_Y \wedge T}(\theta)}{\partial \theta^n} \right|$ is bounded and continuous. Then according to Leibniz rule we have

$$\mathbb{E} \left[\frac{\partial^n M_{\tau_Y \wedge T}(\theta)}{\partial \theta^n} \right] = \frac{\partial^n \mathbb{E} \left[M_{\tau_Y \wedge T}(\theta) \right]}{\partial \theta^n} = 0. \tag{132}$$

First according to [8, Theorem 7.5.1 & Theorem 7.5.5], $\forall \gamma < \infty$, the mean of stopping time τ_Y is bounded, i.e.,

$$\mathbb{E} [l_Y] = \mathbb{E} [Z_{l_Y}^2] = \mathbb{E} [\max\{3\gamma, Z_D^2\}] = 3\gamma + \bar{D} < \infty. \tag{133}$$

To obtain the second-order moment of τ_Y , we compute the 4-th order derivative of $M_{l_Y \wedge T}(\theta)$, i.e.,

$$\frac{\partial^4 M_{l_Y \wedge T}(\theta)}{\partial \theta^4} \Big|_{\theta=0} = Z_{l_Y \wedge T}^4 - 6(l_Y \wedge T) \cdot Z_{l_Y \wedge T}^2 + 3(l_Y \wedge T)^2. \quad (134)$$

Plugging (134) into (132), we have:

$$\mathbb{E}[(l_Y \wedge T)^2] = 2\mathbb{E}\left[\left(\frac{1}{\sqrt{2}}l_Y \wedge T\right) \cdot \sqrt{2}Z_{l_Y \wedge T}^2\right] - \frac{1}{3}\mathbb{E}\left[Z_{l_Y \wedge T}^4\right] \leq \frac{1}{2}\mathbb{E}[(l_Y \wedge T)^2] + \frac{5}{3}\mathbb{E}[Z_{l_Y \wedge T}^2] \quad (135)$$

Therefore, for each $T < \infty$, the second order moment of $(l_Y \wedge T)$ can be upper bounded by:

$$\mathbb{E}[(l_Y \wedge T)^2] \leq \frac{10}{3}\mathbb{E}[Z_{l_Y \wedge T}^2] = \frac{10}{3}\mathbb{E}[\max\{3Y, Z_D^2\}] \leq \frac{10}{3}(3Y + \bar{D}) < \infty. \quad (136)$$

Finally, we can upper bound the fourth order moment of τ_Y by computing the 8-th order derivative of $M_{\tau_Y \wedge T}(\theta)$ at $\theta = 0$, i.e.,

$$\frac{\partial^8 M_{l_Y \wedge T}(\theta)}{\partial \theta^8} = Z_{l_Y \wedge T}^8 - 28(l_Y \wedge T) \cdot Z_{\tau_Y \wedge T}^6 + 210(l_Y \wedge T)^2 \cdot Z_{l_Y \wedge T}^4 - 420(l_Y \wedge T)^3 \cdot Z_{l_Y \wedge T}^2 + 105(l_Y \wedge T)^4. \quad (137)$$

Plugging equation (137) into (132), we have:

$$\begin{aligned} \mathbb{E}[(l_Y \wedge T)^4] &= -\frac{1}{105}\mathbb{E}\left[Z_{l_Y \wedge T}^8 - 28(l_Y \wedge T) \cdot Z_{\tau_Y \wedge T}^6 + 210(l_Y \wedge T)^2 \cdot Z_{l_Y \wedge T}^4 - 420(l_Y \wedge T)^3 \cdot Z_{\tau_Y \wedge T}^2\right] \\ &\leq 4\mathbb{E}\left[(l_Y \wedge T)^3 \cdot Z_{l_Y \wedge T}^2\right] + \frac{3}{15}\mathbb{E}\left[(l_Y \wedge T) \cdot Z_{l_Y \wedge T}^6\right] \\ &\leq 4\left(\mathbb{E}[(l_Y \wedge T)^4]\right)^{3/4} \cdot \left(\mathbb{E}[Z_{l_Y \wedge T}^8]\right)^{1/4} + \frac{3}{15}\left(\mathbb{E}[(l_Y \wedge T)^4]\right)^{1/4} \cdot \left(\mathbb{E}[Z_{l_Y \wedge T}^6]\right)^{3/4} \\ &\leq 4\left(\mathbb{E}[(l_Y \wedge T)^4]\right)^{3/4} \cdot \left((3Y)^4 + 105B\right)^{1/4} + \frac{3}{15}\left(\mathbb{E}[(l_Y \wedge T)^4]\right)^{1/4} \cdot \left((3Y)^4 + 105B\right)^{3/4}. \end{aligned} \quad (138)$$

Inequality (138) implies

$$\mathbb{E}[(l_Y \wedge T)^4] \leq 4^3 \left((3Y)^4 + 105B\right). \quad (139)$$

Let $T \rightarrow \infty$ and then use the dominated convergence theorem on the LHS of inequality (138), we conclude that $\mathbb{E}[l_Y^4]$ is bounded.

$$\mathbb{E}[\tau_Y^4] = -\frac{1}{105}(1 - 28 + 350 - 1708)(3Y)^4 = \frac{277}{21}(3Y)^4. \quad (140)$$

H PROOF OF LEMMA 5

PROOF.

$$\frac{d}{dY}\mathbb{E}\left[\frac{1}{6}\left((3Y)^2 \cdot \mathbb{I}_{(Z_D^2 \leq 3Y)} + Z_D^4 \cdot \mathbb{I}_{(Z_D^2 > 3Y)}\right)\right] = 3Y\Pr(Z_D^2 \leq 3Y). \quad (141)$$

$$\frac{d}{dY}\mathbb{E}\left[\left(3Y \cdot \mathbb{I}_{(Z_D^2 \leq 3Y)} + Z_D^2 \cdot \mathbb{I}_{(Z_D^2 > 3Y)}\right)\right] = 3\Pr(Z_D^2 \leq 3Y) \quad (142)$$

Therefore, the monotonic decreasing characteristics can be verified through

$$\bar{g}'_0(Y) = -\mathbb{E}[\max\{3Y, Z_D^2\}] < 0 \quad (143)$$

$$\bar{g}''_0(Y) = -3\Pr(Z_D^2 \leq 3Y) < 0. \quad (144)$$

Through Taylor expansion, since $\bar{g}_0(Y)$ is monotonically decreasing, for $Y < Y^*$ we have:

$$\bar{g}_0(Y) \geq \bar{g}_0(Y^*) + \bar{g}'_0(Y^*)(Y - Y^*) = -l(Y^*)(Y - Y^*) \geq 0. \quad (145)$$

Since $Y - Y^* < 0$, inequality (145) implies

$$(Y - Y^*)\bar{g}_0(Y) \leq -l(Y^*)(Y - Y^*)^2. \quad (146)$$

And for $Y > Y^*$, we have:

$$\begin{aligned} &\bar{g}_0(Y) \\ &\stackrel{(a)}{\leq} \mathbb{E}\left[\frac{1}{6}\max\{3Y^*, Z_D^2\}^2 - Y\max\{3Y^*, Z_D^2\}\right] \\ &\stackrel{(b)}{=} (Y^* - Y)l(Y^*). \end{aligned} \quad (147)$$

where inequality (a) is obtained because choosing the stopping time to be $\tau = \inf\{t \geq D \mid |Z_t| \geq \sqrt{3Y}\}$ minimizes function $\mathbb{E}\left[\frac{1}{6}Z_\tau^4 - Y\tau\right]$ and equality (b) is because $\mathbb{E}\left[\frac{1}{6}\max\{3Y^*, Z_D^2\}^2 - Y^*\max\{3Y^*, Z_D^2\}\right] = 0$.

Multiplying $(\gamma - \gamma^\star)$ on both sides of (147) we have:

$$(\gamma - \gamma^\star)\bar{g}_0(\gamma) \leq -l(\gamma^\star)(\gamma^\star - \gamma)^2, \forall \gamma > \gamma^\star. \quad (148)$$

Combining (146) and (148) finishes the proof of Lemma 5-(iii). \square

I PROOF OF COROLLARY 2

PROOF. The first conclusion follows directly from Lemma 1.

The conditional expectation of $(X_{S_{k+1}} - X_{S_k})^4$ can be upper bounded by:

$$\mathbb{E}_k[(X_{S_{k+1}} - X_{S_k})^8] = \mathbb{E}_k[\max\{3\gamma_k, Z_D^2\}^4] \leq \left((3\gamma_k)^4 + 105\mathbb{E}[D^4]\right) = (3\gamma_k)^4 + 105B. \quad (149)$$

Finally, we can upper bounded the second order moment of E_k as follows:

$$\begin{aligned} & \mathbb{E}_k[E_k^2] \\ &= \mathbb{E}_k\left[\left(\int_{S_k}^{S_k+D_k} (X_t - X_{S_{k-1}})^2 dt + \int_{S_k+D_k}^{S_{k+1}} (X_t - X_{S_k})^2 dt\right)^2\right] \\ &= \mathbb{E}_k\left[\left(\int_{S_k}^{S_k+D_k} (X_t - X_{S_k} + X_{S_k} - X_{S_{k-1}})^2 dt + \int_{S_k+D_k}^{S_{k+1}} (X_t - X_{S_k})^2 dt\right)^2\right] \\ &= \mathbb{E}_k\left[\left((X_{S_k} - X_{S_{k-1}})^2 D_k + 2(X_{S_k} - X_{S_{k-1}}) \cdot \int_{S_k}^{S_k+D_k} (X_t - X_{S_k}) dt + \int_{S_k}^{S_k+D_k} (X_t - X_{S_k})^2 dt\right)^2\right] \\ &\stackrel{(c)}{\leq} 3\mathbb{E}_k[(X_{S_k} - X_{S_{k-1}})^4 D_k^2] + 12\mathbb{E}_k\left[(X_{S_k} - X_{S_{k-1}})^2 \left(\int_{S_k}^{S_k+D_k} (X_t - X_{S_k}) dt\right)^2\right] + 3\mathbb{E}_k\left[\left(\int_{S_k}^{S_k+D_k} (X_t - X_{S_k})^2 dt\right)^2\right], \end{aligned} \quad (150)$$

where inequality (c) is from Cauchy-Schwartz $\mathbb{E}[(a+b+c)^2] \leq 3\mathbb{E}[a^2 + b^2 + c^2]$.

Since the transmission delay D_k is independent of $X_{S_k} - X_{S_{k-1}}$, the first term on the RHS of inequality (150) can be upper bounded by:

$$\mathbb{E}_k[(X_{S_k} - X_{S_{k-1}})^4 D_k^2] = (X_{S_k} - X_{S_{k-1}})^4 \mathbb{E}[D^2] \leq (X_{S_k} - X_{S_{k-1}})^4 \sqrt{B}. \quad (151)$$

To upper bound the second and third term on the RHS of inequality (150), we introduce the following Lemma, whose proof is provided in Appendix J

LEMMA 8. Recall that Z_t is a wiener process staring from time 0 and let $l_\gamma := \inf\{t \geq D \mid |Z_t| \geq \sqrt{3\gamma}\}$ be the frame length when threshold γ is used. When $\mathbb{E}[D^4] \leq B$, we have the following results:

$$\mathbb{E}\left[\left(\int_{t=0}^{l_\gamma} Z_t dt\right)^2\right] \leq \left(\frac{277}{31}(3\gamma)^4 + B\right) \cdot \left((3\gamma)^2 + \left(\frac{4}{3}\right)^4 \cdot 3\sqrt{B}\right) =: C_1(\gamma, B), \quad (152)$$

$$\mathbb{E}\left[\left(\int_{t=0}^{l_\gamma} Z_t^2 dt\right)^2\right] \leq \left(\frac{277}{31}(3\gamma)^4 + B\right) \cdot \left((3\gamma)^4 + \left(\frac{8}{7}\right)^8 \cdot 105B\right) =: C_2(\gamma, B). \quad (153)$$

Plugging inequality (152) and (153) into (150), we can upper bound $\mathbb{E}[E_k^2 | \gamma_k, X_{S_k} - X_{S_{k-1}}]$ by:

$$\mathbb{E}_k[E_k^2] = 3(X_{S_k} - X_{S_{k-1}})^4 \sqrt{B} + 12C_1(\gamma_k, B)(X_{S_k} - X_{S_{k-1}})^2 + 3C_2(\gamma, B). \quad (154)$$

\square

J PROOF OF LEMMA 8

Recall that $l_\gamma = \inf\{t \geq D \mid |\sqrt{Z_t}| \geq \sqrt{3\gamma}\}$ is a stopping time of the Wiener process Z_t starting from time 0. Then for any time $\tau \geq 0$ we have:

$$\begin{aligned} & \mathbb{E}\left[\left(\int_{t=0}^{l_\gamma} Z_t^p dt\right)^2\right] \\ &\leq \mathbb{E}\left[\left(\int_0^{l_\gamma} \left(\sup_{0 \leq t' \leq l_\gamma} |Z_{t'}|\right)^p dt\right)^2\right] \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E} \left[l_Y^2 \cdot \left(\sup_{0 \leq t' \leq l_Y} |Z_{t'}|^{2p} \right) \right] \\
&\stackrel{(a)}{\leq} \sqrt{\mathbb{E} [l_Y^4] \cdot \mathbb{E} \left[\sup_{0 \leq t' \leq l_Y} |Z_{t'}|^{4p} \right]}, \tag{155}
\end{aligned}$$

where inequality (a) from the Cauchy-Schwartz inequality.

From Lemma 4, $\mathbb{E}[l_Y^4]$ can be bounded as follows:

$$\mathbb{E}[l_Y^4] \leq 256 \left((3\gamma)^4 + 105B \right). \tag{156}$$

Next, we prove $\mathbb{E}[\sup_{0 \leq t' \leq \tau} |Z_{t'}|^{2p}]$ is bounded. Recall that the stopping rule is obtained by:

$$l_Y = \inf \{ t \geq D \mid |Z_t| \geq \sqrt{3\gamma} \}. \tag{157}$$

Then $\mathbb{E} \left[\sup_{0 \leq t' \leq l_Y} |Z_{t'}|^{4p} \right]$ can be upper bounded by:

$$\begin{aligned}
&\mathbb{E} \left[\sup_{0 \leq t' \leq l_Y} |Z_{t'}|^{4p} \right] \\
&= \mathbb{E} \left[\left(\sup_{0 \leq t' \leq l_Y} |Z_{t'}|^{4p} \right) \cdot \mathbb{I}(l_Y > D) \right] + \mathbb{E} \left[\left(\sup_{0 \leq t' \leq l_Y} |Z_{t'}|^{4p} \right) \cdot \mathbb{I}(l_Y \leq D) \right] \\
&\stackrel{(b)}{\leq} (3\gamma)^{2p} + \mathbb{E} \left[\sup_{0 \leq t' \leq D} |Z_{t'}|^{4p} \right], \tag{158}
\end{aligned}$$

where inequality (b) is because if $l_Y > D$, then we have $|Z_t| \leq \sqrt{3\gamma}, \forall t \in [D, l_Y]$ and therefore $\sup_{0 \leq t' \leq l_Y} |Z_{t'}|^{4p} \leq (3\gamma)^{2p} + \sup_{0 \leq t' \leq D} |Z_{t'}|^{4p}$. For each $D < \infty$, since the Wiener process Z_t is a martingale and D is a stopping time, for each $d < \infty$, we can upper bound $\mathbb{E} \left[\sup_{0 \leq t' \leq D} |Z_{t'}|^{4p} \right]$ as follows:

$$\mathbb{E} \left[\sup_{0 \leq t' \leq d} |Z_{t'}|^{4p} \right] \stackrel{(c)}{\leq} \left(\frac{4p}{4p-1} \right)^{4p} \mathbb{E} [Z_d^{4p}] \stackrel{(d)}{=} \left(\frac{4p}{4p-1} \right)^{4p} \frac{(4p)!}{2^{2p}(2p)!} d^{2p}. \tag{159}$$

where inequality (c) is because of the Doob's maximal inequality [20, p.54, Theorem 1.7] and equality (d) is because Z_d follows a Gaussian distribution.

When the transmission delay D is fourth order bounded, for $p = 1$ and 2, plugging $\mathbb{E}[D^4] \leq B$ and $\mathbb{E}[D^2] \leq \sqrt{B}$ into inequality (159) and (158), we have:

$$\mathbb{E} \left[\sup_{0 \leq t' \leq l_Y} |Z_{t'}|^4 \right] \leq (3\gamma)^2 + \left(\frac{4}{3} \right)^4 \cdot 3\mathbb{E}[D^2] = (3\gamma)^2 + \left(\frac{4}{3} \right)^4 \cdot 3\sqrt{B}, \tag{160}$$

$$\mathbb{E} \left[\sup_{0 \leq t' \leq l_Y} |Z_{t'}|^8 \right] \leq (3\gamma)^4 + \left(\frac{8}{7} \right)^8 \cdot 105\mathbb{E}[D^4] = (3\gamma)^4 + \left(\frac{8}{7} \right)^8 \cdot 105B. \tag{161}$$

Plugging inequality (161), (160) and (156) into inequality (155), we have:

$$\mathbb{E} \left[\left(\int_{t=0}^{l_Y} Z_t dt \right)^2 \right] \leq 256 \left((3\gamma)^4 + 105B \right) \cdot \left((3\gamma)^2 + \left(\frac{4}{3} \right)^4 \cdot 3\sqrt{B} \right), \tag{162}$$

$$\mathbb{E} \left[\left(\int_{t=0}^{l_Y} Z_t^2 dt \right)^2 \right] \leq 256 \left((3\gamma)^4 + 105B \right) \cdot \left((3\gamma)^4 + \left(\frac{8}{7} \right)^8 \cdot 105B \right). \tag{163}$$

K PROOF OF INEQUALITY (61)

PROOF. For $l \geq \bar{D}$, let $\Pi_l \triangleq \{\pi \mid \mathbb{E}[Z_l^2] = l, \forall \pi \in \Pi\}$ whose squared error at the time of sample is l . Next, we establish the lower bound of $\mathbb{E} \left[\int_0^\tau Z_t^2 dt \right]$ for any policy $\pi \in \Pi_l$, which can be formulated into the following optimization problem:

$$\inf_{\pi} \mathbb{E} \left[\int_0^\tau Z_t^2 dt \right], \text{ s.t. } \mathbb{E}[\tau] = l, \tau \geq D. \tag{164}$$

As is shown in [24, Theorem 7], the optimum solution to (164) has a threshold structure, and the optimum sampling policy for each sample path is as follows:

$$\tau = \inf\{t \geq D \mid |Z_t| \geq \lambda^*\}, \quad (165)$$

where the selection of λ^* satisfies:

$$\mathbb{E}[\tau] = \mathbb{E}[Z_\tau^2] = l. \quad (166)$$

Through Lemma 7, we can compute the optimum solution to (164) as follows:

$$\mathbb{E}\left[\int_0^\tau Z_t^2 dt\right] = \frac{1}{6}\mathbb{E}[Z_\tau^4] = \frac{1}{6}\mathbb{E}\left[\max\{Z_D^2, 3\gamma\}^2\right]. \quad (167)$$

To finish the proof of inequality (61), it then remains to lower bound (167) as follows:

$$\frac{1}{6}\mathbb{E}\left[\max\{3\gamma, Z_D^2\}^2\right] - \gamma\mathbb{E}\left[\max\{3\gamma, Z_D^2\}\right]. \quad (168)$$

The analysis is divided into the following two cases. For simplicity, denote γ_l to be the threshold such that $\mathbb{E}[\max\{Z_D^2, 3\gamma_l\}] = l$.

- Case 1: $l \geq l(\gamma^*)$, it can be easily verify that $\gamma_l \geq \gamma^*$. Therefore, we have:

$$\begin{aligned} & \frac{1}{6}\mathbb{E}\left[\max\{Z_D^2, 3\gamma_l\}^2\right] \\ &= \frac{1}{6}(3\gamma_l)^2\Pr(Z_D^2 \leq 3\gamma_l) + \frac{1}{6}\mathbb{E}\left[Z_D^4 \cdot \mathbb{I}(Z_D^2 > 3\gamma_l)\right] \\ &= \frac{1}{6}\left(\mathbb{E}\left[(3\gamma^*)^2\mathbb{I}(Z_D^2 \leq 3\gamma^*)\right] + \mathbb{E}\left[Z_D^4\mathbb{I}(Z_D^2 > 3\gamma^*)\right] + \mathbb{E}[(3\gamma_l)^2 - (3\gamma^*)^2]\mathbb{I}(Z_D^2 \leq 3\gamma^*)\right. \\ &\quad \left.+ \mathbb{E}[(3\gamma_l)^2 - Z_D^4]\mathbb{I}(3\gamma^* \leq Z_D^2 \leq 3\gamma_l)\right) \\ &\stackrel{(a)}{\geq} q(\gamma^*) + \frac{1}{6}\mathbb{E}\left[(3\gamma_l - 3\gamma^*)^2\mathbb{I}(Z_D^2 \leq 3\gamma^*)\right] + \frac{1}{6}\mathbb{E}[3\gamma^*(3\gamma_l - 3\gamma^*)\mathbb{I}(Z_D^2 \leq 3\gamma^*)] + \frac{1}{6}\mathbb{E}\left[6\gamma^*(3\gamma_l - Z_D^2)\mathbb{I}(3\gamma^* \leq Z_D^2 \leq 3\gamma_l)\right] \\ &\stackrel{(b)}{\geq} \gamma^*l(\gamma^*) + \frac{1}{6}p_w(l(\gamma_l) - l(\gamma^*))^2 + \gamma^*(l(\gamma_l) - l^*) \\ &= \gamma^*l + \frac{1}{6}p_w(l - l(\gamma^*))^2, \end{aligned} \quad (169)$$

where inequality (a) is obtained because $(3\gamma_l)^2 - (3\gamma^*)^2 = (3\gamma_l - 3\gamma^*)^2 + 2 \times 3\gamma^*(3\gamma_l - 3\gamma^*)$ and for $Z_D^2 = x$ that satisfies $3\gamma^* \leq x \leq 3\gamma_l$, $(3\gamma_l)^2 - x^2 \geq 6\gamma^*(3\gamma_l - x)$; inequality (b) is because $l(\gamma_l) - l(\gamma^*) = \mathbb{E}\left[(3\gamma_l - 3\gamma^*)\mathbb{I}(Z_D^2 \leq 3\gamma^*)\right] + \mathbb{E}\left[(3\gamma_l - Z_D^2)\mathbb{I}(3\gamma^* \leq Z_D^2 \leq 3\gamma_l)\right]$.

- Case 2: $l \leq l(\gamma^*)$, similarly, it can be verified that $3\gamma_l \leq 3\gamma^*$. As a result:

$$\begin{aligned} & \frac{1}{6}\mathbb{E}\left[\max\{Z_D^2, 3\gamma_l\}^2\right] \\ &= \frac{1}{6}\mathbb{E}\left[(3\gamma_l)^2\mathbb{I}(Z_D^2 \leq 3\gamma_l)\right] + \frac{1}{6}\mathbb{E}\left[Z_D^4\mathbb{I}(Z_D^2 > 3\gamma_l)\right] \\ &= \frac{1}{6}\left(\mathbb{E}\left[(3\gamma^*)^2\mathbb{I}(Z_D^2 \leq 3\gamma^*)\right] + \mathbb{E}\left[Z_D^4\mathbb{I}(Z_D^2 > 3\gamma^*)\right] - \mathbb{E}\left[(3\gamma^*)^2 - (3\gamma_l)^2\right]\mathbb{I}(Z_D^2 \leq 3\gamma^*) - \mathbb{E}\left[(Z_D^2 - (3\gamma_l)^2)\mathbb{I}(3\gamma_l \leq Z_D^2 \leq 3\gamma^*)\right]\right) \\ &\stackrel{(c)}{\geq} q(\gamma^*) + \frac{1}{6}\mathbb{E}\left[(3\gamma_l - 3\gamma^*)^2\mathbb{I}(Z_D^2 \leq 3\gamma^*)\right] + \frac{1}{6}\mathbb{E}\left[6\gamma^*(3\gamma_l - 3\gamma^*)\mathbb{I}(Z_D^2 \leq 3\gamma^*)\right] - \frac{1}{6}\mathbb{E}\left[3\gamma^*(3\gamma^* - Z_D^2)\mathbb{I}(3\gamma_l \leq Z_D^2 \leq 3\gamma^*)\right] \\ &= \gamma^*l(\gamma^*) + \gamma^*\mathbb{E}\left[(3\gamma_l - 3\gamma^*)\mathbb{I}(Z_D^2 \leq 3\gamma^*)\right] + \gamma^*\mathbb{E}\left[(3\gamma^* - Z_D^2)\mathbb{I}(3\gamma_l \leq Z_D^2 \leq 3\gamma^*)\right] + \frac{1}{6}p_w(l - l(\gamma^*))^2 \\ &= \gamma^*l + \frac{1}{6}p_w(l - l(\gamma^*))^2, \end{aligned} \quad (170)$$

where inequality (c) is obtained similarly as inequality (a) and (b).

□