

Regret-Optimal and Stability-Enhanced Online Sampling of the Wiener Process for Remote Estimation over an Unreliable Channel with Unknown Statistics

Miao Pan, Haoyue Tang, Jiayu Pan, Tie Qiu, and Jianwei Yin

Abstract—We study the online sampling problem of a Wiener process, where the samples are transmitted to the remote estimator via an unreliable channel. The objective is to minimize the long-term average mean squared error (MSE) of the remote estimator subject to a sampling frequency constraint. To derive the online sampling strategy when the channel statistics is unknown, we reformulate the problem as an optimal stopping problem and review the sufficient condition of the offline MSE-optimal sampling threshold. Utilizing the sufficient condition, we then propose a stochastic approximation algorithm that adaptively approximate the optimum threshold. Our theoretical analysis shows the approximated threshold converges to the optimum sampling threshold almost surely. Moreover, let k be the number of successfully received samples, we proved that: the MSE gap between our online algorithm and the optimum offline algorithm decays with rate $O(1/k)$, while for any online sampling algorithm that does not know the channel statistics in advance, the minimum MSE gap under the worst case delay distribution decays with rate more than $\Omega(1/k)$. To improve the speed of the adaptive learning algorithm, we integrate the Polyak–Ruppert (PR) averaging algorithm when updating the sampling threshold, achieving a faster convergence speed both theoretically and in simulations.

Index Terms—Remote Estimation, Unreliable Channel, Wiener Process, Polyak–Ruppert Averaging

I. INTRODUCTION

The ability to provide timely and accurate status updates is essential in modern networked systems such as remote health monitoring, GPS tracking, and drone control [1]–[3]. These applications rely on real-time estimates of rapidly evolving state information (e.g., location, speed), which are reconstructed from samples transmitted over unreliable or resource-

constrained communication channels. In such settings, the freshness of information—i.e., the difference between the current state and the state estimated from the received data—is a critical factor in maintaining system performance. To maintain accurate state estimates, systems often rely on sampling strategies, which determine when to observe and transmit a new sample based on past observations and system constraints. Specifically, when the source is modeled as a stochastic process, a more fundamental objective is to minimize the mean squared estimation error (MSE) at the receiver, which directly captures the quality of signal reconstruction.

Recent research has studied the optimal MSE-oriented sampling strategies under various idealized conditions [4]–[8]. For instance, when the signal varies according to a Wiener process, a kind of threshold sampling strategy has been shown to be optimal [9]–[11]. These strategies postpone transmission until the estimation error surpasses a predefined threshold, thereby striking a balance between estimation accuracy and communication cost. Such policies have been further characterized using the mathematical framework of free-boundary theory, which offers principled tools for analyzing optimal stopping and control problems [12]. Nonetheless, real-world communication channels often exhibit random delays and unreliability. Recent research have extended the analysis to non-ideal settings such as packet-losses [13]–[15]. These studies confirm that threshold-based sampling remains MSE optimal, however, computing the optimum threshold requires that the channel statistics are known in advance.

When the channel statistics are not known in advance, adaptive algorithms have been developed to approximate the optimum sampling threshold. Several approaches have been proposed based on stochastic control [16]–[19], bandit learning [20]–[22], and reinforcement learning [23]–[25]. For systems with unknown and random delays, stochastic approximation has also been used to design online strategies [26], [27].

We aim to design an online sampling strategy that has a lower regret value relative to offline benchmarks and is robust to estimator variance under unreliable transmission conditions. However, three core challenges remain unresolved when performing MSE-optimal sampling over unknown and unreliable communication channels:

C1: Implicit thresholds and convergence analysis chal-

This work is supported by National Natural Science Foundation of China 62402435, by Ningbo Yongjiang Talent Programme 2023A-398-G, and Natural Science Foundation of Ningbo 2024J205. (Corresponding author: Jiayu Pan.)

M. Pan is with the School of Software Technology, Zhejiang University, Hangzhou, China (e-mail: 22351015@zju.edu.cn).

H. Tang is with Meta Platforms, Inc. She will join the Department of Electrical Engineering, University of Hawaii at Manoa, Honolulu, USA (e-mail: haoyue@hawaii.edu).

J. Pan and T. Qiu are with the School of Computer Science and Engineering, Northeastern University, Shenyang, China (e-mail: panjiayu@cse.neu.edu.cn; qitue@ieee.org).

J. Yin is with the College of Computer Science and Technology, Zhejiang University, Hangzhou, China (e-mail: zjuyjw@cs.zju.edu.cn).

lenges. In unreliable channel scenarios, the optimal threshold v lacks a closed-form solution and must be computed iteratively. For example, in [15], v_n is updated repeatedly until it converges to v . Unlike prior works with explicit thresholds, our setting complicates the analysis of convergence and optimality, requiring careful treatment of approximation errors and value function recursion.

C2: Instability of Robbins–Monro (RM) algorithm. RM is a classical method for solving root-finding problems via noisy observations. However, its stability is highly sensitive to the variance of the gradient estimator, which can be large under unknown and unreliable delays. Existing works rarely address this issue, and no prior method provides variance control or step size guarantees in this context.

To address these challenges, we propose an online stochastic approximation framework for learning the optimal strategy: First, we reformulate the MMSE-optimal sampling problem as a stochastic root-finding problem and apply the RM algorithm to learn minimum MMSE from observed samples. **For C1**, we adopt the iterative procedure from [15] to compute the optimal threshold in settings where no closed-form solution exists due to channel unreliability. Building on this offline framework, we integrate the RM method to develop an online algorithm that dynamically estimates the optimal threshold from observed data and updates the sampling strategy based on estimated MSE. We theoretically prove almost sure convergence to the optimal parameter and establish a sublinear convergence rate. Furthermore, we derive a minimax lower bound via Le Cam’s method [28], which characterizes the fundamental performance limit of any online algorithm. These results provide strong theoretical justification for using regret—the MMSE gap between our online strategy and the offline optimum—as a principled performance metric. **For C2**, we introduce a Polyak–Ruppert (PR) averaging technique to stabilize the RM updates and prove that it achieves the same asymptotic accuracy with improved empirical stability. We also derive the largest admissible step size that guarantees optimization objective convergence under asymptotic convergence variance conditions. Our main contributions are:

- We develop a stochastic approximation-based online learning algorithm that jointly addresses sampling frequency constraints, channel unreliability, and delay uncertainty. Our method achieves sublinear regret and adapts to unknown delay distributions in real time.
- To enhance convergence speed and numerical stability, we incorporate PR averaging into the RM updates. This improves empirical performance while maintaining asymptotic optimality guarantees.
- We establish rigorous theoretical guarantees, including almost sure convergence, sublinear convergence rates, and minimax lower bounds. Additionally, we derive the maximal admissible step size that ensures stable learning under variance conditions.
- Extensive simulations under various delay distributions validate the robustness and effectiveness of our approach. Results show consistent convergence toward the offline-

optimal threshold, even in challenging delay settings.

II. RELATED WORK

Online sampling algorithm. When the transmission delay is known, the problem of minimizing AoI and the problem of minimizing MSE have been widely studied. For example, [7] investigated the nonlinear AoI minimization problem under limited transmission frequency and proposed a binary search algorithm to efficiently compute the optimal threshold. Similarly, [29] focused on MSE minimization under frequency constraints and used free-boundary theory to derive and prove the optimality of the corresponding threshold-based strategy. However, in the actual remote sampling estimation problem, the transmission delay may not be known, and it is difficult to estimate the transmission delay distribution. As a result, designing optimal sampling strategies under unknown transmission delays remains a critical open problem. We propose an adaptive transmission strategy. This strategy estimates the threshold of the optimal strategy based on the historical information of previous sampling, and as the number of transmissions increases, the adaptive strategy threshold can converge to the threshold of the optimal strategy under the unknown delay distribution.

Unstable channel transmission. In real communication systems, the transmission channel is not necessarily stable, and transmission failures often occur. [15] studied the MSE minimization problem when the transmission channel is unstable. [15] not only gave an efficient algorithm for solving the optimal strategy, but also proved the optimality of the strategy through the free boundary theory. While the structural form of the optimal strategy remains consistent with the stable-channel case, the derivation and proof techniques differ significantly due to channel unreliability. When the channel is unstable, the strategy threshold of failed transmission needs to be used to approximate the optimal strategy threshold. The MSE minimization problem is greatly increased in unstable channels. [27] studied the MSE minimization problem of the Wiener process sampling when the transmission delay is unknown. [27] constructed an adaptive optimal sampling strategy using the random approximation method, and proved the optimality of the adaptive strategy under the worst delay distribution using the Le Cam two-point method. However, there is little research on how to minimize the MSE problem in sampling when the transmission delay is unknown and unstable.

Difference with Prior Studies. Prior works [2], [18], [26], [27] focus on stable channels, whereas we establish, for the first time, convergence guarantees for online sampling under unreliable transmissions in the Wiener process. The key challenge lies in handling non-i.i.d. and random-length failures preceding each successful sample, which result in a sequence of heterogeneous, history-dependent value functions J . This setting requires a refined convergence analysis beyond classical techniques. Moreover, we apply Polyak–Ruppert averaging to improve the asymptotic behavior of the Robbins–Monro

iterates, ensuring asymptotic normality and $O(1/k)$ mean-square error decay. While the step size remains $\eta_k = O(1/k)$, our analysis shows that a properly tuned pre-factor achieves the optimal trade-off between convergence speed and asymptotic variance. Although our strategy adopts a threshold structure, it emerges from fundamental optimality, not methodological overlap.

III. MODEL AND PROBLEM FORMULATION

A. System Model

We consider a continuous-time state update system as shown in Fig. 1, in which a sensor obtains samples from the time-varying source and transmits the samples to a destination via a network channel.

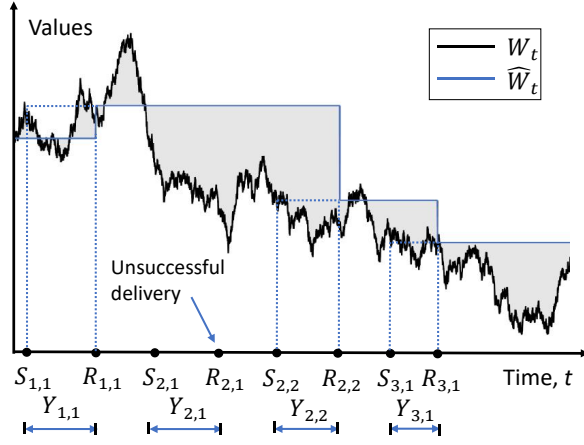


Fig. 1. Illustration of a sample path of the Wiener process W_t and its piecewise-constant MMSE estimate $\hat{W}_t$ at the receiver.

The source process $W_{t \geq 0}$ is modeled as a Wiener process. Let $k = 1, 2, 3, \dots$ be the index of the k -th successfully received packet, and let M_k denote the number of transmission attempts required in round k . The sampler generates the j -th sample in the k -th round at time $S_{k,j}$, and the transmission packet contains both the sampling time and the sampled value $(S_{k,j}, W_{S_{k,j}})$. The unreliable channel has an i.i.d. transmission failure, and we denote $\alpha \in [0, 1)$ as the probability of failure (i.e., the channel condition is OFF). The channel also has an i.i.d. transmission time $D_{k,j}$, and we have $\mathbb{E}[D^2] < \infty$. The transmission time and the channel condition are mutually independent. In this paper, we also assume that the transmission time is lower bounded, i.e., there exists $\epsilon > 0$ (which can be sufficiently small) such that $D_{k,j} \geq \epsilon$. The j -th sample is delivered to the destination at time $R_{k,j}$, where $R_{k,j} = S_{k,j} + D_{k,j}$. At the delivery time $R_{k,j}$, the destination knows the outcome of the transmission of the j -th sample. Only if the transmission was successful, the destination receives the sample message $(S_{k,j}, W_{S_{k,j}})$. After the delivery, the destination immediately sends an error-free acknowledgment indicating success or failure. We assume that the sampler proceeds to generate the next sample only after receiving this feedback. This feedback-driven design

avoids queueing and ensures that each sampled packet can be delivered to the destination as timely as possible to maintain information freshness.

B. Estimator

During the interval $t \in [R_{k-1, M_{k-1}}, R_{k, M_k})$, the freshest (and thus most informative) sample available at the receiver is still $(S_{k-1, M_{k-1}}, W_{S_{k-1, M_{k-1}}})$, where k is the number of successful transmission. Let j be the number of attempts in block k that have been received by time t . Then, we define $\mathcal{F}_t = \{S_{k,n}, D_{k,n}, W_{S_{k,n}}\}_{n=0}^{j-1}$ as the receiver's history up to time t , including all timestamps, transmission delays, and sampled values from past attempts, where $S_{k,0}$ corresponds to the last successfully received sample $S_{k-1, M_{k-1}}$. Using the strong Markov property of the Wiener process, the minimum mean-square error (MMSE) estimator $\hat{W}_t$ at time t is given by:

$$\hat{W}_t = \mathbb{E}[W_t | \mathcal{F}_t] = W_{S_{k-1, M_{k-1}}}, \quad t \in [R_{k-1, M_{k-1}}, R_{k, M_k}). \quad (1)$$

C. Problem Formulation

Our goal is to optimize the long-term average mean square error (MSE) of a communication system under channel transmissions. In addition, a constraint is imposed on the average sampling frequency. Therefore, the optimal sampling problem is formulated as follows:

$$\text{mse}_{\text{opt}} \triangleq \inf_{\pi \in \Pi} \limsup_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \int_{t=0}^T (\hat{W}_t - W_t)^2 dt \right], \quad (2)$$

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

Here, $c(T)$ denotes the total number of samples taken by time T . A sampling strategy π is defined as a sequence of sampling times $S_{i,j}$. The strategy set Π denotes the collection of all causal sampling policies π , where each $S_{i,j}$ is determined based on the information available up to the previous delivery time. $f_{\max}$ is the maximum allowed sampling rate. This constraint reflects practical limitations in real-world systems, particularly in energy-constrained settings like battery-powered IoT devices or wireless sensor networks—where excessive sampling may quickly deplete energy resources or overload the communication channel.

IV. PROBLEM RESOLUTION

Define the time interval between the sampling time S_k of the k -th sample and the sampling time S_{k+1} of the $(k+1)$ -th sample as frame k . Since for each frame k , the difference $W_{k+1} - W_k$ is the Wiener process, independent of the past $\{W_{k'+1} - W_{k'}\}_{k' < k}$, the problem can be reformulated as the following Markov decision process:

$$\text{mse}_{\text{opt}} \triangleq \inf_{\pi \in \Pi} \lim_{K \rightarrow \infty} \frac{\sum_{k=0}^K \mathbb{E} \left[\int_{R_{k,0}}^{R_{k+1,0}} (\hat{W}_t - W_t)^2 dt \right]}{\sum_{k=0}^K \mathbb{E}[R_{k+1,0} - R_{k,0}]}, \quad (4)$$

$$s.t. \quad \lim_{K \rightarrow \infty} \sup \mathbb{E}[R_{k+1,0} - R_{k,0}] \geq \frac{1}{f_{\max}}. \quad (5)$$

R_k denotes the time at which the k -th sample successfully reaches the destination. Since π^* achieves the minimum expected time average MSE among all sampling strategy Π , we have

$$\lim_{K \rightarrow \infty} \sup_{\pi \in \Pi} \frac{\sum_{k=0}^K \mathbb{E} \left[\int_{R_{k,0}}^{R_{k+1,0}} (\hat{W}_t - W_t)^2 dt \right]}{\sum_{k=0}^K \mathbb{E}[R_{k+1,0} - R_{k,0}]} \geq \beta, \quad (6)$$

where β^* is mse_{opt} , which is the average cost of the MDP when the optimum strategy π^* is used. Since $\sum_{k=0}^K \mathbb{E}[R_{k+1,0} - R_{k,0}] > 0$, for any strategy $\pi \in \Pi$, we can rewritten Eq. (6) as:

$$\inf_{\pi \in \Pi} \lim_{K \rightarrow \infty} \sum_{k=0}^K \mathbb{E} \left[\int_{R_{k,0}}^{R_{k+1,0}} (\hat{W}_t - W_t)^2 dt \right] - \beta \sum_{k=0}^K \mathbb{E}[R_{k+1,0} - R_{k,0}] \geq 0. \quad (7)$$

According to Lemma 1 in [15], when mse_{opt} equals to β , Eq. (7) is zero and equivalent to Eq. (4). When mse_{opt} is greater than β , Eq. (7) is greater than zero. Through this transformation, we can find optimal strategy of Eq. (7), which is equivalent to the optimal strategy of Eq. (4). Since each successful transmission can be considered as an update, by the nature of the update process, each successful transmission is independently and identically distributed from one another. In this way, we can only study how to minimize the AoI within a successful transmission.

We focus on the case of unstable channels, thus there are several failed transmissions between each successful transmission of a sampled sample. We denote the number of failed transmissions between each successful transmission with the subscript n . For example, W_{S_k} denotes the k -th successful transmission, and $W_{S_{k,n}}$ denotes the n -th failed transmission between the k -th and the $(k+1)$ -th successful transmission. The estimated signal $\hat{W}_t$ is $W_{S_{k,1}}$. Therefore the two successful transmissions are not independent of each other.

When there are multiple transmissions with non-independent and identically distributed sampling waiting time in the k -th successful transmission round, directly calculating the average estimation error over multiple intervals becomes difficult to handle. To address this issue, we adopt an integral form to evaluate the cumulative error between two successful transmissions. Let w denote the current estimation error at time $S_{k,0}$, i.e., w is actually the estimation error. Then W_t is the Wiener increment from the sampling point to the next delivery. When a single transmission is allowed, the corresponding value function $J_1(w, \beta)$ is defined as:

$$J_1(w, \beta) \triangleq g(w, \beta) = \inf_{\pi \in \Pi} \mathbb{E} \left[\int_{R_{k,0}}^{R_{k,1}} (w + W_t)^2 dt \right. \\ \left. - \beta(R_{k,1} - R_{k,0}) \middle| W_{S_{k,0}} - \hat{W}_{S_{k,0}} = w \right], \quad (8)$$

where $g(w, \beta)$ is the value function of a single transfer. And so on, we have $J_2(w, \beta) = g(w, \beta) + \alpha \mathbb{E}[g(w + W_{Y_1}, \beta)]$, where α is packet loss rate and Y is transmission delay. Therefore, we have:

$$J_n(w, \beta) = \sum_{i=1}^n \alpha^{i-1} \mathbb{E}[g(w + \sum_{j=1}^i W_{Y_j + \tau_j}, \beta)]. \quad (9)$$

Define M_k as the number of transmission attempts required to achieve a successful transmission in the k -th round. We can easily see that when $n \rightarrow M_k$, $J_n \rightarrow J$. The expression of $g(w, \beta)$ is [15]:

$$g(w, \beta) = \frac{1}{2} \mathbb{E}[Y^2] + \frac{1}{6} \max(v^4 - w^4, 0) [Y] \beta \\ + \mathbb{E}[Y] w^2 - (\beta - \mathbb{E}[Y]) \max(v^2 - w^2, 0) - \mathbb{E}[Y] \beta, \quad (10)$$

where v is threshold of sampling strategy. As shown in [15], the optimal sampling strategy for remote estimation of the Wiener process over an unreliable channel adopts a threshold-based strategy. Specifically, a new sample is sampled when the absolute deviation of the process exceeds a predefined threshold v .

V. ONLINE SAMPLING ALGORITHM

Before introducing our online sampling policies that minimize the MSE when the channel statistics is unknown, we first review the MSE-optimal sampling policy $\pi_{mse,unreliable}$ and its properties [15]. The optimum sampling policy has a threshold structure, which means that after a signal is sampled, the system waits until $|w + W_t|$ exceeds a threshold before taking the next sample:

Theorem 1: The sequence of optimal stopping times τ_n 's to the problem is given as follows:

$$\tau = \inf_t \{t \geq 0 : |w + W_t| \geq v(\beta)\}, \quad (11)$$

where $v(\beta) = \lim_{n \rightarrow \infty} v_n(\beta)$ and $v_n(\beta)$ is the unique positive root of the free boundary differential equation:

$$\frac{\partial}{\partial w} J_n(w, \beta) \Big|_{w=v_n(\beta)^+} = \frac{\partial}{\partial w} J_n(w, \beta) \Big|_{w=v_n(\beta)^-}. \quad (12)$$

When β approaches the optimum β^* , the threshold waiting strategy tends to the optimal strategy. When $|w| \leq v(\beta^-)$, according to Eq. (51) in [15], we have:

$$J'_n(w) = G'_n(w) = 2\mathbb{E}[Y]w + \alpha \mathbb{E}[J'_{n-1}(w + W_Y)]. \quad (13)$$

Notably, $g'(w, \beta)$ is $2\mathbb{E}[Y]w$, when $|w| \leq v(\beta^-)$. Therefore, the only difference between J and G is that during the first transmission, G adopts a zero-wait strategy. Similarly, when $|w| \geq v(\beta^-)$, we have:

$$J'_n(w) = 2\beta w - \frac{2}{3} w^3 + \alpha \mathbb{E}[J'_{n-1}(w + W_Y)]. \quad (14)$$

The above theorem is valid for any $\beta > 0$, so we can design a binary search algorithm to iteratively solve the optimal β^* . Because β^* makes $\mathbb{E}_w[J_n(w, \beta^*)]$ equal to 0, we can determine the iteration direction of β based on this condition. In other words, when $\mathbb{E}_w[J_n(w, \beta)]$ is greater than 0, we

increase the lower bound of β . Otherwise, we reduce the upper bound of β . When the upper and lower bounds are close enough, we have found β^* . This theorem gives the solution to the optimal strategy in the unconstrained case.

A. Online Sampling Algorithm

The derivation of the optimal strategy in the known-delay setting relies critically on the knowledge of the delay distribution. As a result, when the distribution of transmission delays is unknown, the offline strategy can no longer be directly applied. It is worth noting that the optimal sampling strategy will make $\mathbb{E}_w[J(w, \beta)]$ equal to 0 when the delay distribution is known, i.e.,

$$\mathbb{E}_w[J_n(w, \beta^*)] = 0. \quad (15)$$

The seminal work in the field of stochastic approximation is the Robbins-Monro algorithm. This is a classical solution to a wide class of stochastic iterative algorithms for solving root finding or solving optimization problems. We develop an online algorithm for estimating β^* . The algorithm is initialized by selecting $\beta_1 = 0$ and $U_1 = 0$. In each frame k , the sampling and updating rules are:

1. Sampling: We treat $\nu_k := \frac{1}{V} V_k^+$ as the dual optimizer, where $V > 0$ is fixed as a constant and V_k denote the cumulative violation of the sampling frequency constraint. The waiting time $\tau_{k,n}$ is selected to minimize the function J_n , and according to Theorem 1, $\tau_{k,n}$ is selected by:

$$\tau_{k,j} = \inf\{t \geq 0 : |w + W_t| \geq v_j(\beta_k - \nu_k)\}, j < n, \quad (16)$$

where v_j can be solved by Eq. (12). The subscript k denotes the k -th estimate. The subscript n denotes the stopping time approximation truncated at the n -th iteration. Since we cannot model the case of $n \rightarrow \infty$, we choose a larger n such that v_n and J_n are close to v and J , respectively.

2. Update β_k : To search for the root $\beta > 0$ of Eq. (15), we update through the Robbins-Monro algorithm. In each frame k , we are given an i.i.d sample $\delta W_j = W_{S_{k,j}+Y_{k,j}+\tau_{k,j}} - W_{S_{k,j}} \sim W_Y$, where $j < n$. Then we can compute β_{k+1} by:

$$\beta_{k+1} = (\beta_k + \eta_k X_k)^+, \quad (17)$$

where $X_k = h_n(\delta W_n, \beta) = \sum_{j=1}^n \alpha^{j-1} g(\delta W_j, \beta)$.

When the channel transmission is unstable, we need to compute the iteration n times to compute J_n . Define $\bar{Y}_{lb}$ to be a non-zero lower bound on the average delay, based on the RM algorithmic step size $\{\eta_k\}$ that can be chosen in the following way:

$$\eta_k = \frac{1}{\bar{Y}_{lb}(2 + k^\zeta)}, \zeta > 1. \quad (18)$$

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

$$V_{k+1} = V_k + \left(\frac{1}{f_{max}} - \frac{1}{n} \sum_{j=1}^n (Y_{k,j} + \tau_{k,j}) \right). \quad (19)$$

Algorithm 1 Online Estimation of β^* via Robbins-Monro

```

1: Input: Step size parameters  $\bar{Y}_{lb}, \zeta > 0$ ; truncation level
    $n$ ; constant  $V > 0$ 
2: Initialize:  $\beta_1 \leftarrow 0, V_1 \leftarrow 0$ 
3: for each frame  $k = 1, 2, \dots$  do
4:   Dual Optimizer:  $\nu_k \leftarrow V_k$ 
5:   for  $j = 1$  to  $n$  do
6:     // Sampling
7:     Update  $G'_j(w, \beta)$  in Eq. (13)
8:     Update  $J'_j(w, \beta)$  in Eq. (14)
9:     Solve  $v_j(\beta)$  in Eq. (12)
10:    Solve  $\tau_{k,j}^{(j)}$  using Eq. (16)
11:    Sample  $\delta W_j \leftarrow W_{S_{k,j}+Y_{k,j}+\tau_{k,j}} - W_{S_{k,j}}$ 
12:  end for
13:  // Estimate based on sampled trajectories
14:   $X_k \leftarrow \sum_{j=1}^n \alpha^{j-1} g(\delta W_j, \beta)$ 
15:   $\eta_k \leftarrow \frac{1}{\bar{Y}_{lb}(2+k^\zeta)}$ 
16:  // Update  $\beta_k$  with projected stochastic observation
17:   $\beta_{k+1} \leftarrow (\beta_k + \eta_k X_k)^+$ 
18:  // Update constraint violation severity
19:   $V_{k+1} = V_k + \left( \frac{1}{f_{max}} - \frac{1}{n} \sum_{j=1}^n (Y_{k,j} + \tau_{k,j}) \right)$ 
20: end for

```

Note on Finite-Sample Approximation. In Algorithm 1, we approximate the true objective J using a finite-sample surrogate J_n . While this introduces an approximation error, its impact is negligible in practice. This is because the threshold values $v_{k,j}$ are decreasing with increasing j , leading to smaller incremental values of the utility function $g(\cdot)$ in later samples. As a result, the tail terms in the summation contribute marginally to the final update, and the difference between J and J_n remains consistently small. Therefore, the use of finite-sample approximations in our estimator is both computationally efficient and theoretically justified.

B. Stability and Acceleration

Stability via Variable Window Averaging. The classic RM algorithm suffers from high variance during iterations, which limits its effectiveness in finite-sample scenarios despite its asymptotic convergence. To address this, we adopt a sliding window averaging strategy that smooths the estimation of β , thereby stabilizing the algorithm. Specifically, at iteration k , the estimate is updated as

$$\Theta_k = \frac{1}{\tilde{w}} \sum_{i=k-\tilde{w}+1}^k \beta_i, \quad (20)$$

where $\tilde{w}$ is the window size. Different from the previous update method, after updating β_k , we take the average over the most recent $\tilde{w}$ values to form the stabilized estimate Θ_k .

To better represent our variable window, we first use interpolation to convert the step size into a continuous scale. The continuous scale $t = \sum_{i=1}^k \eta_i$ replaces the iteration step k of

Algorithm 2 Online Estimation of β^* via Robbins-Monro with Polyak–Ruppert averaging

```

1: Input: Step size parameters  $\bar{Y}_{lb}, \zeta > 0$ ; truncation level  $n$ ; constant  $V > 0$ 
2: Initialize:  $\Theta_1 \leftarrow 0, U_1 \leftarrow 0$ 
3: for each frame  $k = 1, 2, \dots$  do
4:    $\nu_k \leftarrow \frac{1}{V} \cdot U_k^+$ 
5:   for  $j = 1$  to  $n$  do
6:     // Sampling
7:     Update  $G'_j(w, \Theta)$  in Eq. (13)
8:     Update  $J'_j(w, \Theta)$  in Eq. (14)
9:     Solve  $v_j(\Theta)$  in Eq. (12)
10:    Solve  $\tau_{k,j}^{(j)}$  using Eq. (16)
11:    Sample  $\delta W_j \leftarrow W_{S_{k,j}+Y_{k,j}+\tau_{k,j}} - W_{S_{k,j}}$ 
12:  end for
13:  // Estimate based on sampled trajectories
14:   $X_k \leftarrow \sum_{j=1}^n \alpha^{j-1} g(\delta W_j, \Theta)$ 
15:  // Theoretically optimal step size
16:   $\eta_k \leftarrow \frac{\sum_{j=1}^n \alpha^{j-1} (Y_{k,j} + \tau_{k,j})}{k}$ 
17:  // Maintain Polyak–Ruppert averaged estimate
18:   $\Theta_k = \frac{1}{k} \sum_{i=1}^k \Theta_i$ 
19:  // Update  $\Theta_k$  with projected stochastic observation
20:   $\Theta_{k+1} \leftarrow (\Theta_k + \eta_k X_k)^+$ 
21:  // Update constraint violation severity
22:   $V_{k+1} = V_k + \left( \frac{1}{f_{max}} - \frac{1}{n} \sum_{j=1}^n (Y_{k,j} + \tau_{k,j}) \right)$ 
23: end for

```

the discrete scale. Therefore, as the time step t increases, the averaging window becomes larger and larger.

$$\beta_k(t) = \beta_{k+i}, t \in [t_{k+i} - t_k, t_{k+i+1} - t_k), i \geq 0. \quad (21)$$

Let $U_k(t) = \frac{\sqrt{\eta_k}}{t} \sum_{i=k}^{k+t/\eta_k-1} (\beta_i - \beta^*)$ denote the normalized average error. Under these conditions, we have a normalized average error that tends to zero mean and a normal distribution whose variance converges to 0 as t increases.

Acceleration via Increasing Step Size. However, the above convergence guarantee relies on the condition $\eta_k \sim o(1/k)$. To increase step size, we propose the Polyak–Ruppert averaging method, which averages all β_k starting from the first iteration, i.e., $\Theta_k = \frac{1}{k} \sum_{i=1}^k \beta_i$. Under the iterative method of PR averaging, we can not only prove the convergence of the normalized average error variance, but also find the maximum iteration step size $\eta_k = P/k$, where $P = h'_n = \frac{1}{n} \sum_{j=1}^n \alpha^{j-1} (Y_{k,j} + \tau_{k,j})$.

VI. ANALYSIS OF CONVERGENCE

Overview. In this section we theoretically analyze and accelerate the algorithm. First we analyze whether the algorithm converges to the optimal β^* in Theorem 2. Secondly, we analyze the convergence speed of the algorithm in Theorem 3. Third, we analyze whether there exists an algorithm that has a faster convergence rate than the RM algorithm in Theorem 4. Fourth, we stabilize the algorithm with variable window averaging and prove that variance of normalized average error

tends to 0. Finally, we find out the maximum iteration step size while ensuring that the variance of normalized average error converges.

Theorem 2 (Convergence Analysis): (a) The proposed algorithm learns the optimum β^* almost surely, i.e.,

$$\lim_{k \rightarrow \infty} \beta_k = \beta^*, \quad w.p.1. \quad (22)$$

(b) The average MSE under strategy π_{online} converges to mse_{opt} almost surely, i.e.,

$$\limsup_{k \rightarrow \infty} \mathbb{E} \left[\frac{\int_{R_0}^{R_{k+1}} (W_t - \hat{W}_t)^2 dt}{R_{k+1} - R_0} \right] = mse_{opt}, w.p.1. \quad (23)$$

Theorem 2 establishes the *almost sure convergence* of the iterates β_k to the optimal parameter β^* and the convergence of the corresponding MSE to its minimum value. This ensures the algorithm is asymptotically consistent.

Theorem 3 (Convergence Speed Analysis): (a) The second moment of $(\beta_k - \beta^*)$ satisfies:

$$\mathbb{E}[(\beta_k - \beta^*)^2] = O(1/k). \quad (24)$$

(b) The growth rate of the cumulative MSE optimality gap up to the k -th sample can be bounded as follows:

$$\int_{R_0}^{R_{k+1}} (W_t - \hat{W}_t)^2 dt - mse_{opt} \mathbb{E}[R_{k+1} - R_0] = O(\ln k). \quad (25)$$

Theorem 3 quantifies the *convergence rate*, showing that the variance of the estimate decays as $O(1/k)$ and that the cumulative optimality gap of MSE grows only logarithmically. This provides a sublinear rate guarantee, which is typical in stochastic approximation.

Theorem 4 (Minimax Optimality): For any distribution $\mathbb{P}$, let $\pi^*(\mathbb{P})$ denote the MSE minimum sampling strategy when the delay $Y \sim \mathbb{P}$. After k -samples are taken, the minimax optimality gap of cumulative MSE is lower bounded by:

$$\inf_{\hat{\beta}} \sup_{\mathbb{P}} \mathbb{E} [(\hat{\beta} - \beta^*(\mathbb{P}))^2] = \Omega(1/k). \quad (26)$$

The minimax estimation error mse_{opt} is lower bounded by:

$$\int_{R_0}^{R_{k+1}} (W_t - \hat{W}_t)^2 dt - mse_{opt} \mathbb{E}[R_{k+1} - R_0] = \Omega(\ln k). \quad (27)$$

Theorem 4 addresses the *minimax optimality* of our algorithm by proving that no algorithm can achieve a faster convergence rate under arbitrary delay distributions. This result confirms that the $O(1/k)$ variance decay and $\Omega(\ln k)$ MSE gap achieved in Theorem 3 are in fact optimal in the minimax sense.

Theorem 5 (Stability of Variable Window Averaging): If step size satisfies $\eta_k \sim o(1/k)$, then when $t \rightarrow \infty$, $U_k(t)$ converges to a normally distributed random variable with zero-mean and covariance $\bar{V} = (h'_n)^{-1} \sigma (h'_n)^{-1} / t$, where σ is covariance of noise $\delta h_n = \bar{h}_n - h_n$.

The result establishes that, as the continuous time $t \rightarrow \infty$, the distribution of the averaged error converges to a Gaussian distribution centered at zero with covariance $\bar{V}$, which

decays as $1/t$. This indicates that, despite the noise inherent in stochastic approximation, the averaging scheme achieves vanishing error variance in the limit, leading to stable and efficient estimates.

Theorem 6 (Acceleration and Variance Analysis): In the RM algorithm based on step size P/k with Polyak–Ruppert averaging, the normalized error $\hat{U}_k(t) = \frac{1}{k} \sum_{i=0}^k (\beta_i - \beta^*) / \sqrt{\eta_k}$ converges to a normal distribution with zero mean and variance V/k . Moreover, for a kind of step size $O(1/k)$ with PR averaging, the optimal convergence step size is P/k , where P is h'_n .

This theorem establishes two key insights regarding the application of PR averaging in stochastic approximation. First it shows that, under the PR scheme, the normalized average error $\hat{U}_k(t)$ converges to a Gaussian distribution with variance decaying at a rate of $1/k$, ensuring both asymptotic normality and bounded variance. Secondly, it reveals that the asymptotic performance is also achieved when the step size is as large as $\eta_k = O(1/k)$ and the largest step size is chosen as $\eta_k = P/k$, where $P = h'_n = \frac{1}{n} \sum_{j=1}^n \alpha^{j-1} (Y_{k,j} + \tau_{k,j})$, the Jacobian of the mean field function at equilibrium.

VII. SIMULATIONS

To evaluate the effectiveness of the proposed sampling algorithm, we simulate the MSE performance of the proposed adaptive online sampling policy $\pi_{\text{online,unreliable}}$ in Algorithm 1, and the sampling policy through PR Averaging $\pi_{\text{online,PR}}$ in Algorithm 2. We compare the MSE performance of the proposed policies with the following existing sampling strategies:

a) *Signal-independent strategies:* When the signal model is unknown, we use the age of information (AoI) to measure the freshness of the data. In simulations, we plot the MSE performance using the online AoI minimum sampling strategy $\pi_{\text{AoI,online}}$ when the channel statistics is unknown in advance [26].

b) *Signal-aware strategy:* We also consider sampling strategies designed for MSE optimization when the signal follows the Wiener process. We simulated the MSE-minimizing sampling strategy $\pi_{\text{mse,unreliable}}$ [15] when the lossy channel statistics are known, and the adaptive online sampling strategy $\pi_{\text{online,reliable}}$ [27] for reliable channels.

A. MSE without sampling frequency constraint

First, we consider the transmission delay $D \sim \mathcal{U}[0, 2]$. For each transmission attempt, the packet loss probability $\alpha = 0.1$.

Figure 2 shows the average MSE performance as a function of the number of iterations k . Compared with the signal-independent sampling policy $\pi_{\text{AoI,online}}$, our proposed $\pi_{\text{online,unreliable}}$ and $\pi_{\text{online,PR}}$ converge to a lower MSE, indicating that signal-aware strategies better capture temporally relevant information. Compared to $\pi_{\text{online, reliable}}$ that is designed specifically for reliable transmission, policies $\pi_{\text{online,unreliable}}$ and $\pi_{\text{online,PR}}$ achieve smaller MSE, highlighting the importance of correct channel models in data freshness optimization. Furthermore, the MSE obtained by $\pi_{\text{online,unreliable}}$ and $\pi_{\text{online,PR}}$

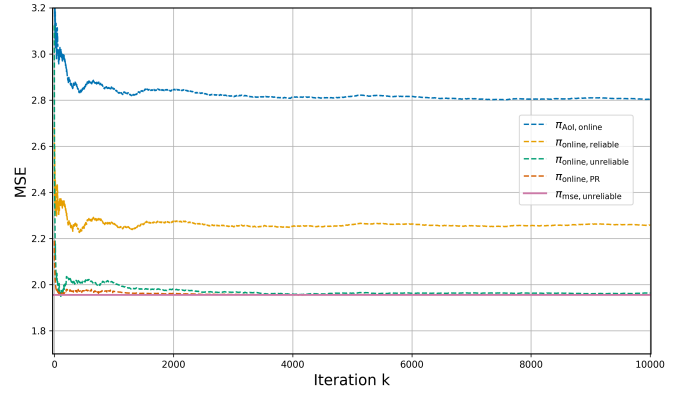


Fig. 2. MSE convergence under different sampling strategies in unreliable channels.

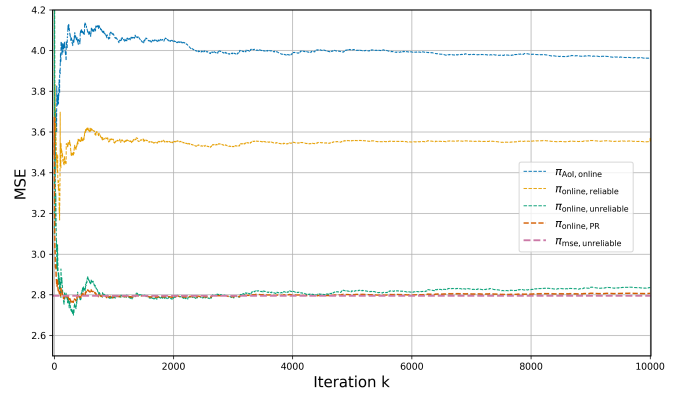


Fig. 3. The MSE performance under different sampling strategies and sampling frequency constraint in unreliable channels.

converge to $\pi_{\text{mse,unreliable}}$, which empirically confirm the theoretical convergence behavior in Theorem 2. The MSE using $\pi_{\text{online,PR}}$ converges to the optimum value at a faster speed than $\pi_{\text{online,unreliable}}$, verifying that $\pi_{\text{online,PR}}$ achieves a faster convergence speed shown in Theorem 6.

B. Sampling frequency constraint

We consider a remote estimation where the transmission delay follows a log-normal distribution. The probability density function of the log-normal distribution is given by $f_D(d) = \frac{1}{d\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln d - \mu)^2}{2\sigma^2}\right)$, where d is the transmission delay. Specifically, we set the distribution parameters as $\mu = 0.7$ and $\sigma = 0.5$, ensuring that $\mathbb{E}[D] = \exp(\mu + \sigma^2/2) \approx 2.28$. The frequency constraint $f_{\max}$ is fixed at 0.3.

Figure 3 shows the MSE convergence under sampling frequency constraint. Under the sampling constraint, the performance of $\pi_{\text{online,unreliable}}$ and $\pi_{\text{online,PR}}$ converges to the optimum value obtained by π_{mse}^* , indicating the convergence property when operating under a sampling frequency constraint. Figure 4 plots the average sampling interval as a function of the number of iterations. Notably, our PR-averaged method exhibits the fastest convergence to the constraint and maintains the most stable behavior throughout training. This

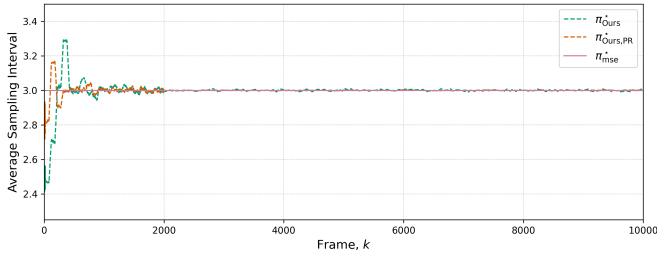


Fig. 4. The average sampling interval over time under transmission frequency constraint.

improvement can be attributed to the variance reduction effect of PR averaging, which enhances numerical stability and accelerates the satisfaction of the constraint. These results highlight the benefit of PR averaging not only in signal estimation accuracy but also in ensuring reliable compliance with long-term transmission constraints.

VIII. PROOF OF THEOREMS

A. Novelty of Theorem 2, 3, 4

Compared to prior works [2], [18], [26], [27] that focus on stable communication channels, our analysis establishes, for the first time, the convergence guarantees of online sampling algorithms under unreliable transmission environments. Although the overall proof strategy builds upon classical stochastic approximation techniques, the key technical novelty arises from the non-i.i.d. and random-length transmission failures that precede each successful sample. To ensure the convergence and convergence rate results hold, we must verify a series of technical conditions involving upper and lower bounds on the per-transmission value function J , which differ across transmissions due to the stochastic and history-dependent nature of the process. Unlike the stable setting, where a unified J function suffices, our analysis requires a sequence of non-identical J functions, each corresponding to a distinct transmission instance. This heterogeneity, combined with the implicit, strategy-dependent threshold structures induced by the unstable Wiener process, introduces substantial analytical challenges and marks a significant departure from previous methods.

B. Proof of Theorem 5

We perform a Taylor expansion on $h_{n,k}$ and substitute it into the update equation for β . We obtain:

$$\beta_{k+1} = \beta_k + \eta_k [h'_{n,k}(\beta^*, \delta W)(\beta_k - \beta^*) + o(\beta_k)]. \quad (28)$$

Since the definition of $U_k(t)$, we obtain:

$$U_k(t) = \int_0^t \nabla h_{n,k} U_k(s) ds + M_k(t) + \tilde{o}_k(t), \quad (29)$$

where $M_k(t) := \sum_{i=1}^k \frac{1}{\sqrt{\eta_i}} \eta_i \xi_i$ and $\tilde{o}_k(t) := \sum_{i=1}^k \frac{1}{\sqrt{\eta_i}} \eta_i o(\beta_i)$. $r(\beta)$ denote higher-order residual terms. We have proved $M_k(t)$ is a scaled sum of martingale differences in Proof of Theorem 2. Under bounded second

moments and independence or mixing conditions, we can use Martingale Functional Central Limit Theorem to show that covariance of $M_k(t)$ converges to $\sigma^{\frac{1}{2}} w(t)$. Since $r(\beta) = o(\|\beta - \beta^*\|)$ and $\|\beta_k - \beta^*\| = O(\sqrt{\eta_k})$, we conclude that $\tilde{o}_k(t) = \frac{1}{\sqrt{\eta_k}} \sum \eta_i r(\beta_i) \rightarrow 0$. Therefore we have constructed:

$$dU(t) = h'_{n,k} U(t) dt + \sigma^{1/2} dw(t). \quad (30)$$

Notice that $h'_{n,k} < 0$ has been proved in proof of Theorem 2. We now proceed to prove the convergence of the window-averaged process. Define the covariance matrix $R(s) = \mathbb{E}[U(t+s)U^T(t)]$, where $U(\cdot)$ is the stationary solution to Eq. (30), and define

$$\bar{R} = \mathbb{E}[U(0)U(0)] = \int_0^\infty e^{As} \sigma e^{As} ds, \quad (31)$$

where A is $h'_{n,k}(\beta^*)$. Then, for any $s \geq 0$, we have $R(s) = e^{As} \bar{R}$, $R(-s) = R(s)$ and

$$\int_{-\infty}^\infty R(s) ds = -A^{-1} \bar{R} - \bar{R} (A^{-1}) = A^{-1} \sigma A^{-1} = \bar{V}. \quad (32)$$

The first equality of Eq. (32) is obvious by evaluating the integral on the left side. The second equality of Eq. (32) follows by noting that $\bar{R}$ defined in Eq. (31) solves the algebraic Riccati equation $-\bar{R}A - A\bar{R} = \sigma$ and multiplying this on the right by $(A^{-1})^T$ and on the left by A^{-1} . Define the processes

$$\tilde{U}_k(t) = \frac{1}{\sqrt{t}} \int_0^t U_k(s) ds, \quad \tilde{U}(t) = \frac{1}{\sqrt{t}} \int_0^t U(s) ds. \quad (33)$$

Since $R(s) \rightarrow 0$ exponentially as $s \rightarrow \infty$, we can write

$$\begin{aligned} \text{Cov } \tilde{U}(t) &= \frac{1}{t} \int_0^t \int_0^t R(s-\tau) ds d\tau \\ &= \int_{-\infty}^\infty R(s) ds + O(1/t) = \bar{V} + O(1/t). \end{aligned} \quad (34)$$

The basic result on the averaged iterates is obtained by relating $U_k(t)$ to $\tilde{U}_k(t)$. Our η_k is slower than $O(1/k)$ and satisfies $\frac{\eta_k}{\eta_{k+1}} = 1 + o(\eta_k)$, which implies that for any $t < \infty$,

$$\max\{i - k : 0 \leq t_i - t_k \leq t\} \cdot \eta_k / t \rightarrow 1. \quad (35)$$

Equivalently, asymptotically, there are t/ϵ_n iterates in the interpolated time interval $[0, t]$ for $U_k(\cdot)$. The relation Eq. (35) would not hold if $\eta_k = O(1/k)$. By the definition of $\tilde{U}_k(\cdot)$, for $i \geq k$

$$\begin{aligned} \sqrt{t} \tilde{U}_k(t) &= \sum_{t_i: t_k \leq t_i \leq t} (\beta_i - \beta^*) \frac{\eta_i^{1/2}}{\eta_k} \\ &= \sum_{t_i: t_k \leq t_i \leq t} \left(\frac{\eta_i^{1/2}}{\eta_k} - \frac{1}{\eta_k^{1/2}} \right) (\beta_i - \beta^*) \\ &\quad + \frac{1}{\eta_k^{1/2}} \sum_{t_i: t_k \leq t_i \leq t} (\beta_i - \beta^*). \end{aligned} \quad (36)$$

By the weak convergence of $U_k(\cdot)$ in ODE, $\tilde{U}_k(\cdot) \Rightarrow \tilde{U}(\cdot)$. Then Eq. (35) imply that the first sum in Eq. (36) goes to

zero in probability as $k \rightarrow \infty$. By Eq. (35) and the weak convergence in ODE, the difference between $\sqrt{t}\tilde{U}_k(t)$ and the second sum in Eq. (36) goes to zero in probability as $k \rightarrow \infty$, which yields the desired conclusion. ■

C. Proof of Theorem 6

Let us suppose that the normalized error $\hat{U}_k$ converges weakly to a zero-mean Gaussian variable as $k \rightarrow \infty$, under suitable regularity assumptions. The classical convergence result [30] shows that, under the Robbins-Monro framework, convergence to the optimum β^* with asymptotic normality holds if the step size satisfies

$$\sum_k \eta_k = \infty, \quad \sum_k \eta_k^2 < \infty, \quad (37)$$

which is satisfied by $\eta_k = \frac{1}{k}$, but not by $\eta_k = \text{const}$ (which violates the second condition), nor by $\eta_k = \frac{1}{k^2}$ (which violates the first condition). Moreover, the asymptotic mean square error of the estimate β_k is typically of the order $\mathbb{E}[\|\beta_k - \beta^*\|^2] \approx \eta_k \cdot \sigma$, where σ is the asymptotic covariance matrix of the limiting distribution. This implies a key trade-off: smaller η_k reduces noise but slows convergence, while larger η_k accelerates updates but leads to higher asymptotic variance. To get the best rate, we choose P to get the best $\tilde{U}$. Under appropriate conditions, it is a classical result that $\hat{U}_k(\cdot) \Rightarrow \hat{U}_0(\cdot)$ [30], where $U_0(\cdot)$ is the stationary solution to:

$$d\hat{U}_0 = (PA + I/2)\hat{U}_0 dt + P\Sigma^{1/2}dw, \quad (38)$$

where $w(\cdot)$ is the standard Wiener process. Next, we prove it. Now, according to definition of U_k , we have

$$\begin{aligned} \hat{U}_{k+1} &= \frac{1}{k} \sum_{i=0}^k \frac{\beta_{k+1} - \beta^*}{\sqrt{\eta_{k+1}}} = \frac{1}{k} \sum_{i=0}^k \frac{\beta_k + \eta_k h_{n,k} - \beta^*}{\sqrt{\eta_{k+1}}} \\ &= \frac{1}{k} \sum_{i=0}^k \frac{\beta_k - \beta^*}{\sqrt{\eta_{k+1}}} + \frac{1}{k} \sum_{i=0}^k \frac{\eta_k}{\sqrt{\eta_{k+1}}} h_{n,k} \\ &= \sqrt{\frac{\eta_k}{\eta_{k+1}}} \hat{U}_k + \frac{1}{k} \sum_{i=0}^k \frac{\eta_k}{\sqrt{\eta_{k+1}}} h_{n,k}. \end{aligned} \quad (39)$$

Since $\eta_k = \frac{P}{k}$, we have $\sqrt{\frac{\eta_k}{\eta_{k+1}}} \approx \sqrt{1 + \frac{1}{k}} \approx 1 + \frac{1}{2k}$. Therefore, we have $\hat{U}_{k+1} = (1 + (2k)^{-1})\hat{U}_k + \sqrt{\eta_k} h_{n,k}$. Since $h_{n,k} \approx h'_{n,k}(\beta_k - \beta^*) + \xi_k$, we have

$$\hat{U}_{k+1} = (1 + (2k)^{-1})\hat{U}_k + \eta_k h'_{n,k} \hat{U}_k + \frac{1}{k} \sum_{i=0}^k \sqrt{\eta_k} \xi_k. \quad (40)$$

Therefore, we obtain:

$$\frac{\hat{U}_{k+1} - \hat{U}_k}{1/k} \approx \left(\frac{1}{2} + P\nabla h_{n,k}\right)\hat{U}_k + \frac{1}{k} \sum_{i=0}^k \sqrt{P}\xi_k. \quad (41)$$

Similar to $M_k(t)$, we have $\frac{1}{k} \sum_{i=0}^k \sqrt{\eta_k} \xi_k$ converges to $P^{\frac{1}{2}}w(t)$. Namely, when $k \rightarrow \infty$, we have

$$d\hat{U} = (PA + I/2)\hat{U}dt + P^{\frac{1}{2}}dw. \quad (42)$$

It is obviously required that $(PA + I/2)$ be a Hurwitz matrix. The stationary covariance matrix of the solution of Eq. (38) is given by Eq. (31) with $PA + I/2$ replacing A and $P^{1/2}(P^{1/2})$ replacing σ . Using $\sigma = A\bar{V}A$, it can be written as

$$\int_0^\infty e^{(B+I/2)t} B\bar{V}B e^{(B+I/2)t} dt, \quad (43)$$

where $B = PA$. By minimizing either the trace of Eq. (43) over B or Eq. (43) over B in the sense of positive definite matrices, the best value of the matrix K is given by $B = -I$, namely, $P = -A^{-1}$. With this asymptotically optimal value of K , the stationary covariance of $U_0(t)$ is just $\bar{V}$. ■

IX. CONCLUSION

In this paper, we investigate the problem of optimally sampling a Wiener process over unreliable channels with unknown transmission delays. To overcome the challenge of non-i.i.d. failures and delay uncertainty, we reformulate the original estimation objective as an infinite-horizon Markov decision problem and derive a recursive characterization of the optimal sampling cost via a threshold-based stopping strategy. The core of the offline solution lies in a novel free-boundary formulation, which allows us to solve for the optimal threshold $v_n(\beta)$ by characterizing the value function as the solution to a second-order differential equation.

Building on this, we develop an online algorithm to learn the optimal parameter β^* in the absence of delay distribution knowledge. Instead of estimating the delay distribution explicitly, we directly apply stochastic approximation to learn β^* , which determines the optimal sampling threshold. Then we enhance RM scheme stability and efficiency through sliding window and PR averaging. Our analysis establishes that the iterates β_k converge almost surely to β^* , and the resulting MSE approaches the offline optimal value.

Importantly, we provide rigorous upper and lower bounds on the convergence behavior. Specifically, we prove that the squared estimation error $\mathbb{E}[(\beta_k - \beta^*)^2] = O(1/k)$, and the cumulative MSE regret admits a tight bound of $\Theta(\ln k)$. The lower bound is established via Le Cam's method, confirming that our online algorithm is *minimax-optimal*. These results not only certify the statistical efficiency of our method but also highlight its theoretical tightness under broad uncertainty.

Overall, our work provides a unified framework for optimal sampling under delay uncertainty—grounded in stochastic control, realized through recursive algorithms, and validated by tight performance bounds.

Future Work. Our framework relies on modeling the signal as the scalar Wiener process and assumes a single-sensor setting without retransmissions or queueing mechanisms. Extending the proposed framework to handle richer signal models, retransmission protocols, and coordinated multi-source sampling under contention would be important directions for future work.

REFERENCES

- [1] S. Liyanaarachchi and S. Ulukus, "Early sampling: Sampling with incomplete queue state information to minimize the age," in *2024 58th Annual Conference on Information Sciences and Systems (CISS)*. IEEE, 2024, pp. 1–6.
- [2] H. He, H. Tang, J. Pan, J. Wang, J. Song, and L. Tassiulas, "Age optimal sampling for unreliable channels under unknown channel statistics," in *2023 21st International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*. IEEE, 2023, pp. 651–656.
- [3] D. Ogbe, C.-C. Wang, and D. J. Love, "On the optimal delay growth rate of multi-hop line networks: Asymptotically delay-optimal designs and the corresponding error exponents," *IEEE Transactions on Information Theory*, 2023.
- [4] B. Hajek, K. Mitzel, and S. Yang, "Paging and registration in cellular networks: Jointly optimal policies and an iterative algorithm," *IEEE Transactions on Information Theory*, vol. 54, no. 2, pp. 608–622, 2008.
- [5] A. Nayyar, T. Başar, D. Teneketzis, and V. V. Veeravalli, "Optimal strategies for communication and remote estimation with an energy harvesting sensor," *IEEE Transactions on Automatic Control*, vol. 58, no. 9, pp. 2246–2260, 2013.
- [6] T. Z. Ornee and Y. Sun, "Sampling and remote estimation for the ornstein-uhlenbeck process through queues: Age of information and beyond," *IEEE/ACM Transactions on Networking*, vol. 29, no. 5, pp. 1962–1975, 2021.
- [7] Y. Sun and B. Cyr, "Sampling for data freshness optimization: Non-linear age functions," *Journal of Communications and Networks*, vol. 21, no. 3, pp. 204–219, 2019.
- [8] C.-H. Tsai and C.-C. Wang, "Unifying aoi minimization and remote estimation—optimal sensor/controller coordination with random two-way delay," *IEEE/ACM Transactions on Networking*, vol. 30, no. 1, pp. 229–242, 2021.
- [9] K. J. Astrom and B. M. Bernhardsson, "Comparison of riemann and lebesgue sampling for first order stochastic systems," in *Proceedings of the 41st IEEE Conference on Decision and Control*, 2002., vol. 2. IEEE, 2002, pp. 2011–2016.
- [10] K. Nar and T. Başar, "Sampling multidimensional wiener processes," in *53rd IEEE Conference on Decision and Control*. IEEE, 2014, pp. 3426–3431.
- [11] J. Chakravorty and A. Mahajan, "Distortion-transmission trade-off in real-time transmission of gauss-markov sources," in *2015 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2015, pp. 1387–1391.
- [12] Y. Sun, Y. Polyanskiy, and E. Uysal, "Sampling of the wiener process for remote estimation over a channel with random delay," *IEEE Transactions on Information Theory*, vol. 66, no. 2, pp. 1118–1135, 2019.
- [13] A. Arafat, K. Banawan, K. G. Seddik, and H. V. Poor, "Sample, quantize, and encode: Timely estimation over noisy channels," *IEEE Transactions on Communications*, vol. 69, no. 10, pp. 6485–6499, 2021.
- [14] M. Klügel, M. H. Mamduhi, S. Hirche, and W. Kellerer, "Aoi-penalty minimization for networked control systems with packet loss," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2019, pp. 189–196.
- [15] J. Pan, Y. Sun, and N. B. Shroff, "Sampling for remote estimation of the wiener process over an unreliable channel," *ACM Sigmetrics 2024, Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 7, no. 3, pp. 1–41, 2023.
- [16] B. Li, "Efficient learning-based scheduling for information freshness in wireless networks," in *IEEE INFOCOM 2021-IEEE Conference on Computer Communications*. IEEE, 2021, pp. 1–10.
- [17] V. Tripathi and E. Modiano, "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*, 2021, pp. 241–250.
- [18] H. Tang, Y. Chen, J. Wang, P. Yang, and L. Tassiulas, "Age optimal sampling under unknown delay statistics," *IEEE Transactions on Information Theory*, vol. 69, no. 2, pp. 1295–1314, 2023.
- [19] C.-H. Tsai and C.-C. Wang, "Age-of-information revisited: Two-way delay and distribution-oblivious online algorithm," in *2020 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2020, pp. 1782–1787.
- [20] S. Banerjee, R. Bhattacharjee, and A. Sinha, "Fundamental limits of age-of-information in stationary and non-stationary environments," in *2020 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2020, pp. 1741–1746.
- [21] E. U. Atay, I. Kadota, and E. Modiano, "Aging wireless bandits: Regret analysis and order-optimal learning algorithm," in *2021 19th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*. IEEE, 2021, pp. 1–8.
- [22] S. Fatale, K. Bhandari, U. Narula, S. Moharir, and M. K. Hanawal, "Regret of age-of-information bandits," *IEEE Transactions on Communications*, vol. 70, no. 1, pp. 87–100, 2021.
- [23] S. Leng and A. Yener, "Age of information minimization for wireless ad hoc networks: A deep reinforcement learning approach," in *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2019, pp. 1–6.
- [24] M. A. Abd-Elmagid, H. S. Dhillon, and N. Pappas, "A reinforcement learning framework for optimizing age of information in rf-powered communication systems," *IEEE Transactions on Communications*, vol. 68, no. 8, pp. 4747–4760, 2020.
- [25] E. T. Ceran, D. Gündüz, and A. György, "A reinforcement learning approach to age of information in multi-user networks with harq," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1412–1426, 2021.
- [26] H. Tang, Y. Chen, J. Sun, J. Wang, and J. Song, "Sending timely status updates through channel with random delay via online learning," in *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*. IEEE, 2022, pp. 1819–1827.
- [27] H. Tang, Y. Sun, and L. Tassiulas, "Sampling of the wiener process for remote estimation over a channel with unknown delay statistics," in *Proceedings of the Twenty-Third International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*, 2022, pp. 51–60.
- [28] H. V. Poor, *An introduction to signal detection and estimation*. Springer Science & Business Media, 2013.
- [29] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the wiener process over a channel with random delay," in *2017 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2017, pp. 321–325.
- [30] T. L. Lai, "Stochastic approximation," *The annals of Statistics*, vol. 31, no. 2, pp. 391–406, 2003.