Timely Remote Estimation and Applications to Situational Awareness

by

Tasmeen Zaman Ornee

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
May 4, 2024

Keywords: Remote estimation, age of information, Gauss-Markov processes, restless multi-armed bandit

Approved by

Yin Sun, Committee Chair, Associate Professor of Electrical and Computer Engineering, Auburn University
Anthony Ephremides, Distinguished University Professor Emeritus, Electrical and Computer Engineering, University of Maryland
Jitendra Tugnait, Professor of Electrical and Computer Engineering, Auburn University
Shiwen Mao, Professor of Electrical and Computer Engineering, Auburn University
SueAnne Griffith, Assistant Professor of Electrical and Computer Engineering, Auburn University
Yazhou Tu, Assistant Professor of Computer Science and Software Engineering, Auburn University

Abstract

In real-time monitoring and networked control systems, sensor observations from vehicles, robots, UAVs, or stock markets, are transmitted to a monitoring or controlling unit, which could be any kind of decision-making device. Real-time services often require fresh and timely data which are usually in the form of a signal. A key performance metric characterizing data freshness is the Age of Information (AoI). However, data signals can exhibit diverse behavior, sometimes evolving slowly and later on evolving very quickly. Therefore, only considering the time difference is insufficient to characterize the variation of a signal. In this dissertation, we investigate the performance of a remote estimation system by considering both the data signal value and its timeliness.

First, we consider the sampling problem for the remote estimation of a scalar Gauss-Markov process. The optimal sampling problem is a constrained continuous-time Markov Decision Process (MDP) with an uncountable state space. Our analysis reveals that the optimal sampling policy is a threshold policy on instantaneous estimation error and the threshold is found. If the sampler has no knowledge of the process, the optimal sampling problem reduces to an MDP for minimizing nonlinear age functions. In both problems, the optimal sampling policies can be computed by low-complexity algorithms.

Next, We generalize this study from single-source, single-channel to multiple-source, multiple-channel and formulate a scheduling problem for the remote estimation of multiple Gauss-Markov processes. This problem is a continuous-time Restless Multi-armed Bandit (RMAB) with a continuous state space. We prove that all bandits are indexable and derive an exact expression of the Whittle index. Our results unite two theoretical frameworks that are used for remote estimation and AoI minimization: threshold-based sampling and Whittle index-based scheduling. In these investigations, the numerical evidence shows that our proposed policy achieves high-performance gain over existing policies.

Finally, we study a scheduling problem for maximizing situational awareness in safety-critical systems where a centralized monitor pulls updates from multiple agents monitoring several safety-critical situations. Based on the received updates, multiple estimators determine the current safety-critical situations. We provide a novel framework that quantifies

the loss due to the unawareness of potential danger which depends on the AoI and the observed signal value. To minimize the penalty, we study an RMAB problem and provide a low-complexity scheduling algorithm that is asymptotically optimal. Numerical evidence shows that our scheduling policy can achieve up to 100 times performance gain over periodic updating and up to 10 times over randomized policy.

Acknowledgments

I would like to express my heartiest gratitude to my advisor, Prof. Yin Sun, for his unwavering support, invaluable guidance, and encouragement throughout my Ph.D. journey. It has been an honor to work with him. Under his guidance, I gained invaluable experience in conducting research and refining my writing and presentation skills, and I found the courage and motivation to solve challenging problems.

I am also immensely grateful to my committee members, Prof. Anthony Ephremides, Prof. Jitendra Tugnait, Prof. Shiwen Mao, and Prof. SueAnne Griffith, for their insightful feedback. I am also grateful to Prof. Yazhou Tu for serving as the University reader and reviewing my work. I appreciate their valuable suggestions and time to improve my work.

I express my gratitude to the Dept. of ECE at Auburn University for providing me with the resources and environment to conduct this research. I am also thankful to my current and former colleagues and collaborators: Dr. Clement Kam, Dr. Ahmed Bedewy, Shaoyi Li, and Md Kamran Chowdhury Shisher.

I would also like to express my gratitude to the following funding agencies: National Science Foundation (NSF), Office of Naval Research (ONR), and Army Research Office (ARO) for providing funding support during my Ph.D.

Finally, I express my heartfelt appreciation to my parents, Md Moniruzzaman and Ferdousi Islam for believing in me, and my brother Tasnim-uz-zaman Ornob for supporting me. I also want to express my sincere love to my grandfathers Md Alek Chand and Dr. Nurul Islam. I believe they must be very proud of me. I want to express my gratitude to my grandmothers Helana Begum and Solema Begum, my aunt Fatema Tuz Zohora, my uncles Md Shafiqul Islam, Md Jobayer Islam, Md Jonayed Islam, and Md. Aklasur Rahman. Last but not least, I want to thank my husband Md Kamran Chowdhury Shisher for his utmost support, patience, and encouragement throughout this journey. His presence has been a source of strength and I am deeply grateful to him.

Table of Contents

List of Figures

Chapter 1

Introduction

Networked control and cyber-physical systems have become extensively dominant over the recent years. This rapid growth underscores the significance of timely updates of the system states for accurate state estimation and effective decision making. For instance, a timely and precise estimate of the nearby vehicles and pedestrians is essential in autonomous driving. The real-time knowledge of surgical robot movements is crucial in remote surgery. Beyond these domains, such as UAV navigation, factory automation, environment watch, augmented/virtual reality applications, etc, real-time state estimation in optimizing the performance of networked systems is of paramount importance.

To evaluate the freshness of state updates, the concept of *Age of Information*, or simply *age*, was introduced to measure the timeliness of state updates received from a remote transmitter [1–3]. Let $U(t)$ be the generation time of the freshest received state update at time t. The age of information, as a function of t, is defined as $\Delta(t) = t - U(t)$, which is the time difference between the freshest updates available at the transmitter and receiver. Recently, the AoI has emerged as a crucial metric for the extensive applications of state updates among systems connected over networks. As shown in Figure 1.1, the age $\Delta(t)$ grows linearly with time and drops to a smaller value whenever a packet is delivered. In addition to the linear AoI $\Delta(t)$, recent advancements have shown that nonlinear functions $p(\Delta(t))$ of the AoI can serve as valuable metrics for information freshness in signal estimation, control, and, wireless communications (e.g., the freshness of channel state information).

In many real-time systems, the information of interest — e.g., the trajectory of UAV mobility trajectory, sensor measurements, and stock prices is conveyed through the value of a time-varying signal X_t, which may change slowly at some time and exhibit more rapid fluctuations later. Hence, the time difference described by the AoI $\Delta(t) = t - U(t)$ or its nonlinear functions cannot fully characterize how much the signal value has varied during the same time period, i.e., $X_t - X_{U_t}$. Hence, the status-update policy that minimizes the AoI is insufficient for minimizing the signal estimation error. One important research question is *how to design efficient status updating policies that improve the system performance better*

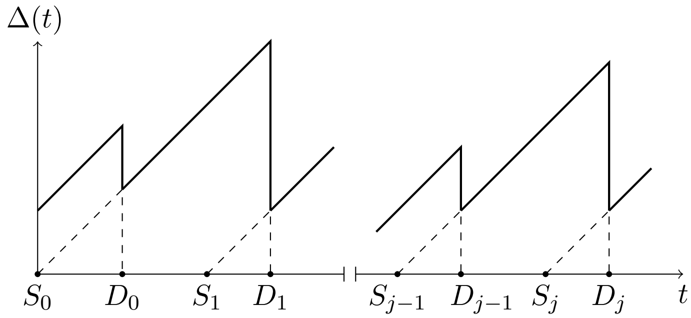


Figure 1.1: An evolution of the age $\Delta(t)$ over time where S_i is the generation time of the i-th packet and D_i is the delivery time of the i-th packet.

beyond just relying on AoI-based status updating policies, especially in more complex real-time scenarios. Motivated by this, we seek the answer to the above-mentioned research question to find structural properties of efficient status updating policies to handle more general signal models.

In this dissertation, we aim to answer the following questions: *(i) how to design an optimal sampling policy for the remote estimation of scalar Gauss-Markov signal process over a random delay channel*, *(ii) how to design a sampling and transmission scheduling policy for the remote estimation of multiple Gauss-Markov processes over multiple random delay channels*, and *(iii) how to design a transmission scheduling policy to maximize the situational awareness in safety-critical systems.*

1.1 Outline and Main Contributions

In Chapter 2, we answer the first question: *(i) how to design an optimal sampling policy for the remote estimation of the Gauss-Markov signal process over a random delay channel?* To answer this question, we consider a remote estimation system, where samples of a scalar Gauss–Markov signal are taken at a source node and forwarded to a remote estimator through an *i.i.d.* random delay channel. The estimator reconstructs an estimate of the real-time signal value from causally received samples. The optimal sampling policy for minimizing the mean square estimation error is a threshold policy, in which a new sample is taken once the instantaneous estimation error exceeds a predetermined threshold. When the sampler does not know current and history signal values, the optimal sampling problem reduces to a nonlinear AoI minimization problem. A new sample is taken in the AoI-optimal sampling policy once the expected estimation error exceeds a threshold. The threshold can be computed by low-complexity algorithms, and the insights behind these algorithms are provided. These optimal sampling results were established (i) for general service time distributions of the queueing server, (ii) for both stable and unstable scalar Gauss–Markov signals, and (iii) for sampling problems both with and without a sampling rate constraint. In chapter 3, we consider a noisy sample of the scalar Gauss-Markov process over a noisy channel. We obtain the performance upper and lower bounds for the remote estimation of the noisy Gauss-Markov process.

In Chapter 4, we answer the second question: *ii) how to design a sampling and transmission scheduling policy for the remote estimation of multiple Gauss-Markov processes over multiple random delay channels?* To answer this question, we generalize the sampling problem of scalar Gauss-Markov signals in multi-source, multi-channel scenario, where a scheduler determines when to take samples from multiple Gauss-Markov processes and send them to remote estimators over multiple *i.i.d.* random delay channels. The objective of the scheduler is to minimize the weighted sum of the time-average expected estimation errors of these Gauss-Markov sources. This problem is a continuous-time Restless Multi-armed Bandit (RMAB) with a continuous state space. We prove that all bandits are indexable and derive an exact expression of the Whittle index. To the extent of our knowledge, this is the first Whittle index policy for multi-source signal-aware remote estimation of Gauss-Markov processes. We further investigate signal-agnostic remote estimation and develop a Whittle index policy for multi-source AoI minimization. Our results unite two theoretical frameworks that were used for remote estimation and AoI minimization: threshold-based sampling and Whittle index-based scheduling. In the single-source, single-channel scenario, we demonstrate that the optimal solution to the sampling and scheduling problem can be equivalently expressed as both a threshold-based sampling strategy and a Whittle index-based scheduling policy. Notably, the Whittle index is equal to zero if and only if two conditions are satisfied: (i) the channel is idle, and (ii) the estimation error is precisely equal to the threshold in the threshold-based sampling strategy.

In Chapter 5, we answer the third question: *(iii) how to design a transmission scheduling policy to maximize the situational awareness in safety-critical systems?* To answer this question, we investigate a status-updating system consisting of multiple agent-estimator pairs. A centralized monitor pulls updates from multiple agents that are monitoring several safety-critical situations (e.g., carbon monoxide density in forest fire detection, machine safety in industrial automation, and road safety). Based on the received updates, multiple estimators determine the current safety-critical situations. Due to transmission errors and limited communication resources, the updates may not be fresh, resulting in the possibility of misunderstanding the current situation. In particular, if a dangerous situation is misinterpreted as safe, the safety risk is high. In this study, we introduce a novel framework that quantifies the penalty due to the unawareness of a potentially dangerous situation. This situation-unaware penalty function depends on two key factors: the AoI and the observed signal value. To minimize the penalty, we study a pull-based multi-agent, multi-channel transmission scheduling problem. Our analysis reveals that for optimal estimators, it is always beneficial to keep the channels busy. Due to communication resource constraints, the scheduling problem can be modeled as a Restless Multi-armed Bandit (RMAB) problem. By utilizing relaxation and Lagrangian decomposition of the RMAB, we provide a low-complexity scheduling algorithm that is asymptotically optimal. Our results hold for both reliable and unreliable channels. Numerical evidence shows that our scheduling policy can achieve up to 100 times performance gain over periodic updating and up to 10 times over randomized policy.

1.2 Literature Review

Next, I present the literature review of prior works.

AoI-based Sampling and Scheduling

There exists a significantly large number of studies on the AoI $\Delta(t)$, e.g., [1, 4–25]. In [1], the authors provided a simple example of a status updating system, where samples of a Wiener process W_t are forwarded to a remote estimator. The age of the delivered sample is $\Delta(t) = t - U(t)$ if $U(t)$ is the generation time of the latest received sample. Furthermore, the MMSE estimate of W_t is $\hat{W}_t = W_{U(t)}$ and the variance of this estimator is $\mathbb{E}[(W_t - \hat{W}_t)^2] = \Delta(t)$. In [5], the authors proposed a sampling policy for a discrete-time source process by incorporating mutual information as a measure for maximizing the information freshness. The results in [5] were further extended for both continuous and discrete-time source processes in [4] where the non-linear functions of the age were used to measure data freshness. In [6], sampling and scheduling policy for multi-source systems was studied by analyzing the peak age and peak average age. In [7], the authors analyzed the status age when the message may take various routes in the network for queueing systems. In [8], the optimal control for information updates traveled from a source to a remote destination was studied and the optimal tradeoff between the updated policy and the AoI was found. The authors also showed that in many cases, the optimal policy is to wait a certain amount before sending the next update. The average age and average peak age have been analyzed for various queueing systems in, e.g., [1, 7, 9, 10]. The optimality of the Last-Come, First-Served (LCFS) policy, or more generally the Last-Generated, First-Served (LGFS) policy, was established for various queueing system models in [13–15, 19]. Optimal sampling policies for minimizing non-linear age functions were developed in, e.g., [4, 5, 8, 24]. Age-optimal transmission scheduling of wireless networks was investigated in, e.g., [11, 12, 16–18, 20, 21]. In [25], a game-theoretic perspective of the age was studied and the authors proposed a sampling policy by studying the timeliness of the status update where an attacker sabotages the system by jamming the channel and maximizing the age of information. which does not have a signal model. A broad survey in the area of AoI is presented in [26].

Remote Estimation

The results in this dissertation also have a tight connection with the area of remote estimation, e.g., [27–33] by adding a queue between the sampler and estimator. In [27], remote state estimation in first-order linear time-invariant (LTI) discrete-time systems was considered with a quadratic cost function and finite time horizon. They showed that a time-dependent threshold-based sampler and Kalman-like estimator are jointly optimal. In [28], the authors investigated the joint optimization of paging and registration policies in cellular networks, which is essentially the same as a joint sampling and estimation optimization problem with an indicator-type cost function and an infinite time horizon. They used majorization theory and Riesz's rearrangement inequality to show that, if the state process is modeled as a symmetric or Gaussian random walk, a threshold-based sampler and a nearest distance estimator are jointly optimal. This is the first study pointing out that the sampler and estimator have different information patterns. In [29], The authors considered a remote estimation problem with an energy-harvesting sensor and a remote estimator, where the sampling decision at the sensor is constrained by the energy level of the battery. They proved that an energy-level dependent threshold-based sampler and a Kalman-like estimator are jointly optimal. In [30], [31], optimal sampling of Wiener processes was studied, where the transmission time from the sampler to the estimator is zero. Optimal sampling of OU processes was also considered in [30], which is solved by discretizing time and using dynamic programming to solve the discrete-time optimal stopping problems. In [34], the optimal sampler of OU processes is obtained analytically. In the optimal sampling policy, sampling is suspended when the server is busy and is reactivated once the server becomes idle. In addition, the threshold precisely was also characterized. The optimal sampling policy for the Wiener process in [35] is a limiting case. Remote estimation of the Wiener process with random two-way delay was considered in [36]. Remote estimation over several different channel models was recently studied in, e.g., [32, 33]. In [27–34], the optimal sampling policies were proven to be threshold policies. Because of the queueing model, the optimal sampling policy in [34] has a different structure from those in [27–33]. Specifically, In [37], a jointly optimal sampler, quantizer, and estimator design was found for a class of continuous-time Markov processes under a bit-rate constraint. In [38], the quantization and coding schemes on the estimation performance are studied. A recent survey on remote estimation systems was presented in [39].

Restless Multi-armed Bandit Problems

AoI-based scheduling for timely status updating has been studied extensively in, e.g., [11, 16, 24, 40–47]. In [11], the authors showed that under inference constraints, the scheduling problem for minimizing the age in wireless networks is NP-hard. In [40], the authors minimized the weighted-sum peak AoI in a multi-source status updating system, subject to constraints on per-source battery lifetime. A joint sampling and scheduling problem for minimizing increasing AoI functions was considered in [24]. When the system state follows a binary ON-OFF Markov process, Whittle index scheduling policies for remote estimation were developed in [48]. AoI minimization in single-hop networks was considered in [45]. AoI-based scheduling with timely throughput constraints was considered in [16]. A Whittle index-based scheduling algorithm for minimizing AoI for stochastic arrivals was considered in [42]. In [43], [41], the Whittle index policy to minimize age functions for reliable and unreliable channels was proposed. A Whittle index policy for multiple source scheduling for binary Markov sources was studied in [47]. For signal-agnostic remote estimation, a Whittle index policy was obtained in [49] for minimizing increasing AoI functions. In [46], the authors proposed a Whittle index policy for minimizing non-monotonic AoI functions. Besides Whittle index-based policies that require an indexability condition, non-indexable scheduling policies were also studied in [50–54]. In this dissertation, we solve two RMAB problems in Chapter 4 (by establishing indexability and developing a Whittle Index policy) and Chapter 5 (by developing a Maximum Gain First policy that does not need to satisfy indexability).

Related Metrics other than AoI

There exists a large number of studies on minimizing linear and nonlinear AoI functions [4, 8, 24, 26, 34, 55, 56]. One limitation of AoI is that it only captures the timeliness of the information while neglecting the actual influence of the conveyed information. To address this, several performance metrics were introduced in conjunction with AoI [26, 47, 57–64]. In [58], the concept of Age of Incorrect Information (AoII) was introduced which is characterized as a function of both age and estimation error. In [57], Age of Synchronization (AoS) was considered along with AoI to measure the freshness of a local cache. Urgency of Information (UoI) was proposed in [59] that captures the context-dependence of the status information along with AoI. Version AoI was introduced in [26] which represents how many versions are outdated at the receiver compared to the transmitter. An AoI at Query (QAoI) metric was investigated in [60], [62], [63] to capture the freshness only when required in a pull-based communication system. Value of Information (VoI), defined by the Shannon mutual information was investigated in [64]. In [61], the authors studied the cost of actuation error which is a goal-oriented measure to capture the costs associated with decisions. Shanon conditional entropy was utilized as the performance metric in [47] as the Uncertainty of Information. In addition, several research papers studied information-theoretic measures to evaluate the impact of information freshness along with information content [4, 46, 47, 54, 64–67]. In [4, 64–66], the authors employed Shannon's mutual information to quantify the information carried by received data messages regarding the current signal value at the source and used Shannon's conditional entropy to measure the uncertainty about the current signal value. Based on the studies of [4, 64–66], the authors in [47] utilized Uncertainty of Information (UoI) by using the Shannon's conditional entropy. However, there exists a disparity between these information-theoretic metrics and the performance of real-time applications such as remote estimation and inference. In [46,54,67], a generalized conditional entropy associated with a loss function L, or L-conditional entropy $H_L(Y_t|, \Delta(t), X_{t-\Delta(t)})$ was utilized to address this disparity, where Y_t is the true state of the source and $X_{t-\Delta(t)}$ is the observed value. In Chapter 5, we consider a signal-aware scheduling scheme while the earlier studies focused on signal-agnostic scenarios.

1.3 Thesis Organization

The thesis is organized as follows: In Chapter 2, we develop a low-complexity sampling policy for the remote estimation of scalar Gauss-Markov process in single-source, single-channel for general *i.i.d.* service time distributions. In Chapter 3, we extend this result by considering noisy samples and noisy channels for which we derive an explicit expression of the performance upper and lower bounds. In Chapter 4, we generalize the sampling problem for the remote estimation of Gauss-Markov processes in single-source, single-channel case to transmission scheduling problem in multi-source, multi-channel case. In Chapter 5, we consider a status updating problem for safety critical systems to maximize situational awareness. Finally, we present the concluding remarks and possible future research directions in Chapter 6.

Chapter 2

Remote Estimation of Single Gauss-Markov Process over Single Channel

2.1 Introduction

Timely updates of the system state are of significant importance for state estimation and decision-making in networked control and cyber-physical systems, such as UAV navigation, robotics control, mobility tracking, and environment monitoring systems. To evaluate the freshness of state updates, the concept of AoI was introduced to measure the timeliness of state samples received from a remote transmitter [1–3]. Let $U(t)$ be the generation time of the freshest received state sample at time t. The AoI, as a function of t, is defined as $\Delta(t) = t - U(t)$, which is the time difference between the freshest samples available at the transmitter and receiver.

Recently, the AoI concept has received significant attention, because of the extensive applications of state updates among systems connected over communication networks. The states of many systems, such as UAV mobility trajectory and sensor measurements, are in the form of a signal X_t, that may change slowly at some time and vary more dynamically later. Hence, the time difference described by the age $\Delta(t) = t - U(t)$ only partially characterizes the variation $X_t - X_{U(t)}$ of the system state, and the state update policy that minimizes the AoI does not minimize the state estimation error. This result was first shown in [35], where a sampling problem of Wiener processes was solved and the optimal sampling policy was shown to have an intuitive structure. As the results therein hold only for signals that can be modeled as a Wiener process, one would wonder how to, and whether it is possible to, extend [35] for handling more general signal models.

In this chapter, we generalize [35] by exploring the problem of sampling a Gauss-Markov process X_t. From the obtained results, we hope to find useful structural properties of the optimal sampler design that can be potentially applied to more general signal models. The Gauss-Markov process X_t is defined as the solution to the stochastic differential equation

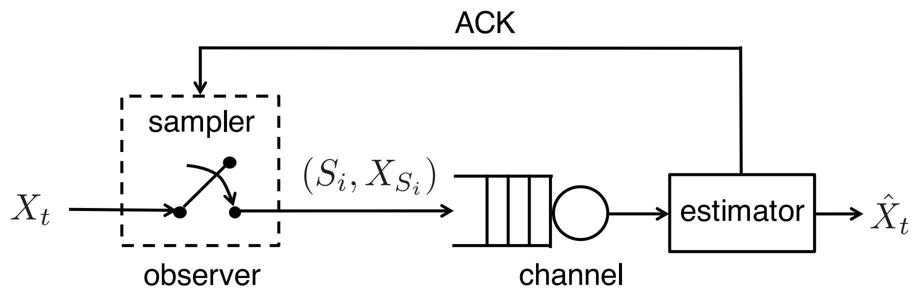


Figure 2.1: A single-source, single-channel remote estimation system.

(SDE) [68, 69]

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t, \tag{2.1}$$

where μ, θ, and $\sigma > 0$ are parameters and W_t represents a Wiener process. Based on the parameter θ, X_t can be classified into three different cases. If $\theta > 0$, then it is called a stable Ornstein-Uhlenbeck (OU) process which is the continuous-time analog of the well-known first-order autoregressive process, i.e., AR(1) process. It is the only nontrivial continuous-time process that is stationary, Gaussian, and Markovian [69]. If $\theta = 0$, then it becomes a Wiener process. If $\theta < 0$, it is known as the unstable Ornstein-Uhlenbeck (OU) process. Examples of first-order systems that can be described as the Gauss-Markov process include interest rates, currency exchange rates, and commodity prices (with modifications) [70], control systems such as node mobility in mobile ad-hoc networks, robotic swarms, and UAV systems [71, 72], and physical processes such as the transfer of liquids or gases in and out of a tank [27].

As shown in Figure 2.1, samples of a Gauss-Markov process are forwarded to a remote estimator through a channel in a first-come, first-served (FCFS) fashion. The samples experience *i.i.d.* random transmission times over the channel, which is caused by random sample size, channel fading, interference, congestions, etc. For example, UAVs flying close to WiFi access points may suffer from long communication delay and instability issues, because they receive strong interference from the WiFi access points [73]. We assume that at any time only one sample can be served by the channel. The samples that are waiting to be sent are stored in a queue at the transmitter. Hence, the channel is modeled as an FCFS queue with *i.i.d.* service times. The service time distributions considered in this paper are quite general: they are only required to have a finite mean. This queueing model is helpful in analyzing the robustness of remote estimation systems with occasionally long transmission times.

The estimator utilizes causally received samples to construct an estimate $\hat{X}_t$ of the real-time signal value X_t. The quality of remote estimation is measured by the time-average mean-squared estimation error, i.e.,

$$\mathsf{mse} = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T (X_t - \hat{X}_t)^2 dt\right]. \tag{2.2}$$

Our goal is to find the optimal sampling policy that minimizes mse by causally choosing the sampling times subject to a maximum sampling rate constraint. In practice, the cost (e.g., energy, CPU cycle, storage) for state updates increases with the average sampling rate. Hence, we are striving to find the optimum tradeoff between estimation error and update cost. In addition, the unconstrained problem is also solved. The contributions of this paper are summarized as follows:

- The optimal sampling problem for minimizing the mse under a sampling rate constraint is formulated as a constrained continuous-time Markov decision process (MDP) with an uncountable state space. Because of the curse of dimensionality, such problems often lack low-complexity solutions that are arbitrarily accurate. However, we were able to solve this MDP exactly: The optimal sampling policy is proven to be a threshold policy on *instantaneous* estimation error, where the threshold is a non-linear function $v(\beta)$ of a parameter β. The value of β is equal to the summation of the optimal objective value of the MDP and the optimal Lagrangian dual variable associated with the sampling rate constraint. If there is no sampling rate constraint, the Lagrangian dual variable is zero, and hence β is exactly the optimal objective value. Among the technical tools developed to prove this result is a free boundary method [74], [75] for finding the optimal stopping time of diffusion processes.

- The optimal sampler design of the Wiener process in [35] is a limiting case of the above result. By comparing the optimal sampling policies of OU process and the Wiener process, we find that the threshold function $v(\beta)$ changes according to the signal model, where the parameter β is determined in the same way for both signal models.

- Further, we consider a class of signal-agnostic sampling policies, where the sampling times are determined without using knowledge of the signal value of the observed Gauss-Markov process; the parameters of the process are known. The optimal signal-agnostic sampling problem is equivalent to an MDP for minimizing the time-average of a nonlinear age function $p(\Delta(t))$, which has been solved recently in [4]. The age-optimal sampling policy is a threshold policy on *expected* estimation error, where the threshold function is simply $v(\beta) = \beta$ and the parameter β is determined in the same way as above.

- The above results hold for (i) general service time distributions with a finite mean and (ii) sampling problems both with and without a sampling rate constraint. Numerical results suggest that the optimal sampling policy is better than zero-wait sampling and classic uniform sampling.

One interesting observation from these results is that the threshold function $v(\beta)$ varies with respect to the signal model and sampling problem, but the parameter β is determined in the same way.

2.2 Model

This section describes the single-source, single-channel model as shown in Figure 2.1.

2.2.1 System Model

We consider the remote estimation system illustrated in Figure 2.1, where an observer takes samples from a Gauss-Markov process X_t and forwards the samples to an estimator through a communication channel. The channel is modeled as a single-server FCFS queue with *i.i.d.* service times. The system starts to operate at time $t = 0$. The i-th sample is generated at time S_i and is delivered to the estimator at time D_i with a service time Y_i, which satisfy $S_i \leq S_{i+1}$, $S_i + Y_i \leq D_i$, $D_i + Y_{i+1} \leq D_{i+1}$, and $0 < \mathbb{E}[Y_i] < \infty$ for all i.

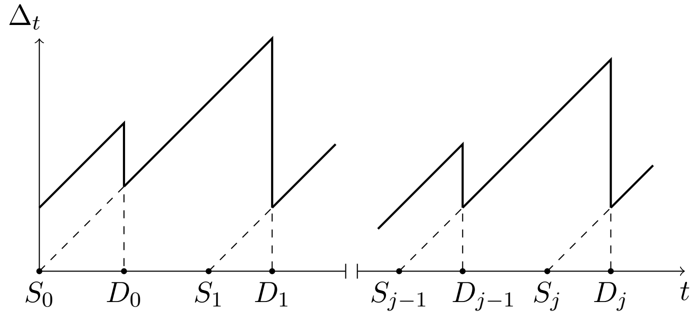


Figure 2.2: Evolution of the age $\Delta(t)$ over time.

Each sample packet (S_i, X_{S_i}) contains the sampling time S_i and the sample value X_{S_i}. Let $U(t) = \max\{S_i : D_i \leq t\}$ be the sampling time of the latest received sample at time t. The *age of information*, or simply *age*, at time t is defined as [1, 2]

$$\Delta(t) = t - U(t) = t - \max\{S_i : D_i \leq t\}, \tag{2.3}$$

which is shown in Fig. 2.2. Because $D_i \leq D_{i+1}$, $\Delta(t)$ can be also expressed as

$$\Delta(t) = t - S_i, \text{ if } t \in [D_i, D_{i+1}), \; i = 0, 1, 2, \ldots \tag{2.4}$$

The initial state of the system is assumed to satisfy $S_0 = 0$, $D_0 = Y_0$, X_0 and $\Delta(0)$ are finite constants. The parameters μ, θ, and σ in (2.1) are known at both the sampler and estimator.

Let $I_t \in \{0, 1\}$ represent the idle/busy state of the server at time t. We assume that whenever a sample is delivered, an acknowledgment is sent back to the sampler with zero delay. By this, the idle/busy state I_t of the server is known at the sampler. Therefore, the information that is available at the sampler at time t can be expressed as $\{X_s, I_s : 0 \leq s \leq t\}$.

2.2.2 Sampling Policies

In causal sampling policies, each sampling time S_i is determined based on the up-to-date information that is available at the sampler, without using any future information. In probability theory, such sampling times are represented by *stopping times*.

To define *stopping time* precisely, the concepts of σ-field and filtration are needed. Let us define the σ-field

$$\mathcal{N}_t = \sigma(X_s, I_s : 0 \leq s \leq t),$$

which is the set of events whose occurrence is determined by the realization of the process $\{X_s, I_s, 0 \leq s \leq t\}$ up to time t. Filtration is a non-decreasing sequence of σ-fields. Our analysis requires a strong Markov property, which is satisfied when the filtration is right-continuous. Define

$$\mathcal{N}_t^+ = \cap_{s>t} \mathcal{N}_s, \tag{2.5}$$

then $\{\mathcal{N}_t^+, t \geq 0\}$ is a right-continuous filtration of the information process $\{X_s, I_s, t \geq 0\}$ [76]. In a causal sampling policy, each sampling time is a stopping time with respect to $\{\mathcal{N}_t^+, t \geq 0\}$, i.e.,

$$\{S_i \leq t\} \in \mathcal{N}_t^+, \ \forall t \geq 0. \tag{2.6}$$

In other words, whether sample i has been generated by time t (i.e., whether $\{S_i \leq t\}$ or $\{S_i > t\}$) is determined by the realization of the process $\{X_s, I_s, 0 \leq s \leq t\}$ up to time t.

Let $\pi = (S_1, S_2, ...)$ represent a sampling policy. We use Π to represent the set of *causal* sampling policies that satisfy two conditions: (i) Each sampling policy $\pi \in \Pi$ satisfies (2.6) for all i. (ii) The sequence of inter-sampling times $\{T_i = S_{i+1} - S_i, i = 0, 1, \ldots\}$ forms a *regenerative process* [77, Section 6.1]: There exists an increasing sequence $0 \leq k_1 < k_2 < \ldots$ of almost surely finite random integers such that the post-k_j process $\{T_{k_j+i}, i = 0, 1, \ldots\}$ has the same distribution as the post-k_0 process $\{T_{k_0+i}, i = 0, 1, \ldots\}$ and is independent of the pre-k_j process $\{T_i, i = 0, 1, \ldots, k_j - 1\}$; further, we assume that $\mathbb{E}[k_{j+1} - k_j] < \infty$, $\mathbb{E}[S_{k_1}] < \infty$, and $0 < \mathbb{E}[S_{k_{j+1}} - S_{k_j}] < \infty, \ j = 1, 2, \ldots$[1]

[1]We will optimize $\limsup_{T\to\infty} \mathbb{E}[\int_0^T (X_t - \hat{X}_t)^2 dt]/T$, but operationally a nicer criterion is $\limsup_{i\to\infty} \mathbb{E}[\int_0^{D_i} (X_t - \hat{X}_t)^2 dt]/\mathbb{E}[D_i]$. These criteria correspond to two definitions of "average cost per unit time" that are widely used in the literature of semi-Markov decision processes. These two criteria are equivalent, if $\{T_1, T_2, \ldots\}$ is a regenerative process, or more generally, if $\{T_1, T_2, \ldots\}$ has only one ergodic

From this, we can obtain that S_i is finite almost surely for all i. We assume that the Gauss-Markov process $\{X_t, t \geq 0\}$ and the service times $\{Y_i, i = 1, 2, \dots\}$ are mutually independent, and do not change according to the sampling policy.

A sampling policy $\pi \in \Pi$ is said to be *signal-agnostic* (*signal-aware*), if π is (not necessarily) independent of $\{X_t, t \geq 0\}$. Let $\Pi_{\text{signal-agnostic}} \subset \Pi$ denote the set of signal-agnostic sampling policies, defined as

$$\Pi_{\text{signal-agnostic}} = \{\pi \in \Pi : \pi \text{ is independent of } \{X_t, t \geq 0\}\}. \tag{2.7}$$

2.2.3 MMSE Estimator

According to (2.6), S_i is a finite stopping time. By using the expression of OU process for stable scenario [83, Eq. (3)] and the strong Markov property of the OU process [74, Eq. (4.3.27)], a solution to (3.1) for $t \in [S_i, \infty)$ given by the following three cases:

$$X_t = \begin{cases} X_{S_i} e^{-\theta(t-S_i)} + \mu\big[1 - e^{-\theta(t-S_i)}\big] + \frac{\sigma}{\sqrt{2\theta}} e^{-\theta(t-S_i)} W_{e^{2\theta(t-S_i)}-1}, & \text{if } \theta > 0, \\ \sigma W_t, & \text{if } \theta = 0, \\ X_{S_i} e^{-\theta(t-S_i)} + \mu\big[1 - e^{-\theta(t-S_i)}\big] + \frac{\sigma}{\sqrt{-2\theta}} e^{-\theta(t-S_i)} W_{1-e^{2\theta(t-S_i)}}, & \text{if } \theta < 0. \end{cases} \tag{2.8}$$

At any time $t \geq 0$, the estimator uses causally received samples to construct an estimate $\hat{X}_t$ of the real-time signal value X_t. The information available to the estimator consists of two parts: (i) $M_t = \{(S_i, X_{S_i}, D_i) : D_i \leq t\}$, which contains the sampling time S_i, sample value X_{S_i}, and delivery time D_i of the samples that have been delivered by time t and (ii) the fact that no sample has been received after the last delivery time $\max\{D_i : D_i \leq t\}$. Similar to [30, 35, 84], we assume that the estimator neglects the second part of the information.[2] Then, as shown in Appendix 2.C, if $t \in [D_i, D_{i+1}), i = 0, 1, 2, \dots$, the minimum mean square

class. If no condition is imposed, however, they are different. The interested readers are referred to [78–82] for more discussions.

[2] We note that this assumption can be removed by considering a joint sampler and estimator design problem. Specifically, it was shown in [27–29,32,33] that when the sampler and estimator are jointly optimized in discrete-time systems, the optimal estimator has the same expression no matter with or without the second part of information. As pointed out in [28, p. 619], such a structure-property of the MMSE estimator can be also established for continuous-time systems. The goal of this paper is to find the closed-form expression of the optimal sampler under this assumption. The remaining task of finding the jointly optimal sampler and estimator design can be done by further using the majorization techniques developed in [27–29, 32, 33]; see [37] for a recent treatment on this task.

error (MMSE) estimator is determined by

$$\hat{X}_t = \mathbb{E}[X_t|M_t] = \begin{cases} X_{S_i} e^{-\theta(t-S_i)} + \mu\big[1 - e^{-\theta(t-S_i)}\big], & \text{if } \theta \neq 0, \\ \sigma W_{S_i}, & \text{if } \theta = 0. \end{cases} \tag{2.9}$$

Hence, the estimation error of the MMSE estimator for $t \in [D_i, D_{i+1}), i = 0, 1, 2, \ldots$ is

$$X_t - \hat{X}_t = \begin{cases} \frac{\sigma}{\sqrt{2\theta}} e^{-\theta(t-S_i)} W_{e^{2\theta(t-S_i)}-1}, & \text{if } \theta > 0, \\ \sigma(W_t - W_{S_i}), & \text{if } \theta = 0, \\ \frac{\sigma}{\sqrt{-2\theta}} e^{-\theta(t-S_i)} W_{1-e^{2\theta(t-S_i)}}, & \text{if } \theta < 0. \end{cases} \tag{2.10}$$

2.3 Problem Formulation

The goal of this paper is to find the optimal sampling policy that minimizes the mean-squared estimation error subject to an average sampling-rate constraint, which is formulated as the following problem:

$$\mathsf{mse}_{\text{opt}} = \inf_{\pi \in \Pi} \ \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}\left[\int_0^T (X_t - \hat{X}_t)^2 dt\right] \tag{2.11}$$

$$\text{s.t.} \ \liminf_{n \to \infty} \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^{n} (S_{i+1} - S_i)\right] \geq \frac{1}{f_{\max}}, \tag{2.12}$$

where $\mathsf{mse}_{\text{opt}}$ is the optimum value of (2.11) and $f_{\max}$ is the maximum allowed sampling rate. When $f_{\max} = \infty$, this problem becomes an unconstrained problem.

2.4 Signal-aware Sampling

2.4.1 Optimal Sampler without Sampling Rate Constraint

Problem (2.11) is a constrained continuous-time MDP with a continuous state space. However, we found an exact solution to this problem.

To present this solution, let us consider a Gauss-Markov process O_t with the initial state $O_0 = 0$ and parameter $\mu = 0$. According to (4.4), O_t can be expressed as

$$O_t = \begin{cases} \frac{\sigma}{\sqrt{2\theta}} e^{-\theta t} W_{e^{2\theta t}-1}, & \text{if } \theta > 0, \\ \sigma(W_t - W_{S_i}), & \text{if } \theta = 0, \\ \frac{\sigma}{\sqrt{-2\theta}} e^{-\theta t} W_{1-e^{2\theta t}}, & \text{if } \theta < 0. \end{cases} \tag{2.13}$$

Define

$$\mathsf{mse}_{Y_i} = \mathbb{E}[O_{Y_i}^2] = \frac{\sigma^2}{2\theta}\mathbb{E}[1 - e^{-2\theta Y_i}], \tag{2.14}$$

$$\mathsf{mse}_{\infty} = \mathbb{E}[O_{\infty}^2] = \frac{\sigma^2}{2\theta}. \tag{2.15}$$

In the sequel, we will see that mse_{Y_i} and mse_{∞} are the lower and upper bounds of $\mathsf{mse}_{\text{opt}}$, respectively. According to (2.10) and (2.13)-(2.15), mse_{Y_i} represents the estimation error when the estimation is made based on a sample that was generated Y_i seconds ago, and mse_{∞} represents the estimation error for the case that no sample has been delivered to the estimator before. We will also need to use the function[3]

$$G(x) = \frac{e^{x^2}}{x}\int_0^x e^{-t^2} dt = \frac{e^{x^2}}{x}\frac{\sqrt{\pi}}{2}\operatorname{erf}(x), \ x \in [0, \infty), \tag{2.16}$$

where $\operatorname{erf}(\cdot)$ is the error function [85], defined as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2} dt. \tag{2.17}$$

We first consider the unconstrained optimal sampling problem, i.e., $f_{\max} = \infty$, such that the rate constraint (2.12) can be removed. In this scenario, the optimal sampler is provided in the following theorem.

Theorem 2.1 *(Sampling without Rate Constraint). If $f_{\max} = \infty$ and the Y_i's are i.i.d. with $0 < \mathbb{E}[Y_i] < \infty$, then $(S_1(\beta), S_2(\beta), \ldots)$ with a parameter β is an optimal solution to* (2.11),

[3]If $x = 0$, $G(x)$ is defined as its right limit $G(0) = \lim_{x\to 0^+} G(x) = 1$.

where

$$S_{i+1}(\beta) = \inf\left\{t \geq D_i(\beta) : \left|X_t - \hat{X}_t\right| \geq v(\beta)\right\}, \tag{2.18}$$

$D_i(\beta) = S_i(\beta) + Y_i$, $v(\beta)$ *is defined by*

$$v(\beta) = \frac{\sigma}{\sqrt{\theta}} G^{-1}\left(\frac{\mathsf{mse}_\infty - \mathsf{mse}_{Y_i}}{\mathsf{mse}_\infty - \beta}\right), \tag{2.19}$$

$G^{-1}(\cdot)$ *is the inverse function of* $G(\cdot)$ *in* (2.16) *and* β *is the unique root of*

$$\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} (X_t - \hat{X}_t)^2 dt\right] - \beta \mathbb{E}[D_{i+1}(\beta) - D_i(\beta)] = 0. \tag{2.20}$$

The optimal objective value to (2.11) *is given by*

$$\mathsf{mse}_{\text{opt}} = \frac{\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} (X_t - \hat{X}_t)^2 dt\right]}{\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)]}. \tag{2.21}$$

Furthermore, β *is exactly the optimal value to* (2.11)*, i.e.,* $\beta = \mathsf{mse}_{\text{opt}}$.

The proof of Theorem 2.1 is explained in Section 2.6. The optimal sampling policy in Theorem 3.1 has a nice structure. Specifically, the $(i+1)$-th sample is taken at the earliest time t satisfying two conditions: (i) The i-th sample has already been delivered by time t, i.e., $t \geq D_i(\beta)$, and (ii) the estimation error $|X_t - \hat{X}_t|$ is no smaller than a pre-determined threshold $v(\beta)$, where $v(\cdot)$ is a non-linear function defined in (2.19). In Section 2.6, it is shown that $\mathsf{mse}_{Y_i} \leq \beta < \mathsf{mse}_\infty$. Further, it is not hard to show that $G(x)$ is strictly increasing on $[0, \infty)$ and $G(0) = 1$. Hence, its inverse function $G^{-1}(\cdot)$ and the threshold $v(\beta)$ are properly defined and $v(\beta) \geq 0$.

Three Algorithms for Solving (2.20)

We now present three algorithms for computing the root of (2.20). Because the $S_i(\beta)$'s are stopping times, numerically calculating the expectations in (2.20) appears to be a difficult

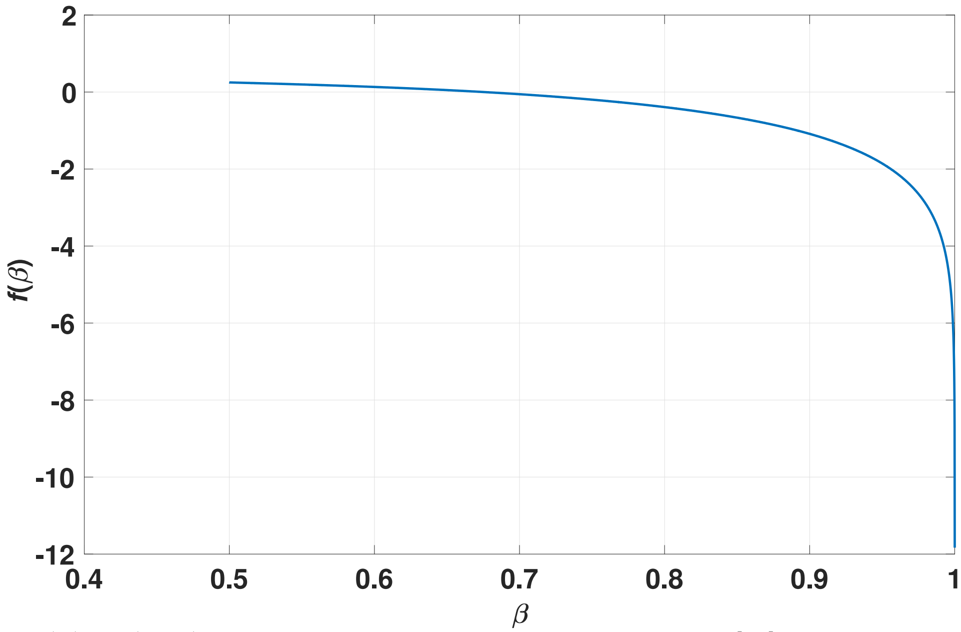


Figure 2.3: $f(\beta)$ in (2.30) for *i.i.d.* exponential service time with $\mathbb{E}[Y_i] = 1$, where the parameters of the Gauss-Markov process are $\sigma = 1$ and $\theta = 0.5$. For these parameters, $\mathsf{mse}_{Y_i} = 0.5$ and $\mathsf{mse}_\infty = 1$.

task. Nonetheless, this challenge can be solved by resorting to the following lemma, which is obtained by using Dynkin's formula [75, Theorem 7.4.1] and the optional stopping theorem.

Lemma 2.1 *In Theorem 2.1, it holds that*

$$\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)] = \mathbb{E}[\max\{R_1(v(\beta)) - R_1(O_{Y_i}), 0\}] + \mathbb{E}[Y_i], \tag{2.22}$$

$$\begin{aligned}\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} (X_t - \hat{X}_t)^2 dt\right] =& \mathbb{E}[\max\{R_2(v(\beta)) - R_2(O_{Y_i}), 0\}] \\ &+ \mathsf{mse}_\infty[\mathbb{E}(Y_i) - \gamma] + \mathbb{E}\left[\max\{v^2(\beta), O_{Y_i}^2\}\right]\gamma, \end{aligned} \tag{2.23}$$

where

$$\gamma = \frac{1}{2\theta}\mathbb{E}[1 - e^{-2\theta Y_i}], \tag{2.24}$$

$$R_1(v) = \frac{v^2}{\sigma^2}\, {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta}{\sigma^2}v^2\right), \tag{2.25}$$

$$R_2(v) = -\frac{v^2}{2\theta} + \frac{v^2}{2\theta}\, {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta}{\sigma^2}v^2\right). \tag{2.26}$$

Proof 2.1 *See Appendix 2.I.*

In (2.25) and (2.26), we have used the generalized hypergeometric function, which is defined by [86, Eq. 16.2.1]

$$ {}_pF_q(a_1, a_2, \cdots, a_p; b_1, b_2, \cdots b_q; z) = \sum_{n=0}^{\infty} \frac{(a_1)_n (a_2)_n \cdots (a_p)_n}{(b_1)_n (b_2)_n \cdots (b_p)_n} \frac{z^n}{n!}, \tag{2.27} $$

where

$$ (a)_0 = 1, \tag{2.28} $$

$$ (a)_n = a(a+1)(a+2)\cdots(a+n-1), \ n \geq 1. \tag{2.29} $$

Using Lemma 1, the expectations in (2.20) can be evaluated by Monte Carlo simulations of scalar random variables O_{Y_i} and Y_i, which is much simpler than directly simulating the entire random process $\{O_t, t \geq 0\}$.

For notational simplicity, we rewrite (2.20) as

$$ f(\beta) = f_1(\beta) - \beta f_2(\beta) = 0, \tag{2.30} $$

where $f_1(\beta) = \mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} (X_t - \hat{X}_t)^2 dt\right]$ and $f_2(\beta) = \mathbb{E}[D_{i+1}(\beta) - D_i(\beta)]$. The function $f(\beta)$ has several nice properties, which are asserted in the following lemma and illustrated in Figure 2.3.

Lemma 2.2 *The function $f(\beta)$ has the following properties:*

(i) $f(\beta)$ is concave, continuous, and strictly decreasing in β,

(ii) $f(\mathsf{mse}_{Y_i}) > 0$ and $\lim_{\beta \to \mathsf{mse}_{\infty}^-} f(\beta) = -\infty$.

Proof 2.2 *See Appendix 2.A.*

The uniqueness of the root of $f(\beta)$ follows immediately from Lemma 2.2.

Because $f(\beta)$ is decreasing and has a unique root, one can use a bisection search method to solve (2.20), which is illustrated in Algorithm 1. The bisection search method has a globally linear convergence speed.

Algorithm 1 Bisection search method for solving (2.20)

given $l = \mathsf{mse}_{Y_i}$, $u = \mathsf{mse}_{\infty}$, tolerance $\epsilon > 0$.
repeat
  $\beta := (l + u)/2$.
  $o := f_1(\beta) - \beta f_2(\beta)$.
  if $o \geq 0$, $l := \beta$; **else**, $u := \beta$.
until $u - l \leq \epsilon$.
return β.

Algorithm 2 Newton's method for solving (2.20)

given tolerance $\epsilon > 0$.
Pick initial value $\beta_0 \in [\mathsf{mse}_{\text{opt}}, \mathsf{mse}_{\infty})$.
repeat
  $\beta_{k+1} := \beta_k - \frac{f(\beta_k)}{f'(\beta_k)}$.
until $|\frac{f(\beta_k)}{f'(\beta_k)}| \leq \epsilon$.
return β_{k+1}.

Algorithm 3 Fixed-point iterations for solving (2.20)

given tolerance $\epsilon > 0$.
Pick initial value $\beta_0 \in [\mathsf{mse}_{\text{opt}}, \mathsf{mse}_{\infty})$.
repeat
  $\beta_{k+1} := \frac{f_1(\beta_k)}{f_2(\beta_k)}$.
until $|\beta_{k+1} - \frac{f_1(\beta_k)}{f_2(\beta_k)}| \leq \epsilon$.
return β_{k+1}.

To achieve an even faster convergence speed, we can use Newton's method [87]

$$\beta_{k+1} = \beta_k - \frac{f(\beta_k)}{f'(\beta_k)} \tag{2.31}$$

to solve (2.20), as shown in Algorithm 2. We suggest choosing the initial value β_0 of Newton's method from the set $[\mathsf{mse}_{\text{opt}}, \mathsf{mse}_{\infty})$, i.e., β_0 is larger than the root $\mathsf{mse}_{\text{opt}}$. Such an initial value β_0 can be found by taking a few bisection search iterations, or by using the mse of a sub-optimal sampling policy [88]. Because $f(\beta)$ is a concave function, the choice of initial value $\beta_0 \in [\mathsf{mse}_{\text{opt}}, \mathsf{mse}_{\infty})$ ensures that β_k is a decreasing sequence converging to $\mathsf{mse}_{\text{opt}}$ [89]. Moreover, because $R_1(\cdot)$ and $R_2(\cdot)$ are twice continuously differentiable, the function $f(\beta)$ is twice continuously differentiable. Therefore, Newton's method is known to have a locally quadratic convergence speed in the neighborhood of the root $\mathsf{mse}_{\text{opt}}$ [87, Chapter 2].

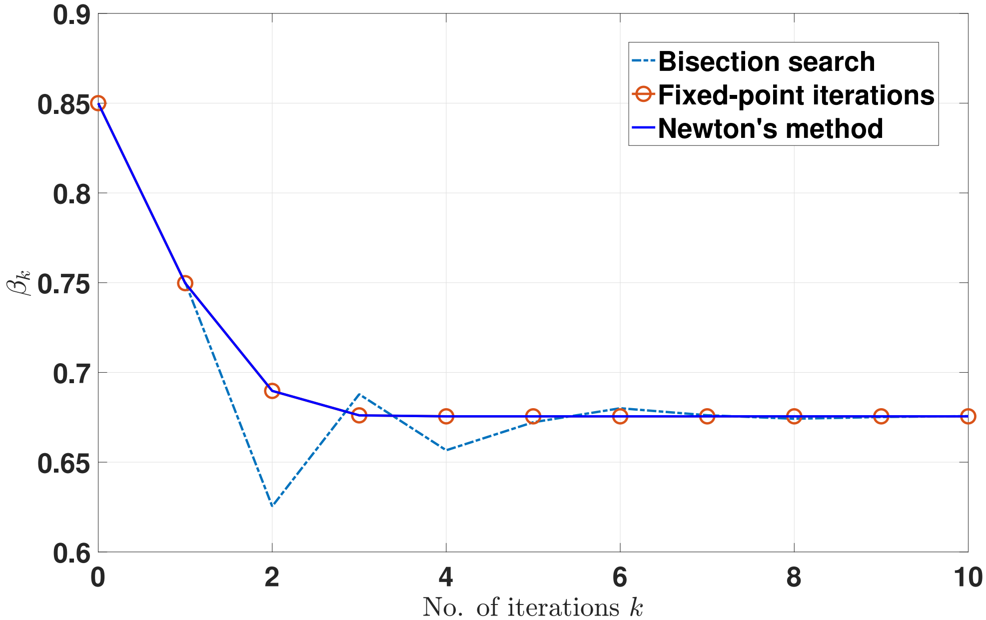


Figure 2.4: Convergence of three algorithms for solving (4.56), where the service times are *i.i.d.* exponential with mean $\mathbb{E}[Y_i] = 1$, the parameters of the Gauss-Markov process are $\sigma = 1$ and $\theta = 0.5$.

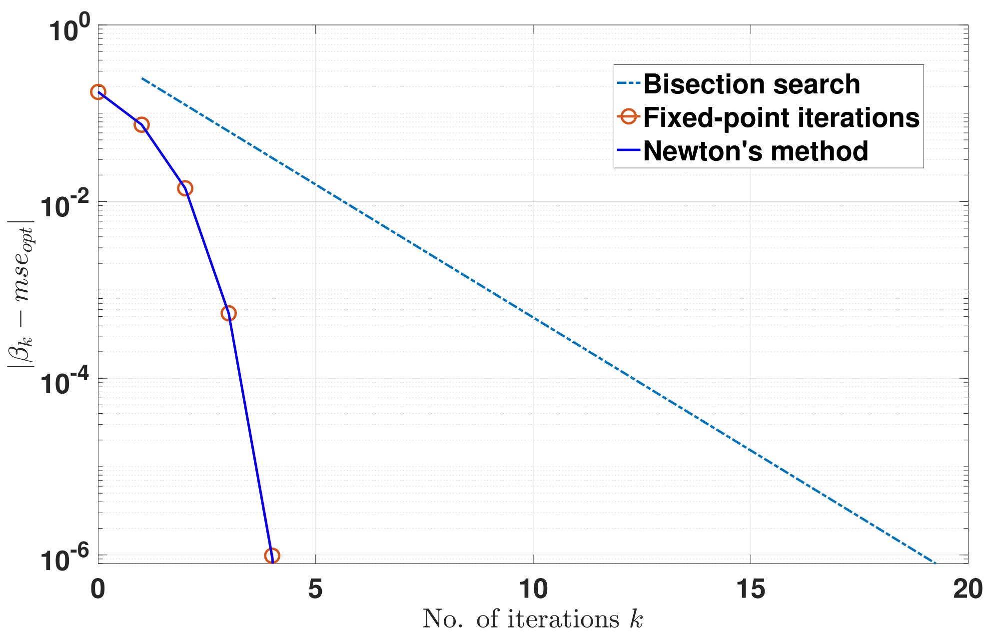


Figure 2.5: Convergence of three algorithms for solving (4.56), where the service times are *i.i.d.* exponential with mean $\mathbb{E}[Y_i] = 1$, the parameters of the Gauss-Markov process are $\sigma = 1$ and $\theta = 0.5$. For bisection search, we plot the difference $|u - l|$ between the upper bound u and lower bound l, which is an upper bound of $|\beta_k - \mathsf{mse}_{\text{opt}}|$.

Newton's method requires to compute the gradient $f'(\beta_k)$, which can be solved by a finite-difference approximation, as in the secant method [87]. In the sequel, we introduce another approximation approach of Newton's method, which is of independent interest. In Theorem 2.1, we have shown that

$$\mathsf{mse}_{\rm opt} = \underset{\beta\in[\mathsf{mse}_{Y_i},\mathsf{mse}_\infty)}{\rm argmax} \frac{f_1(\beta)}{f_2(\beta)}. \tag{2.32}$$

Hence, the gradient of $f_1(\beta)/f_2(\beta)$ is equal to zero at the optimal solution $\beta = \mathsf{mse}_{\rm opt}$, which leads to

$$f_1'(\mathsf{mse}_{\rm opt})f_2(\mathsf{mse}_{\rm opt}) - f_1(\mathsf{mse}_{\rm opt})f_2'(\mathsf{mse}_{\rm opt}) = 0. \tag{2.33}$$

Therefore,

$$\mathsf{mse}_{\rm opt} = \frac{f_1(\mathsf{mse}_{\rm opt})}{f_2(\mathsf{mse}_{\rm opt})} = \frac{f_1'(\mathsf{mse}_{\rm opt})}{f_2'(\mathsf{mse}_{\rm opt})}. \tag{2.34}$$

Because $f_1(\beta)$ and $f_2(\beta)$ are smooth functions, when β_k is in the neighborhood of $\mathsf{mse}_{\rm opt}$, (2.34) implies that $f_1'(\beta_k) - \beta_k f_2'(\beta_k) \approx f_1'(\mathsf{mse}_{\rm opt}) - \mathsf{mse}_{\rm opt} f_2'(\mathsf{mse}_{\rm opt}) = 0$. Substituting this into (2.31), yields

$$\begin{aligned}\beta_{k+1} =& \beta_k - \frac{f_1(\beta_k) - \beta_k f_2(\beta_k)}{f_1'(\beta_k) - f_2(\beta_k) - \beta_k f_2'(\beta_k)}\\ \approx& \beta_k - \frac{f_1(\beta_k) - \beta_k f_2(\beta_k)}{-f_2(\beta_k)}\\ =& \frac{f_1(\beta_k)}{f_2(\beta_k)},\end{aligned} \tag{2.35}$$

which is a fixed-point iterative algorithm (see Algorithm 3) that was recently proposed in [88]. Similar to Newton's method, the fixed-point updates in (2.35) converge to $\mathsf{mse}_{\rm opt}$ if the initial value $\beta_0\in[\mathsf{mse}_{\rm opt}, \mathsf{mse}_\infty)$. Moreover, (2.35) has a locally quadratic convergence speed, see [88] for a proof of this result. A numerical comparison of these three algorithms is shown in Figure 2.4 and Figure 2.5. One can observe that the fixed-point updates and Newton's method converge faster than the bisection search.

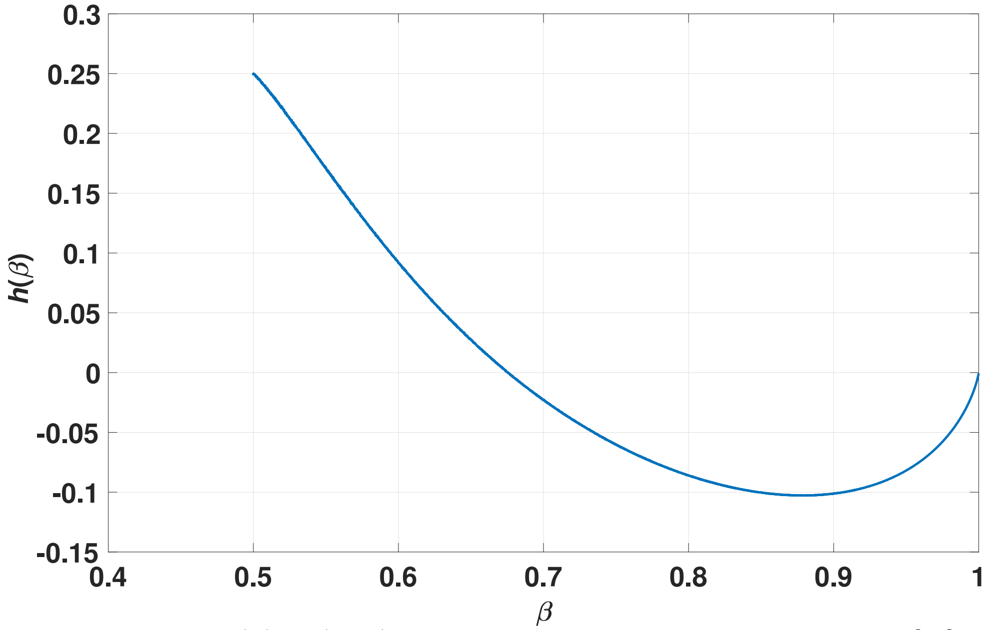


Figure 2.6: The function $h(\beta)$ in (2.36) for *i.i.d.* exponential service time with $\mathbb{E}[Y_i] = 1$, where the parameters of the Gauss-Markov process are $\sigma = 1$ and $\theta = 0.5$. For these parameters, $\mathsf{mse}_{Y_i} = 0.5$ and $\mathsf{mse}_{\infty} = 1$.

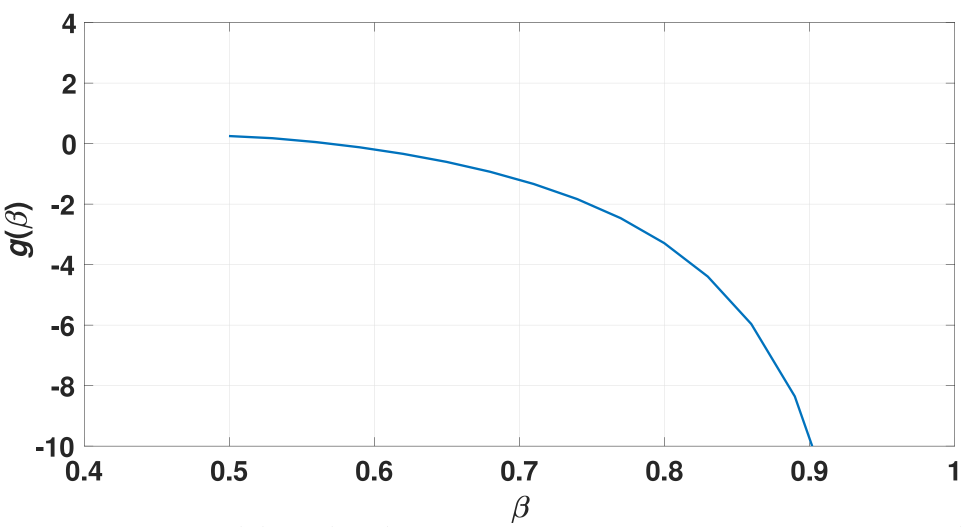


Figure 2.7: The function $g(\beta)$ in (2.40) for *i.i.d.* exponential service time with $\mathbb{E}[Y_i] = 1$ and $f_{\max} = 0.8$, where the parameters of the Gauss-Markov process are $\sigma = 1$ and $\theta = 0.5$. For these parameters, $\mathsf{mse}_{Y_i} = 0.5$ and $\mathsf{mse}_{\infty} = 1$.

We note that although (2.20), and equivalently (2.30), has a unique root $\mathsf{mse}_{\text{opt}}$, the fixed-point equation

$$h(\beta) = \frac{f_1(\beta)}{f_2(\beta)} - \beta = \frac{f_1(\beta) - \beta f_2(\beta)}{f_2(\beta)} = 0 \tag{2.36}$$

has two roots $\mathsf{mse}_{\text{opt}}$ and mse_{∞}. See Figure 2.6 for an illustration of the two roots of $h(\beta)$. As shown in Appendix 2.O, the correct root for computing the optimal threshold is $\mathsf{mse}_{\text{opt}}$. Interestingly, Algorithms 1-3 converge to the desired root $\mathsf{mse}_{\text{opt}}$, instead of mse_{∞}. Finally, we remark that these three algorithms can be used to find the optimal threshold in the age-optimal sampling problem studied in, e.g., [4, 5].

2.4.2 Optimal Sampler with Sampling Rate Constraint

When the sampling rate constraint (2.12) is taken into consideration, a solution to (2.11) is expressed in the following theorem:

Theorem 2.2 *(Sampling with Rate Constraint). If the Y_i's are i.i.d. with $0 < \mathbb{E}[Y_i] < \infty$, then (2.18)-(2.20) is an optimal solution to (2.11). The value of $\beta \geq 0$ is determined in two cases: β is the unique root of (2.20) if the root of (2.20) satisfies*

$$\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)] > 1/f_{\max}; \tag{2.37}$$

otherwise, β is the unique root of

$$\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)] = 1/f_{\max}. \tag{2.38}$$

The optimal objective value to (2.11) is given by

$$\mathsf{mse}_{\text{opt}} = \frac{\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} (X_t - \hat{X}_t)^2 dt\right]}{\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)]}. \tag{2.39}$$

The proof of Theorem 2.2 is explained in Section 2.6. One can see that Theorem 2.1 is a special case of Theorem 2.2 when $f_{\max} = \infty$.

Algorithm 4 Bisection search method for solving (2.38)

given $l = \mathsf{mse}_{Y_i}$, $u = \mathsf{mse}_\infty$, tolerance $\epsilon > 0$.
repeat
  $\beta := (l+u)/2$.
  $o := \mathbb{E}[D_{i+1}(\beta) - D_i(\beta)]$.
  if $o \geq 1/f_{\max}$, $u := \beta$; **else**, $l := \beta$.
until $u - l \leq \epsilon$.
return β.

In Theorem 2.2, the calculation of β falls into two cases: In one case, β can be computed by solving (2.20) via Algorithms 1-3. For this case to occur, the sampling rate constraint (2.12) needs to be inactive at the root of (2.20). Because $D_i(\beta) = S_i(\beta) + Y_i$, we can obtain $\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)] = \mathbb{E}[S_{i+1}(\beta) - S_i(\beta)]$ and hence (2.37) holds when the sampling rate constraint (2.12) is inactive.

In the other case, β is selected to satisfy the sampling rate constraint (2.12) with equality, as required in (2.38). Before we solve (2.38), let us first use $f_2(\beta)$ to express (2.38) as

$$g(\beta) = \frac{1}{f_{\max}} - f_2(\beta) = 0. \tag{2.40}$$

Lemma 2.3 *The function $g(\beta)$ has the following properties:*

(i) $g(\beta)$ is continuous and strictly decreasing in β,

(ii) $g(\mathsf{mse}_{Y_i}) \geq 0$ and $\lim_{\beta \to \mathsf{mse}_\infty^-} g(\beta) = -\infty$ if the root of (4.56) *does not satisfy* (2.37).

Proof 2.3 *See Appendix 2.B.*

According to Lemma 2.3, (2.38) has a unique root in $[\mathsf{mse}_{Y_i}, \mathsf{mse}_\infty)$, which is denoted as β^*. In addition, the numerical results in Figure 2.7 suggest that $g(\beta)$ should be concave, for which we do not have proof.

The root β^* can be solved by using bisection search and Newton's method, which are explained in Algorithms 4-5, respectively. Similar to the discussions in Section 2.4.1, the convergence of Algorithm 4 is ensured by Lemma 2.3. Moreover, if $g(\beta)$ is concave and $\beta_0 \in [\beta^*, \mathsf{mse}_\infty)$, β_k in Algorithm 5 is a decreasing sequence converging to the root β^* of (2.38) [89].

Algorithm 5 Newton's method for solving (2.38)

given tolerance $\epsilon > 0$.
Pick initial value $\beta_0 \in [\beta^*, \mathsf{mse}_\infty)$.
repeat
    $\beta_{k+1} := \beta_k - \frac{g(\beta_k)}{g'(\beta_k)}$.
until $|\frac{g(\beta_k)}{g'(\beta_k)}| \leq \epsilon$.
return β_{k+1}.

2.4.3 Special Case: Sampling of the Wiener Process

In the limiting case that $\sigma = 1$ and $\theta \to 0$, the Gauss-Markov process X_t in (3.1) becomes a Wiener process $X_t = W_t$. In this case, the MMSE estimator in (2.9) is given by

$$\hat{X}_t = W_{S_i}, \text{ if } t \in [D_i, D_{i+1}). \tag{2.41}$$

As shown in Appendix 2.E, $v(\cdot)$ defined by (2.19) tends to

$$v(\beta) = \sqrt{3(\beta - \mathbb{E}[Y_i])}. \tag{2.42}$$

Theorem 2.3 *If $\sigma = 1$, $\theta \to 0$, and the Y_i's are i.i.d. with $0 < \mathbb{E}[Y_i] < \infty$, then* $(S_1(\beta), S_2(\beta),$
$\ldots)$ *with a parameter β is an optimal solution to* (2.11)*, where*

$$S_{i+1}(\beta) = \inf\left\{t \geq D_i(\beta) : \left|X_t - \hat{X}_t\right| \geq \sqrt{3(\beta - \mathbb{E}[Y_i])}\right\}, \tag{2.43}$$

$D_i(\beta) = S_i(\beta) + Y_i$. *The value of $\beta \geq 0$ is determined in two cases: β is the unique root of*

$$\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} (X_t - \hat{X}_t)^2 dt\right] - \beta\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)] = 0, \tag{2.44}$$

if the root of (2.44) *satisfies $\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)] > 1/f_{\max}$; otherwise, β is the unique root of $\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)] = 1/f_{\max}$. The optimal objective value to* (2.11) *is given by*

$$\mathsf{mse}_{\text{opt}} = \frac{\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} (X_t - \hat{X}_t)^2 dt\right]}{\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)]}. \tag{2.45}$$

Theorem 2.3 is an alternative form of Theorem 1 in [35] and hence its proof is omitted. The benefit of the new expression in Theorem 2.3 is that it allows to character β based on the optimal objective value $\mathsf{mse}_{\text{opt}}$ and the sampling rate constraint (2.12), in the same way as in Theorems 2.1-2.2. This appears to be more fundamental than the expression in [35]. The new form of optimal sampling policy of Wiener processes was also discovered in [36] without considering the constraint on (2.12).

2.5 Signal-agnostic Sampling

In signal-agnostic sampling policies, the sampling times S_i are determined based only on the service times Y_i, but not on the observed Gauss-Markov process $\{X_t, t \geq 0\}$.

Lemma 2.4 *If $\pi \in \Pi_{signal\text{-}agnostic}$, then the mean-squared estimation error of the Gauss-Markov process X_t at time t is*

$$p(\Delta(t)) = \mathbb{E}\left[(X_t - \hat{X}_t)^2 \middle| \pi, Y_1, Y_2, \ldots\right] = \frac{\sigma^2}{2\theta}\left(1 - e^{-2\theta\Delta(t)}\right), \tag{2.46}$$

which is a strictly increasing function of the age $\Delta(t)$.

Proof 2.4 *See Appendix 2.D.*

According to Lemma 2.4, for every policy $\pi \in \Pi_{\text{signal-agnostic}}$,

$$\mathbb{E}\left[\int_0^T (X_t - \hat{X}_t)^2 dt\right] = \mathbb{E}\left[\int_0^T p(\Delta(t)) dt\right]. \tag{2.47}$$

Hence, minimizing the mean-squared estimation error among signal-agnostic sampling policies can be formulated as the following MDP for minimizing the expected time-average of the nonlinear age function $p(\Delta(t))$ in (2.46):

$$\mathsf{mse}_{\text{age-opt}} = \inf_{\pi \in \Pi_{\text{signal-agnostic}}} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T p(\Delta(t)) dt\right] \tag{2.48}$$

$$\text{s.t.} \quad \liminf_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^{n}(S_{i+1} - S_i)\right] \geq \frac{1}{f_{\max}}, \tag{2.49}$$

where $\mathsf{mse}_{\text{age-opt}}$ is the optimal value of (2.48). By (2.46), $p(\Delta(t))$ and $\mathsf{mse}_{\text{age-opt}}$ are bounded. Because $\Pi_{\text{signal-agnostic}} \subset \Pi$, it follows immediately that $\mathsf{mse}_{\text{opt}} \leq \mathsf{mse}_{\text{age-opt}}$.

Problem (2.48) is one instance of the problems recently solved in Corollary 3 of [4] for general strictly increasing functions $p(\cdot)$. From this, a solution to (2.48) for signal-agnostic sampling is given by

Theorem 2.4 *If the Y_i's are i.i.d. with $0 < \mathbb{E}[Y_i] < \infty$, then $(S_1(\beta), S_2(\beta), \ldots)$ with a parameter β is an optimal solution to* (2.48)*, where*

$$S_{i+1}(\beta) = \inf \left\{ t \geq D_i(\beta) : \mathbb{E}[(X_{t+Y_{i+1}} - \hat{X}_{t+Y_{i+1}})^2] \geq \beta \right\}, \tag{2.50}$$

$D_i(\beta) = S_i(\beta) + Y_i$ and β is the unique root of

$$\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} (X_t - \hat{X}_t)^2 dt\right] - \beta \mathbb{E}[D_{i+1}(\beta) - D_i(\beta)] = 0, \tag{2.51}$$

if the root of (2.51) *satisfies $\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)] > 1/f_{\max}$; otherwise, β is the unique root of*

$$\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)] = 1/f_{\max}. \tag{2.52}$$

The optimal objective value to (2.48) *is given by*

$$\mathsf{mse}_{\text{age-opt}} = \frac{\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} (X_t - \hat{X}_t)^2 dt\right]}{\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)]}. \tag{2.53}$$

Theorem 2.4 follows from Corollary 3 of [4] and Lemma 2.4. Similar to the case of signal-aware sampling, the roots of (2.51) and (2.52) can be solved by using Algorithms 1-5. In fact, Algorithms 1-5 can be used for minimizing general non-decreasing age penalty [4].

2.5.1 Discussions and Remarks

The difference among Theorems 2.1-2.4 is only in the expressions (2.18), (2.43), (2.50) of threshold policies. In signal-aware sampling policies (2.18) and (2.43), the sampling time is determined by the *instantaneous* estimation error $|X_t - \hat{X}_t|$, and the threshold function

$v(\cdot)$ is determined by the specific signal model. In the signal-agnostic sampling policy (2.50), the sampling time is determined by the *expected* estimation error $\mathbb{E}[(X_{t+Y_{i+1}} - \hat{X}_{t+Y_{i+1}})^2]$ at time $t + Y_{i+1}$. We note that if $t = S_{i+1}(\beta)$, then $t + Y_{i+1} = S_{i+1}(\beta) + Y_{i+1} = D_{i+1}(\beta)$ is the delivery time of the new sample. Hence, (2.50) requires that the expected estimation error upon the delivery of the new sample is no less than β. The parameter β in Theorems 2.1-2.4 is determined by the optimal objective value and the sampling rate constraint in the same manner. Later on in (2.67), we will further see that β is exactly equal to the summation of the optimal objective value of the MDP and the optimal Lagrangian dual variable associated with the sampling rate constraint. Finally, it is worth noting that Theorems 2.1-2.4 hold for all distributions of the service times Y_i satisfying $0 < \mathbb{E}[Y_i] < \infty$, and for both constrained and unconstrained sampling problems.

2.6 Proof of the Main Results

We first provide the proof of Theorem 2.2. After that Theorem 2.1 follows immediately because it is a special case of Theorem 2.2. We prove Theorem 2.2 in four steps: (i) We first show that sampling should be suspended when the server is busy, which can be used to simplify (2.11). (ii) We use an extended Dinkelbach's method [90] and Lagrangian duality method to decompose the simplified problem into a series of mutually independent per-sample MDP. (iii) We utilize the free boundary method from optimal stopping theory [74] to solve the per-sample MDPs analytically. (iv) Finally, we use a geometric multiplier method [91] to show that the duality gap is zero. The above proof framework is an extension to that used in [4, 35], and the most challenging part is Step (iii).

2.6.1 Preliminaries

The Gauss-Markov process O_t in (2.13) with initial state $O_t = 0$ and parameter $\mu = 0$ is the solution to the SDE

$$dO_t = -\theta O_t dt + \sigma dW_t. \tag{2.54}$$

In addition, the infinitesimal generator of O_t is [92, Eq. A1.22]

$$\mathcal{G} = -\theta u \frac{\partial}{\partial u} + \frac{\sigma^2}{2} \frac{\partial^2}{\partial u^2}. \tag{2.55}$$

According to (2.18) and (2.9), the estimation error $(X_t - \hat{X}_t)$ is of the same distribution with O_{t-S_i}, if $t \in [D_i, D_{i+1})$. By using Dynkin's formula and the optional stopping theorem, we obtain the following lemma.

Lemma 2.5 *Let $\tau \geq 0$ be a stopping time of the Gauss-Markov process O_t with $\mathbb{E}[\tau] < \infty$, then*

$$\mathbb{E}\left[\int_0^\tau O_t^2 dt\right] = \mathbb{E}\left[\frac{\sigma^2}{2\theta}\tau - \frac{1}{2\theta}O_\tau^2\right]. \tag{2.56}$$

If, in addition, τ is the first exit time of a bounded set, then

$$\mathbb{E}[\tau] = \mathbb{E}[R_1(O_\tau)], \tag{2.57}$$

$$\mathbb{E}\left[\int_0^\tau O_t^2 dt\right] = \mathbb{E}[R_2(O_\tau)], \tag{2.58}$$

where $R_1(\cdot)$ and $R_2(\cdot)$ are defined in (2.25) *and* (2.26)*, respectively.*

Proof 2.5 *See Appendix 2.F.*

2.6.2 Suspend Sampling when the Server is Busy

By using the strong Markov property of the Gauss-Markov process X_t and the orthogonality principle of MMSE estimation, we obtain the following useful lemma:

Lemma 2.6 *Suppose that a feasible sampling policy for problem* (2.11) *is π, in which at least one sample is taken when the server is busy processing an earlier generated sample. Then, there exists another feasible policy π' for problem* (2.11) *which has a smaller estimation error than policy π. Therefore, in* (2.11)*, it is suboptimal to take a new sample before the previous sample is delivered.*

Proof 2.6 *See Appendix 2.G.*

A similar result was obtained in [35] for the sampling of Wiener processes.

Because $\{X_t - \hat{X}_t, t \in [D_i, D_{i+1})\}$ and $\{O_{t-S_i}, t \in [D_i, D_{i+1})\}$ are of the same distribution, for each $i = 1, 2, \ldots$,

$$\mathbb{E}\left[\int_{D_i}^{D_{i+1}} (X_t - \hat{X}_t)^2 dt\right] = \mathbb{E}\left[\int_{D_i}^{D_{i+1}} O_{t-S_i}^2 dt\right] = \mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i+Y_{i+1}} O_s^2 ds\right]. \tag{2.59}$$

Because T_i is a regenerative process, the renewal theory [93] tells us that $\frac{1}{n}\mathbb{E}[S_n]$ is a convergent sequence and

$$\begin{aligned}\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T (X_t - \hat{X}_t)^2 dt\right] &= \lim_{n\to\infty} \frac{\mathbb{E}\left[\int_0^{D_n} (X_t - \hat{X}_t)^2 dt\right]}{\mathbb{E}[D_n]} \\ &= \lim_{n\to\infty} \frac{\sum_{i=1}^n \mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i+Y_{i+1}} O_s^2 ds\right]}{\sum_{i=1}^n \mathbb{E}\left[Y_i + Z_i\right]}.\end{aligned} \tag{2.60}$$

Hence, (2.11) can be rewritten as the following MDP:

$$\begin{aligned}\mathsf{mse}_{\text{opt}} = \inf_{\pi\in\Pi_1} \lim_{n\to\infty} &\frac{\sum_{i=1}^n \mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i+Y_{i+1}} O_s^2 ds\right]}{\sum_{i=1}^n \mathbb{E}\left[Y_i + Z_i\right]} \\ \text{s.t.} \ \lim_{n\to\infty} &\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[Y_i + Z_i\right] \geq \frac{1}{f_{\max}},\end{aligned} \tag{2.61}$$

where $\mathsf{mse}_{\text{opt}}$ is the optimal value of (2.61).

2.6.3 Reformulation of Problem (2.61)

In order to solve (2.61), let us consider the following MDP with a parameter $c \geq 0$:

$$\begin{aligned}h(c) = \inf_{\pi\in\Pi_1} \lim_{n\to\infty} &\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i+Y_{i+1}} O_s^2 ds - c(Y_i + Z_i)\right] \\ \text{s.t.} \ \lim_{n\to\infty} &\frac{1}{n}\sum_{i=1}^n \mathbb{E}\left[Y_i + Z_i\right] \geq \frac{1}{f_{\max}},\end{aligned} \tag{2.62}$$

where $h(c)$ is the optimum value of (2.62). Similar to Dinkelbach's method [90] for nonlinear fractional programming, the following lemma holds for the MDP (2.61):

Lemma 2.7 *[35] The following assertions are true:*

(a). $\mathsf{mse}_{\text{opt}} \gtreqless c$ *if and only if* $h(c) \gtreqless 0$.

(b). If $h(c) = 0$*, the solutions to* (2.61) *and* (2.62) *are identical.*

Hence, the solution to (2.61) can be obtained by solving (2.62) and seeking $c = \mathsf{mse}_{\text{opt}} \geq 0$ such that

$$h(\mathsf{mse}_{\text{opt}}) = 0. \tag{2.63}$$

2.6.4 Lagrangian Dual Problem of (2.62)

Next, we use the Lagrangian dual approach to solve (2.62) with $c = \mathsf{mse}_{\text{opt}}$. We define the Lagrangian associated with (2.62) as

$$L(\pi; \lambda) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i+Y_{i+1}} O_s^2 ds - (\mathsf{mse}_{\text{opt}} + \lambda)(Y_i + Z_i) \right] + \frac{\lambda}{f_{\max}}, \tag{2.64}$$

where $\lambda \geq 0$ is the dual variable. Let

$$e(\lambda) = \inf_{\pi\in\Pi_1} L(\pi; \lambda). \tag{2.65}$$

Then, the dual problem of (2.62) is defined by

$$d = \max_{\lambda\geq 0} e(\lambda), \tag{2.66}$$

where d is the optimum value of (2.66). Weak duality [91] implies $d \leq h(\mathsf{mse}_{\text{opt}})$. In Section 2.6.6, we will establish strong duality, i.e., $d = h(\mathsf{mse}_{\text{opt}})$.

In the sequel, we decompose (2.65) into a sequence of mutually independent per-sample MDPs. Let us define

$$\beta = \mathsf{mse}_{\text{opt}} + \lambda. \tag{2.67}$$

As shown in Appendix 2.H, by using Lemma 2.5, we can obtain

$$\mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i+Y_{i+1}} O_s^2 ds - \beta(Y_i+Z_i)\right] = \mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i} (O_s^2 - \beta)ds + \gamma O_{Y_i+Z_i}^2\right] + \frac{\sigma^2}{2\theta}[\mathbb{E}(Y_{i+1}) - \gamma] - \beta\mathbb{E}[Y_{i+1}], \tag{2.68}$$

where γ is defined in (2.24). For any $s \geq 0$, define the σ-fields $\mathcal{F}_t^s = \sigma(O_{s+r} - O_s : r \in [0,t])$ and the right-continuous filtration $\mathcal{F}_t^{s+} = \cap_{r>t}\mathcal{F}_r^s$. Then, $\{\mathcal{F}_t^{s+}, t \geq 0\}$ is the filtration of the time-shifted Gauss-Markov process $\{O_{s+t} - O_s, t \in [0,\infty)\}$. Define $\mathfrak{M}_s$ as the set of integrable stopping times of $\{O_{s+t} - O_s, t \in [0,\infty)\}$, i.e.,

$$\mathfrak{M}_s = \{\tau \geq 0 : \{\tau \leq t\} \in \mathcal{F}_t^{s+}, \mathbb{E}[\tau] < \infty\}. \tag{2.69}$$

By using a sufficient statistic of (2.65), we can obtain

Lemma 2.8 *An optimal solution* $(Z_0, Z_1, \ldots)$ *to* (2.65) *satisfies*

$$\inf_{Z_i \in \mathfrak{M}_{Y_i}} \mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i} (O_s^2 - \beta)ds + \gamma O_{Y_i+Z_i}^2 \middle| O_{Y_i}, Y_i\right], \tag{2.70}$$

where $\beta \geq 0$ *and* $\gamma \geq 0$ *are defined in* (2.67) *and* (2.24), *respectively.*

Proof 2.7 *See Appendix 2.J.*

By this, (2.65) is decomposed as a series of per-sample MDP (2.70).

2.6.5 Analytical Solution to Per-sample MDP (2.70)

We solve (2.70) by using the free-boundary approach for optimal stopping problems [74]. Let us consider a Gauss-Markov process V_t with initial state $V_0 = v$ and parameter $\mu = 0$. Define the σ-fields $\mathcal{F}_t^V = \sigma(V_s : s \in [0,t])$, $\mathcal{F}_t^{V+} = \cap_{r>t}\mathcal{F}_r^V$, and the filtration $\{\mathcal{F}_t^{V+}, t \geq 0\}$ associated to $\{V_t, t \geq 0\}$. Define $\mathfrak{M}_V$ as the set of integrable stopping times of $\{V_t, t \in [0,\infty)\}$, i.e.,

$$\mathfrak{M}_V = \{\tau \geq 0 : \{\tau \leq t\} \in \mathcal{F}_t^{V+}, \mathbb{E}[\tau] < \infty\}. \tag{2.71}$$

Our goal is to solve the following optimal stopping problem for any given initial state $v \in \mathbb{R}$

$$\sup_{\tau \in \mathfrak{M}_V} \mathbb{E}_v \left[-\gamma V_\tau^2 - \int_0^\tau (V_s^2 - \beta) ds \right], \tag{2.72}$$

where $\mathbb{E}_v[\cdot]$ is the conditional expectation for given initial state $V_0 = v$, γ and β are given by (2.24) and (2.67), respectively. Hence, (2.70) is one instance of (2.72) with $v = O_{Y_i}$, where the supremum is taken over all stopping times τ of V_t. In this subsection, we focus on the case that β in (2.72) satisfies $\mathsf{mse}_{Y_i} \leq \beta < \mathsf{mse}_\infty$. Later on in Section 2.6.6, we will show that this condition is indeed satisfied by the optimal solution to (2.62).

To solve (2.72), we first find a candidate solution to (2.72) by solving a free boundary problem; then we prove that the free boundary solution is indeed the value function of (4.128):

A Candidate Solution to (2.72)

Now, we show how to solve (2.72). The general optimal stopping theory in Chapter I of [74] tells us that the following guess of the stopping time should be optimal for Problem (2.72):

$$\tau_* = \inf\{t \geq 0 : |V_t| \geq v_*\}, \tag{2.73}$$

where $v_* \geq 0$ is the optimal stopping threshold to be found. Observe that in this guess, the continuation region $(-v_*, v_*)$ is assumed symmetric around zero. This is because the Gauss-Markov process is symmetric, i.e., the process $\{-V_t, t \geq 0\}$ is also a Gauss-Markov process started at $-V_0 = -v$. Similarly, we can also argue that the value function of problem (2.72) should be even.

According to [74, Chapter 8], and [75, Chapter 10], the value function and the optimal stopping threshold v_* should satisfy the following free boundary problem:

$$\frac{\sigma^2}{2}H''(v) - \theta v H'(v) = v^2 - \beta, \quad v \in (-v_*, v_*), \tag{2.74}$$

$$H(\pm v_*) = -\gamma v_*^2, \tag{2.75}$$

$$H'(\pm v_*) = \mp 2\gamma v_*. \tag{2.76}$$

In Appendix 2.K, we use the integrating factor method [94, Sec. I.5] to find the general solution to (2.74), which is given by

$$H(v) = \\ -\frac{v^2}{2\theta} + \left(\frac{1}{2\theta} - \frac{\beta}{\sigma^2}\right) {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta}{\sigma^2}v^2\right) v^2 + C_1 \text{erfi}\left(\frac{\sqrt{\theta}}{\sigma}v\right) + C_2, \ v \in (-v_*, v_*), \tag{2.77}$$

where C_1 and C_2 are constants to be found for satisfying (2.75)-(2.76), and $\text{erfi}(x)$ is the imaginary error function, i.e.,

$$\text{erfi}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{t^2} dt. \tag{2.78}$$

Because $H(v)$ should be even but $\text{erfi}(x)$ is odd, we should choose $C_1 = 0$. Further, in order to satisfy the boundary condition (2.75), C_2 is chosen as

$$C_2 = \frac{1}{2\theta}\mathbb{E}\left(e^{-2\theta Y_i}\right) v_*^2 - \left(\frac{1}{2\theta} - \frac{\beta}{\sigma^2}\right) {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta}{\sigma^2}v_*^2\right) v_*^2, \tag{2.79}$$

where we have used (2.24). With this, the expression of $H(v)$ is obtained in the continuation region $(-v_*, v_*)$. In the stopping region $|v| \geq v_*$, the stopping time in (2.73) is simply $\tau_* = 0$, because $|V_0| = |v| \geq v_*$. Hence, if $|v| \geq v_*$, the objective value achieved by the sampling time (2.73) is

$$\mathbb{E}_v\left[-\gamma v^2 - \int_0^0 (V_s^2 - \beta) ds\right] = -\gamma v^2. \tag{2.80}$$

Combining (2.77)-(2.80), we obtain a candidate of the value function for (2.72):

$$H(v) = \begin{cases} -\frac{v^2}{2\theta} + \left(\frac{1}{2\theta} - \frac{\beta}{\sigma^2}\right) {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta}{\sigma^2} v^2\right) v^2 + C_2, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } \ |v| < v_*, \\ -\gamma v^2, \qquad\qquad\qquad\qquad\qquad\qquad \text{if } \ |v| \geq v_*. \end{cases} \tag{2.81}$$

Next, we find a candidate value of the optimal stopping threshold v_*. By taking the gradient of $H(v)$, we get

$$H'(v) = -\frac{v}{\theta} + \left(\frac{\sigma}{\theta^{\frac{3}{2}}} - \frac{2\beta}{\sigma\sqrt{\theta}}\right) F\left(\frac{\sqrt{\theta}}{\sigma} v\right), \quad v \in (-v_*, v_*), \tag{2.82}$$

where

$$F(x) = e^{x^2} \int_0^x e^{-t^2} dt. \tag{2.83}$$

The boundary condition (2.76) implies that v_* is the root of

$$-\frac{v}{\theta} + \left(\frac{\sigma}{\theta^{\frac{3}{2}}} - \frac{2\beta}{\sigma\sqrt{\theta}}\right) F\left(\frac{\sqrt{\theta}}{\sigma} v\right) = -2\gamma v. \tag{2.84}$$

Substituting (3.10), (2.15), and (2.24) into (2.84), yields that v_* is the root of

$$(\mathsf{mse}_\infty - \beta)\, G\left(\frac{\sqrt{\theta}}{\sigma} v\right) = \mathsf{mse}_\infty - \mathsf{mse}_{Y_i}, \tag{2.85}$$

where $G(\cdot)$ is defined in (2.16). Because $\mathsf{mse}_{Y_i} \leq \beta < \mathsf{mse}_\infty$, $G(x)$ is strictly increasing on $[0, \infty)$, and $G(0) = 1$, we know that (2.85) has a unique non-negative root v_*. Further, the root v_* can be expressed as a function $v(\beta)$ of β, where $v(\beta)$ is defined in (2.19). By this, we obtain a candidate solution to (2.72).

Verification of the Optimality of the Candidate Solution

Next, we use Itô's formula to verify the above candidate solution is indeed optimal, as stated in the following theorem:

Theorem 2.5 *If* $\mathsf{mse}_{Y_i} \leq \beta < \mathsf{mse}_\infty$*, then for all* $v \in \mathbb{R}$*,* $H(v)$ *in* (2.81) *is the value function of the optimal stopping problem* (2.72)*. In addition, the optimal stopping time for solving* (2.72) *is* τ_* *in* (2.73)*, where* $v_* = v(\beta)$ *is given by* (2.19)*.*

In order to prove Theorem 4.7, we need to establish the following properties of $H(v)$ in (2.81), for the case that $\mathsf{mse}_{Y_i} \leq \beta < \mathsf{mse}_\infty$ is satisfied in (2.72):

Lemma 2.9 $H(v) = \mathbb{E}_v\big[-\gamma V_{\tau_*}^2 - \int_0^{\tau_*}(V_s^2 - \beta)ds\big]$.

Proof 2.8 *See Appendix 2.L.*

Lemma 2.10 $H(v) \geq -\gamma v^2$ *for all* $v \in \mathbb{R}$*.*

Proof 2.9 *See Appendix 2.M.*

A function $f(v)$ is said to be *excessive* for the process V_t if

$$\mathbb{E}_v f(V_t) \leq f(v), \forall t \geq 0, v \in \mathbb{R}. \tag{2.86}$$

By using Itô's formula in stochastic calculus, we can obtain

Lemma 2.11 *The function* $H(v)$ *is excessive for the process* V_t*.*

Proof 2.10 *See Appendix 2.N.*

Now, we are ready to prove Theorem 4.7.

Proof 2.11 (Proof of Theorem 2.5) *In Lemmas 2.9-2.11, we have shown that* $H(v) = \mathbb{E}_v\big[-\gamma V_{\tau_*}^2 - \int_0^{\tau_*}(V_s^2 - \beta)ds\big]$*,* $H(v) \geq -\gamma v^2$*, and* $H(v)$ *is an excessive function. Moreover, from the proof of Lemma 2.9, we know that* $\mathbb{E}_v[\tau_*] < \infty$ *holds for all* $v \in \mathbb{R}$*. Hence,* $\mathbb{P}_v(\tau_* < \infty) = 1$ *for all* $v \in \mathbb{R}$*. These conditions and Theorem 1.11 in [74, Section 1.2] imply that* τ_* *is an optimal stopping time of* (2.72)*. This completes the proof.*

Because (2.70) is a special case of (2.72), we can get from Theorem 2.5 that

Corollary 2.1 *If* $\mathsf{mse}_{Y_i} \leq \beta < \mathsf{mse}_\infty$*, then a solution to* (2.70) *is* $(Z_1(\beta), Z_2(\beta), \ldots)$*, where*

$$Z_i(\beta) = \inf\{t \geq 0 : |O_{Y_i+t}| \geq v(\beta)\}, \tag{2.87}$$

and $v(\beta)$ *is defined in* (2.19)*.*

2.6.6 Zero Duality Gap between (2.62) and (2.66)

Strong duality is established in the following theorem:

Theorem 2.6 *If the service times Y_i are i.i.d. with $0 < \mathbb{E}[Y_i] < \infty$, then the duality gap between* (2.62) *and* (2.66) *is zero. Further,* $(Z_0(\beta), Z_1(\beta), \ldots)$ *is an optimal solution to both* (2.62) *and* (2.66)*, where $Z_i(\beta)$ is determined by*

$$Z_i(\beta) = \inf\{t \geq 0 : |O_{Y_i+t}| \geq v(\beta)\}, \tag{2.88}$$

$v(\beta)$ is defined in (2.19)*, $\beta \geq 0$ is the root of*

$$\mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i(\beta)+Y_{i+1}} O_t^2 dt\right] - \beta\mathbb{E}[Y_i + Z_i(\beta)] = 0, \tag{2.89}$$

if $\mathbb{E}[Y_i + Z_i(\beta)] > 1/f_{\max}$; otherwise, β is the root of $\mathbb{E}[Y_i + Z_i(\beta)] = 1/f_{\max}$. In both cases, $\mathsf{mse}_{Y_i} \leq \beta < \mathsf{mse}_\infty$ is satisfied, and hence (2.19) *is well-defined. Further, the optimal objective value to* (2.61) *is given by*

$$\mathsf{mse}_{\text{opt}} = \frac{\mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i(\beta)+Y_{i+1}} O_t^2 dt\right]}{\mathbb{E}[Y_i + Z_i(\beta)]}. \tag{2.90}$$

Proof 2.12 *We use [91, Prop. 6.2.5] to find a* geometric multiplier *for* (2.62)*. This suggests that the duality gap between* (2.62) *and* (2.66) *must be zero because otherwise there exists no geometric multiplier [91, Prop. 6.2.3(b)]. The details are provided in Appendix 2.O.*

Hence, Theorem 2.2 follows from Theorem 2.6. Because Theorem 2.1 is a special case of Theorem 2.2, Theorem 2.1 is also proven.

2.7 Numerical Comparisons

In this section, we evaluate the estimation error achieved by the following four sampling policies:

1. *Uniform sampling*: Periodic sampling with a period given by $S_{i+1} - S_i = 1/f_{\max}$.

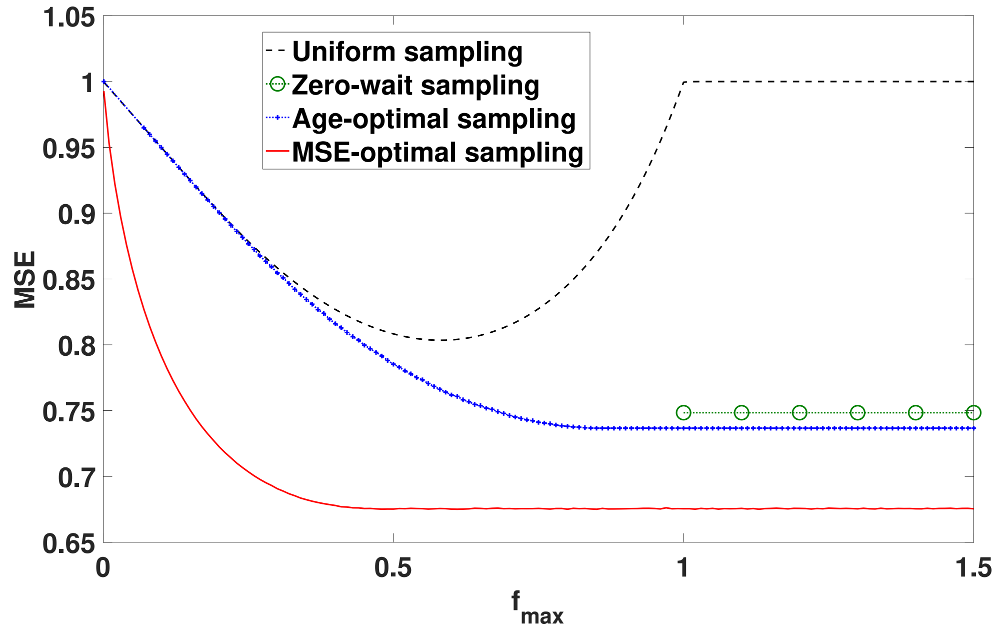


Figure 2.8: MSE vs $f_{\max}$ tradeoff for *i.i.d.* exponential service time with $\mathbb{E}[Y_i] = 1$, where the parameters of the Gauss-Markov process are $\sigma = 1$ and $\theta = 0.5$.

2. *Zero-wait sampling* [1,8]: The sampling policy given by

$$S_{i+1} = S_i + Y_i, \tag{2.91}$$

   which is infeasible when $f_{\max} < 1/\mathbb{E}[Y_i]$.

3. *Age-optimal sampling* [4]: The sampling policy given by Theorem 2.4.

4. *MSE-optimal sampling*: The sampling policy given by Theorem 3.1.

Let $\mathsf{mse}_{\text{uniform}}$, $\mathsf{mse}_{\text{zero-wait}}$, $\mathsf{mse}_{\text{age-opt}}$, and $\mathsf{mse}_{\text{opt}}$, be the MSEs of uniform sampling, zero-wait sampling, age-optimal sampling, MSE-optimal sampling, respectively. We can obtain

$$\begin{aligned} &\mathsf{mse}_{Y_i} \leq \mathsf{mse}_{\text{opt}} \leq \mathsf{mse}_{\text{age-opt}} \leq \mathsf{mse}_{\text{uniform}} \leq \mathsf{mse}_{\infty}, \\ &\mathsf{mse}_{\text{age-opt}} \leq \mathsf{mse}_{\text{zero-wait}} \leq \mathsf{mse}_{\infty}, \end{aligned} \tag{2.92}$$

whenever zero-wait sampling is feasible, which fit with our numerical results. The expectations in (2.25) and (2.26) are evaluated by taking the average over 1 million samples. The parameters of the Gauss-Markov process are given by $\sigma = 1$, $\theta = 0.5$, and μ can be chosen arbitrarily because it does not affect the estimation error.

Figure 2.8 illustrates the tradeoff between the MSE and $f_{\max}$ for *i.i.d.* exponential service times with mean $\mathbb{E}[Y_i] = 1$. Because $\mathbb{E}[Y_i] = 1$, the maximum throughput of the queue is 1. The lower bound mse_{Y_i} is 0.5 and the upper bound mse_{∞} is 1. In fact, as Y_i is an exponential random variable with mean 1, $\frac{\sigma^2}{2\theta}(1 - e^{-2\theta Y_i})$ has a uniform distribution on $[0, 1]$. Hence, $\mathsf{mse}_{Y_i} = 0.5$. For small values of $f_{\max}$, age-optimal sampling is similar to uniform sampling, and hence $\mathsf{mse}_{\text{age-opt}}$ and $\mathsf{mse}_{\text{uniform}}$ are close to each other in the regime. However, as $f_{\max}$ grows, $\mathsf{mse}_{\text{uniform}}$ reaches the upper bound mse_{∞} and remains constant for $f_{\max} \geq 1$. This is because the queue length of uniform sampling is large at high sampling frequencies. In particular, when $f_{\max} \geq 1$, the queue length of uniform sampling is infinite. On the other hand, $\mathsf{mse}_{\text{age-opt}}$ and $\mathsf{mse}_{\text{opt}}$ decrease with respect to $f_{\max}$. The reason behind this is that the set of feasible sampling policies satisfying the constraint in (2.11) and (2.48) becomes larger as $f_{\max}$ grows, and hence the optimal values of (2.11) and (2.48) are decreasing in $f_{\max}$. As we expected, $\mathsf{mse}_{\text{zero-wait}}$ is larger than $\mathsf{mse}_{\text{opt}}$ and $\mathsf{mse}_{\text{age-opt}}$. Moreover, all of them are between the lower bound mse_{Y_i} and upper bound mse_{∞}.

Figures 2.9 and 2.10 depict the MSE of *i.i.d.* normalized log-normal service time for $f_{\max} = 0.8$ and $f_{\max} = 1.2$, respectively, where $Y_i = e^{\alpha X_i}/\mathbb{E}[e^{\alpha X_i}]$, $\alpha > 0$ is the scale parameter of log-normal distribution, and $(X_1, X_2, \dots)$ are *i.i.d.* Gaussian random variables with zero mean and unit variance. Because $\mathbb{E}[Y_i] = 1$, the maximum throughput of the queue is 1. In Fig. 2.9, since $f_{\max} < 1$, zero-wait sampling is not feasible and hence is not plotted. As the scale parameter α grows, the tail of the log-normal distribution becomes heavier.

In both figures, $\mathsf{mse}_{\text{age-opt}}$ and $\mathsf{mse}_{\text{opt}}$ drop with α. This phenomenon may look surprising at first sight, because $\mathsf{mse}_{\text{age-opt}}$ and $\mathsf{mse}_{\text{opt}}$ grow quickly in α in the previous study [35] on the Wiener process. To understand this phenomenon, let us consider the age penalty function $p(\Delta(t))$ in (2.46) for the OU process. As the scale parameter α grows, the service time tends to become either shorter or much longer than the mean $\mathbb{E}[Y_i]$, rather than being close to $\mathbb{E}[Y_i]$. When $\Delta(t)$ is small, $p(\Delta(t))$ reduces quickly in $\Delta(t)$, and hence the service time smaller than $\mathbb{E}[Y_i]$ leads to a fast decrease in the average age penalty; when $\Delta(t)$ is quite large, $p(\Delta(t))$ cannot increase much because it is upper bounded by mse_{∞}, hence the service time much

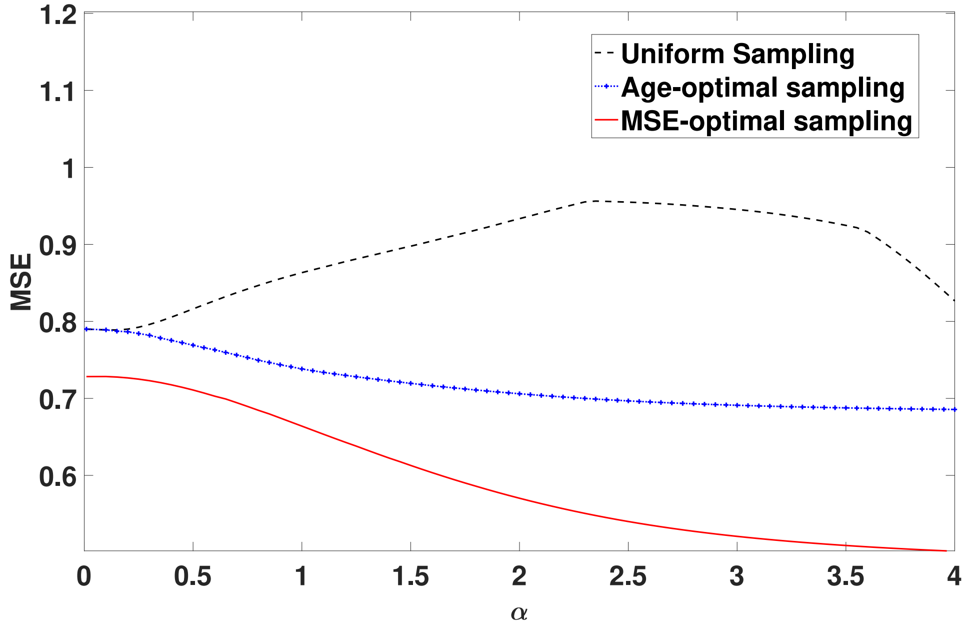


Figure 2.9: MSE vs. the scale parameter α of *i.i.d.* normalized log-normal service time distribution with $\mathbb{E}[Y_i] = 1$ and $f_{\max} = 0.8$, where the parameters of the Gauss-Markov process are $\sigma = 1$ and $\theta = 0.5$. Zero-wait sampling is not feasible here as $f_{\max} < 1/\mathbb{E}[Y_i]$ and hence is not plotted.

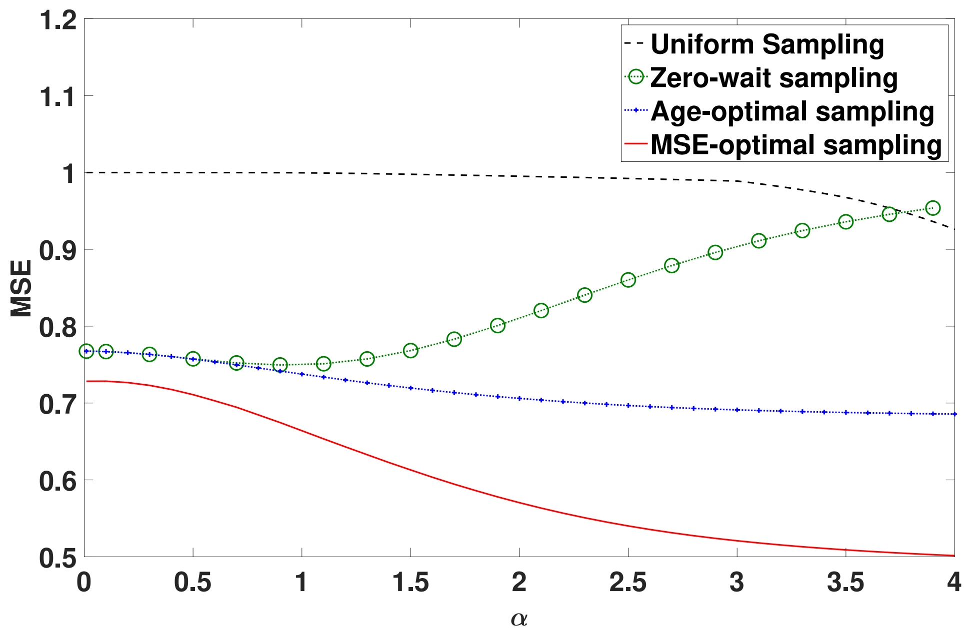


Figure 2.10: MSE vs. the scale parameter α of *i.i.d.* normalized log-normal service time distribution $\mathbb{E}[Y_i] = 1$ and $f_{\max} = 1.2$, where the parameters of the Gauss-Markov process are $\sigma = 1$, $\theta = 0.5$.

longer than $\mathbb{E}[Y_i]$ would not increase the average age penalty by much. By combining these two trends, the average age penalty $\mathsf{mse}_{\text{age-opt}}$ decreases in α. The dropping of $\mathsf{mse}_{\text{opt}}$ in α can be understood similarly. On the other hand, the age penalty function of the Wiener process is $p(\Delta(t)) = \Delta(t)$, which is quite different from the case considered here. We also observe that in both figures, the gap between $\mathsf{mse}_{\text{opt}}$ and $\mathsf{mse}_{\text{age-opt}}$ increases as α grows.

2.8 Conclusion

In this chapter, we have studied the optimal sampler design for remote estimation of the scalar Gauss-Markov processes through queues. We have developed optimal causal sampling policies that minimize the estimation error of Gauss-Markov processes subject to a sampling rate constraint. These optimal sampling policies have nice structures and are easy to compute. A connection between remote estimation and nonlinear age metrics has been found. The structural properties of the optimal sampling policies shed light on the possible structure of the optimal sampler designs for more general signal models, such as Feller processes, which is an important future research direction.

Appendix

2.A Proof of Lemma 2.2

Part (i): According to (2.64) and (2.67), the Lagrangian $L(\pi, \beta)$ is linear and strictly decreasing in β. Further, (2.65) tells us that $f(\beta)$ is the infimum of $L(\pi, \beta)$ among all policies $\pi \in \Pi_1$. Because the infimum of a linear and strictly decreasing function is concave and strictly decreasing, $f(\beta)$ is concave and strictly decreasing in β. Moreover, because $f(\beta)$ is concave, it is also continuous.

Part (ii): We first show that $f(\mathsf{mse}_{Y_i}) > 0$. According to (2.19), $v(\mathsf{mse}_{Y_i}) = 0$. This, together with (2.25), (2.26), and (2.30), implies

$$\begin{aligned} f(\mathsf{mse}_{Y_i}) =& f_1(\mathsf{mse}_{Y_i}) - \mathsf{mse}_{Y_i} f_2(\mathsf{mse}_{Y_i}) \\ =& \mathsf{mse}_\infty\{\mathbb{E}[Y_i] - \gamma\} + \mathbb{E}[O_{Y_i}^2]\gamma - \mathsf{mse}_{Y_i}\mathbb{E}[Y_i] \\ =& \frac{\sigma^2}{2\theta}\left\{\mathbb{E}[Y_i] - \frac{1}{2\theta}\mathbb{E}\left[1 - e^{-2\theta Y_i}\right] + \frac{1}{2\theta}\left\{\mathbb{E}\left[1 - e^{-2\theta Y_i}\right]\right\}^2 - \mathbb{E}\left[1 - e^{-2\theta Y_i}\right]\mathbb{E}[Y_i]\right\}. \end{aligned} \tag{2.93}$$

Therefore, it suffices to prove that

$$\mathbb{E}[Y_i] - \frac{1}{2\theta}\mathbb{E}[1 - e^{-2\theta Y_i}] + \frac{1}{2\theta}\left\{\mathbb{E}[1 - e^{-2\theta Y_i}]\right\}^2 - \mathbb{E}[1 - e^{-2\theta Y_i}]\mathbb{E}[Y_i] > 0, \tag{2.94}$$

which can be simplified as

$$\left(\mathbb{E}[Y_i] - \frac{1}{2\theta}\mathbb{E}[1 - e^{-2\theta Y_i}]\right)\mathbb{E}[e^{-2\theta Y_i}] > 0. \tag{2.95}$$

Because $x > 1 - e^{-x}$ for all $x > 0$ and $\mathbb{E}[Y_i] > 0$, we get

$$\mathbb{E}[2\theta Y_i] > \mathbb{E}[1 - e^{-2\theta Y_i}]. \tag{2.96}$$

By this, $f(\mathsf{mse}_{Y_i}) > 0$ is proven.

Finally, we prove that $f(\mathsf{mse}_\infty) < 0$. When $\beta \to \mathsf{mse}_\infty^-$, (2.19) tells us that $v(\beta)$ grows to infinite. Further, according to (2.25) and (2.26), $R_1(v(\beta))$ and $R_2(v(\beta))$ are quite large compared to $R_1(O_{Y_i})$ and $R_2(O_{Y_i})$. Therefore,

$$\lim_{\beta\to\mathsf{mse}_\infty^-} f(\beta) = -\frac{1}{2\theta}\lim_{\beta\to\mathsf{mse}_\infty^-} v^2(\beta)\mathbb{E}[e^{-2\theta Y_i}] - \mathsf{mse}_\infty\gamma$$
$$= -\infty. \tag{2.97}$$

This completes the proof.

2.B Proof of Lemma 2.3

Part (i): From (2.22), it is evident that the function $f_2(\beta)$ is continuous and hence, from (2.40), $g(\beta)$ is also continuous. The derivatives of $R_1(v)$ in (2.25) and $v(\beta)$ in (2.19) are given by

$$R_1'(v) = \frac{\sqrt{\pi}}{\sigma\sqrt{\theta}}\mathrm{erf}\left(\frac{\sqrt{\theta}}{\sigma}v\right)e^{\frac{\theta}{\sigma^2}v^2}, \tag{2.98}$$

$$v'(\beta) = \frac{\sigma}{\sqrt{\theta}}\left\{G^{-1}\left(\frac{\mathsf{mse}_\infty - \mathsf{mse}_{Y_i}}{\mathsf{mse}_\infty - \beta}\right)\right\}'. \tag{2.99}$$

Denote

$$G^{-1}\left(\frac{\mathsf{mse}_\infty - \mathsf{mse}_{Y_i}}{\mathsf{mse}_\infty - \beta}\right) = y. \tag{2.100}$$

Then, by using the derivative of inverse function [95], $v'(\beta)$ in (2.99) becomes

$$v'(\beta) = \frac{\sigma}{\sqrt{\theta}}\frac{1}{G'(y)}\frac{\mathsf{mse}_\infty - \mathsf{mse}_{Y_i}}{(\mathsf{mse}_\infty - \beta)^2}, \tag{2.101}$$

where

$$G'(x) = \sqrt{\pi}e^{x^2}\mathrm{erf}(x) - \frac{e^{x^2}}{x^2}\frac{\sqrt{\pi}}{2}\mathrm{erf}(x) + \frac{1}{x} > 0 \tag{2.102}$$

for all $x > 0$. Hence, $v(\beta)$ is strictly increasing in β. From (2.98), we know $R_1'(v) > 0$, i.e., $R_1(v)$ is strictly increasing in v. Therefore, $R_1(v(\beta))$ is strictly increasing in β. This further implies that in (2.22), $\max\{R_1(v(\beta)) - R_1(O_{Y_i}), 0\}$ is strictly increasing in β. Therefore,

$\mathbb{E}[\max\{R_1(v(\beta)) - R_1(O_{Y_i}), 0\}]$ is also strictly increasing in β and hence, $f_2(\beta)$ is strictly increasing in β. Then, by (2.40), $g(\beta)$ is strictly decreasing in β. This completes the proof.

Part (ii): We first show that $g(\mathsf{mse}_{Y_i}){\geq}0$. If the root of (2.20) does not satisfy (2.37), then, let β^* is the root of (2.40). Therefore, $g(\beta^*) = 0$. As $\mathsf{mse}_{Y_i} \leq \beta \leq \mathsf{mse}_\infty$ and from part (i), $g(\beta)$ is strictly decreasing in β, we get that

$$g(\mathsf{mse}_{Y_i}) \geq g(\beta^*) = 0. \tag{2.103}$$

Hence, $g(\mathsf{mse}_{Y_i}) \geq 0$.

Finally, as $\beta{\to}\mathsf{mse}_\infty^-$, because $v(\beta)$ grows to infinite, $R_1(v(\beta))$ becomes quite large compared to $R_1(O_{Y_i})$. Hence,

$$\begin{aligned}\lim_{\beta\to\mathsf{mse}_\infty^-} g(\beta) =&\frac{1}{f_{\max}} - \lim_{\beta\to\mathsf{mse}_\infty^-} R_1(v(\beta))\\ =&-\infty.\end{aligned} \tag{2.104}$$

This complete the proof.

2.C Proof of Equation (2.9)

The MMSE estimator $\hat{X}_t$ can be expressed as

$$\begin{aligned}\hat{X}_t &= \mathbb{E}[X_t|M_t]\\ &= \mathbb{E}[X_t|\{(S_j, X_{S_j}, D_j) : D_j \leq t\}].\end{aligned} \tag{2.105}$$

Recall that $U_t = \max\{S_i : D_i \leq t\}$ is the generation time of the latest received sample at time t. According to the strong Markov property of X_t [74, Eq. (4.3.27)] and the fact that the Y_i's are independent of $\{X_t, t \geq 0\}$, $\{U_t, X_{U_t}\}$ is a sufficient statistic for estimating X_t based on $\{(S_j, X_{S_j}, D_j) : D_j \leq t\}$. If $t \in [D_i, D_{i+1})$, (2.4) suggests that $U_t = S_i$ and

$X_{U_t} = X_{S_i}$. This and (2.8) tell us that, if $t \in [D_i, D_{i+1})$, then

$$\begin{aligned}\hat{X}_t =& \ \mathbb{E}[X_t|\{(S_i, X_{S_i}, D_i) : D_i \leq t\}] \\ =& \ \mathbb{E}[X_t|S_i, X_{S_i}] \\ =& \ X_{S_i} e^{-\theta(t-S_i)} + \mu\big[1 - e^{-\theta(t-S_i)}\big]. \end{aligned} \tag{2.106}$$

This completes the proof.

2.D Proof of Lemma 2.4

In any signal-ignorant policy, because the sampling times S_i and the service times Y_i are both independent of $\{X_t, t \geq 0\}$, the delivery times D_i are also independent of $\{X_t, t \geq 0\}$. Hence, for any $t \in [D_i, D_{i+1})$,

$$\begin{aligned}& \mathbb{E}\left[(X_t - \hat{X}_t)^2 \middle| S_i, D_i, D_{i+1}\right] \\ \overset{(a)}{=} & \ \mathbb{E}\left[\frac{\sigma^2}{2\theta} e^{-2\theta(t-S_i)} W^2_{e^{2\theta(t-S_i)}-1} \middle| S_i, D_i, D_{i+1}\right] \\ \overset{(b)}{=} & \ \frac{\sigma^2}{2\theta}\left[1 - e^{-2\theta(t-S_i)}\right], \end{aligned} \tag{2.107}$$

where Step (a) is due to (2.8)-(2.9) and Step (b) is due to $\mathbb{E}[W_t^2] = t$ for all constant $t \geq 0$. We note that in signal-aware sampling policies,

$$(X_t - \hat{X}_t)^2 = \frac{\sigma^2}{2\theta} e^{-2\theta(t-S_i)} W^2_{e^{2\theta(t-S_i)}-1} \tag{2.108}$$

could be correlated with (S_i, D_i, D_{i+1}) and hence Step (b) of (2.107) may not hold. Substituting (2.4) into (2.107), yields that for all $t \geq 0$

$$\mathbb{E}\left[(X_t - \hat{X}_t)^2 \middle| \pi, Y_1, Y_2, \ldots\right] = \frac{\sigma^2}{2\theta}\left(1 - e^{-2\theta\Delta_t}\right), \tag{2.109}$$

which is strictly increasing in Δ_t. This completes the proof.

2.E Proof of Equation (2.42)

When $\sigma = 1$, (2.85) can be expressed as

$$(1 - 2\theta\beta)\, G\left(\sqrt{\theta}v\right) = \mathbb{E}\left[e^{-2\theta Y_i}\right], \tag{2.110}$$

The error function $\mathrm{erf}(x)$ has a Maclaurin series representation, given by

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\left[x - \frac{x^3}{3} + o(x^3)\right]. \tag{2.111}$$

Hence, the Maclaurin series representation of $G(x)$ in (2.16) is

$$G(x) = 1 + \frac{2x^2}{3} + o(x^2). \tag{2.112}$$

Let $x = \sqrt{\theta}v$, we get

$$G\left(\sqrt{\theta}v\right) = 1 + \frac{2}{3}\theta v^2 + o(\theta). \tag{2.113}$$

In addition,

$$\mathbb{E}\left[e^{-2\theta Y_i}\right] = 1 - 2\theta\mathbb{E}[Y_i] + o(\theta). \tag{2.114}$$

Hence, (2.110) can be expressed as

$$(1 - 2\beta\theta)\left[1 + \frac{2}{3}\theta v^2 + o(\theta)\right] = 1 - 2\theta\mathbb{E}[Y_i] + o(\theta). \tag{2.115}$$

Expanding (2.115), yields

$$2\theta\mathbb{E}[Y_i] - 2\beta\theta + \frac{2}{3}\theta v^2 + o(\theta) = 0. \tag{2.116}$$

Dividing by θ and letting $\theta \to 0$ on both sides of (2.116), yields

$$v^2 - 3(\beta - \mathbb{E}[Y_i]) = 0. \tag{2.117}$$

Equation (2.117) has two roots $-\sqrt{3(\beta - \mathbb{E}[Y_i])}$, and $\sqrt{3(\beta - \mathbb{E}[Y_i])}$. If $v_* = -\sqrt{3(\beta - \mathbb{E}[Y_i])}$, the free boundary problem in (2.74)-(2.76) are invalid. Hence, as $\theta \to 0$ and $\sigma = 1$, the root of (2.19) is $v_* = \sqrt{3(\beta - \mathbb{E}[Y_i])}$. This completes the proof.

2.F Proof of Lemma 2.5

We first prove (4.100). It is known that the OU process O_t is a Feller process [96, Section 5.5]. By using a property of Feller process in [96, Theorem 3.32], we get that

$$\begin{aligned} &O_t^2 - \int_0^t \mathcal{G}(O_s^2)ds \\ =&O_t^2 - \int_0^t (-\theta O_s 2 O_s + \sigma^2)ds \\ =&O_t^2 - \sigma^2 t + 2\theta \int_0^t O_s^2 ds \end{aligned} \tag{2.118}$$

is a martingale, where $\mathcal{G}$ is the infinitesimal generator of O_t defined in (4.89). According to [76], the minimum of two stopping times is a stopping time and constant times are stopping times. Hence, $t \wedge \tau$ is a bounded stopping time for every $t \in [0, \infty)$, where $x \wedge y = \min\{x, y\}$. Then, by Theorem 8.5.1 of [76], for every $t \in [0, \infty)$

$$\mathbb{E}\left[\int_0^{t\wedge\tau} {O_s}^2 ds\right] = \mathbb{E}\left[\frac{\sigma^2}{2\theta}(t\wedge\tau)\right] - \mathbb{E}\left[\frac{1}{2\theta}O_{t\wedge\tau}^2\right]. \tag{2.119}$$

Because $\mathbb{E}\left[\int_0^{t\wedge\tau} O_s^2 ds\right]$ and $\mathbb{E}[t\wedge\tau]$ are positive and increasing with respect to t, by using the monotone convergence theorem [76, Theorem 1.5.5], we get

$$\lim_{t\to\infty} \mathbb{E}\left[\int_0^{t\wedge\tau} O_s^2 ds\right] = \mathbb{E}\left[\int_0^{\tau} O_s^2 ds\right], \tag{2.120}$$

$$\lim_{t\to\infty} \mathbb{E}[(t\wedge\tau)] = \mathbb{E}[\tau]. \tag{2.121}$$

In addition, according to [97, Theorem 2.2],

$$\mathbb{E}\left[\max_{0\le s\le\tau} O_s^2\right] \le \frac{C}{\theta}\mathbb{E}\left[\log\left(1 + \frac{\theta\tau}{\sigma}\right)\right] \le \frac{C}{\sigma}\mathbb{E}\left[\tau\right] < \infty. \tag{2.122}$$

Because $O^2_{t\wedge\tau} \leq \sup_{0\leq s\leq\tau} O_s^2$ for all t and $\sup_{0\leq s\leq\tau} O_s^2$ is integratable, by invoking the dominated convergence theorem [76, Theorem 1.5.6], we have

$$\lim_{t\to\infty} \mathbb{E}\left[O^2_{t\wedge\tau}\right] = \mathbb{E}\left[O^2_{\tau}\right]. \tag{2.123}$$

Combining (4.105)-(4.108), (4.100) is proven.

We now prove (4.101) and (4.102). By using the solution of the ODE in Appendix 2.K, one can show that $R_1(v)$ in (4.37) is the solution to the following ODE

$$\frac{\sigma^2}{2}R_1''(v) - \theta v R_1'(v) = 1, \tag{2.124}$$

and $R_2(v)$ in (4.38) is the solution to the following ODE

$$\frac{\sigma^2}{2}R_2''(v) - \theta v R_2'(v) = v^2. \tag{2.125}$$

In addition, $R_1(v)$ and $R_2(v)$ are twice continuously differentiable. According to Dynkin's formula in [75, Theorem 7.4.1 and the remark afterwards], because the initial value of O_t is $O_0 = 0$, if τ is the first exit time of a bounded set, then

$$\mathbb{E}_0[R_1(O_\tau)] = R_1(0) + \mathbb{E}_0\left[\int_0^\tau 1 ds\right] = R_1(0) + \mathbb{E}_0[\tau], \tag{2.126}$$

$$\mathbb{E}_0[R_2(O_\tau)] = R_2(0) + \mathbb{E}_0\left[\int_0^\tau O_s^2 ds\right]. \tag{2.127}$$

Because $R_1(0) = R_2(0) = 0$, (4.101) and (4.102) follow. This completes the proof.

2.G Proof of Lemma 2.6

Suppose that in the sampling policy π, sample i is generated when the server is busy sending another sample, and hence sample i needs to wait for some time before being submitted to the server, i.e., $S_i < G_i$. Let us consider a *virtual* sampling policy $\pi' = \{S_0, \ldots, S_{i-1}, G_i, S_{i+1}, \ldots\}$ such that the generation time of sample i is postponed from S_i to G_i. We call policy π' a virtual policy because it may happen that $G_i > S_{i+1}$. However, this will not affect our proof below. We will show that the MSE of the sampling policy π' is smaller than that of the sampling policy $\pi = \{S_0, \ldots, S_{i-1}, S_i, S_{i+1}, \ldots\}$.

Note that $\{X_t : t \in [0, \infty)\}$ does not change according to the sampling policy, and the sample delivery times $\{D_0, D_1, D_2, \ldots\}$ remain the same in policy π and policy π'. Hence, the only difference between policies π and π' is that *the generation time of sample i is postponed from S_i to G_i*. The MMSE estimator under policy π is given by (2.9) and the MMSE estimator under policy π' is given by

$$\begin{aligned}\hat{X}'_t =& \mathbb{E}[X_t | (S_j, X_{S_j}, D_j)_{j \leq i-1}, (G_i, X_{G_i}, D_i)] \\ =& \begin{cases} \mathbb{E}[X_t | X_{G_i}, G_i], & t \in [D_i, D_{i+1}); \\ \mathbb{E}[X_t | X_{S_j}, S_j], & t \in [D_j, D_{j+1}),\ j \neq i. \end{cases}\end{aligned} \tag{2.128}$$

Next, we consider a third virtual sampling policy π'' in which the samples (X_{G_i}, G_i) and (X_{S_i}, S_i) are both delivered to the estimator at the same time D_i. Clearly, the estimator under policy π'' has more information than those under policies π and π'. By following the arguments in Appendix 2.C, one can show that the MMSE estimator under policy π'' is

$$\begin{aligned}\hat{X}''_t =& \mathbb{E}[X_t | (S_j, X_{S_j}, D_j)_{j \leq i}, (G_i, X_{G_i}, D_i)] \\ =& \begin{cases} \mathbb{E}[X_t | X_{G_i}, G_i], & t \in [D_i, D_{i+1}); \\ \mathbb{E}[X_t | X_{S_j}, S_j], & t \in [D_j, D_{j+1}),\ j \neq i. \end{cases}\end{aligned} \tag{2.129}$$

Notice that, because of the strong Markov property of OU process, the estimator under policy π'' uses the fresher sample (X_{G_i}, G_i), instead of the stale sample (X_{S_i}, S_i), to construct $\hat{X}''_t$ during $[D_i, D_{i+1})$. Because the estimator under policy π'' has more information than that under policy π, one can imagine that policy π'' has a smaller estimation error than policy π, i.e.,

$$\mathbb{E}\left\{\int_{D_i}^{D_{i+1}} (X_t - \hat{X}_t)^2 dt\right\} \geq \mathbb{E}\left\{\int_{D_i}^{D_{i+1}} (X_t - \hat{X}''_t)^2 dt\right\}. \tag{2.130}$$

To prove (4.82), we invoke the orthogonality principle of the MMSE estimator [98, Prop. V.C.2] under policy π'' and obtain

$$\mathbb{E}\left\{\int_{D_i}^{D_{i+1}} 2(X_t - \hat{X}_t'')(\hat{X}_t'' - \hat{X}_t)dt\right\} = 0, \tag{2.131}$$

where we have used the fact that (X_{G_i}, G_i) and (X_{S_i}, S_i) are available by the MMSE estimator under policy π''. Next, from (4.83), we can get

$$\begin{aligned}
&\mathbb{E}\left\{\int_{D_i}^{D_{i+1}} (X_t - \hat{X}_t)^2 dt\right\} \\
=&\mathbb{E}\left\{\int_{D_i}^{D_{i+1}} (X_t - \hat{X}_t'')^2 + (\hat{X}_t'' - \hat{X}_t)^2 dt\right\} \\
&+ \mathbb{E}\left\{\int_{D_i}^{D_{i+1}} 2(X_t - \hat{X}_t'')(\hat{X}_t'' - \hat{X}_t)dt\right\} \\
=&\mathbb{E}\left\{\int_{D_i}^{D_{i+1}} (X_t - \hat{X}_t'')^2 + (\hat{X}_t'' - \hat{X}_t)^2 dt\right\} \\
\geq&\mathbb{E}\left\{\int_{D_i}^{D_{i+1}} (X_t - \hat{X}_t'')^2 dt\right\}.
\end{aligned} \tag{2.132}$$

In other words, the estimation error of policy π'' is no greater than that of policy π. Furthermore, by comparing (4.80) and (4.81), we can see that the MMSE estimators under policies π'' and π' are exact the same. Therefore, the estimation error of policy π' is no greater than that of policy π.

By repeating the above arguments for all samples i satisfying $S_i < G_i$, one can show that the sampling policy $\{S_0, G_1, \ldots, G_{i-1}, G_i, G_{i+1}, \ldots\}$ is better than the sampling policy $\pi = \{S_0, S_1, \ldots, S_{i-1}, S_i, S_{i+1}, \ldots\}$. This completes the proof.

2.H Proof of Equation (2.68)

According to Lemma 4.7,

$$\mathbb{E}\left[\int_{Y_i+Z_i}^{Y_i+Z_i+Y_{i+1}} O_s^2 ds\right] = \frac{\sigma^2}{2\theta}\mathbb{E}[Y_{i+1}] - \frac{1}{2\theta}\mathbb{E}\left[O^2_{Y_i+Z_i+Y_{i+1}} - O^2_{Y_i+Z_i}\right]. \tag{2.133}$$

The second term in (2.133) can be expressed as

$$\begin{aligned}\mathbb{E}\left[O_{Y_i+Z_i+Y_{i+1}}^2 - O_{Y_i+Z_i}^2\right] =&\mathbb{E}\left[\left(O_{Y_i+Z_i}e^{-\theta Y_{i+1}} + \frac{\sigma}{\sqrt{2\theta}}e^{-\theta Y_{i+1}}W_{e^{2\theta Y_{i+1}}-1}\right)^2 - O_{Y_i+Z_i}^2\right]\\ =&\mathbb{E}\left[O_{Y_i+Z_i}^2(e^{-2\theta Y_{i+1}}-1) + \frac{\sigma^2}{2\theta}e^{-2\theta Y_{i+1}}W_{e^{2\theta Y_{i+1}}-1}^2\right]\\ &+\mathbb{E}\left[2O_{Y_i+Z_i}e^{-\theta Y_{i+1}}\frac{\sigma}{\sqrt{2\theta}}e^{-\theta Y_{i+1}}W_{e^{2\theta Y_{i+1}}-1}\right]. \end{aligned} \tag{2.134}$$

Because Y_{i+1} is independent of $O_{Y_i+Z_i}$ and W_t, we have

$$\mathbb{E}\left[O_{Y_i+Z_i}^2(e^{-2\theta Y_{i+1}}-1)\right] = \mathbb{E}\left[O_{Y_i+Z_i}^2\right]\mathbb{E}\left[e^{-2\theta Y_{i+1}}-1\right], \tag{2.135}$$

and

$$\begin{aligned}&\mathbb{E}\left[2O_{Y_i+Z_i}e^{-\theta Y_{i+1}}\frac{\sigma}{\sqrt{2\theta}}e^{-\theta Y_{i+1}}W_{e^{2\theta Y_{i+1}}-1}\right]\\ =&\mathbb{E}\left[2O_{Y_i+Z_i}\right]\mathbb{E}\left[e^{-\theta Y_{i+1}}\frac{\sigma}{\sqrt{2\theta}}e^{-\theta Y_{i+1}}W_{e^{2\theta Y_{i+1}}-1}\right]\\ \overset{(a)}{=}&\mathbb{E}\left[2O_{Y_i+Z_i}\right]\mathbb{E}\left[\mathbb{E}\left[e^{-\theta Y_{i+1}}\frac{\sigma}{\sqrt{2\theta}}e^{-\theta Y_{i+1}}W_{e^{2\theta Y_{i+1}}-1}\middle|Y_{i+1}\right]\right]. \end{aligned} \tag{2.136}$$

where Step (a) is due to the law of iterated expectations. Because $\mathbb{E}[W_t]=0$ for all constant $t\geq 0$, it holds for all realizations of Y_{i+1} that

$$\mathbb{E}\left[e^{-\theta Y_{i+1}}\frac{\sigma}{\sqrt{2\theta}}e^{-\theta Y_{i+1}}W_{e^{2\theta Y_{i+1}}-1}\middle|Y_{i+1}\right] = 0. \tag{2.137}$$

Hence,

$$\mathbb{E}\left[2O_{Y_i+Z_i}e^{-\theta Y_{i+1}}\frac{\sigma}{\sqrt{2\theta}}e^{-\theta Y_{i+1}}W_{e^{2\theta Y_{i+1}}-1}\right] = 0. \tag{2.138}$$

In addition,

$$\begin{aligned}
&\mathbb{E}\left[\frac{\sigma^2}{2\theta}e^{-2\theta Y_{i+1}}W^2_{e^{2\theta Y_{i+1}}-1}\right]\\
\overset{(a)}{=}&\frac{\sigma^2}{2\theta}\mathbb{E}\left[\mathbb{E}\left[e^{-2\theta Y_{i+1}}W^2_{e^{2\theta Y_{i+1}}-1}\middle|Y_{i+1}\right]\right]\\
\overset{(b)}{=}&\frac{\sigma^2}{2\theta}\mathbb{E}\left[1-e^{-2\theta Y_{i+1}}\right],
\end{aligned}\tag{2.139}$$

where Step (a) is due to the law of iterated expectations and Step (b) is due to $\mathbb{E}[W_t^2]=t$ for all constant $t\geq 0$. Hence,

$$\begin{aligned}
&\mathbb{E}\left[\int_{Y_i+Z_i}^{Y_i+Z_i+Y_{i+1}}O_s^2ds\right]\\
=&\frac{\sigma^2}{2\theta}\mathbb{E}[Y_{i+1}]+\gamma\mathbb{E}\left[O^2_{Y_i+Z_i}\right]-\frac{\sigma^2}{4\theta^2}\mathbb{E}\left[1-e^{-2\theta Y_{i+1}}\right]\\
=&\frac{\sigma^2}{2\theta}[\mathbb{E}(Y_{i+1})-\gamma]+\mathbb{E}\left[O^2_{Y_i+Z_i}\right]\gamma,
\end{aligned}\tag{2.140}$$

where γ is defined in (2.24). Using this, (2.68) can be shown readily.

2.I Proof of Lemma 2.1

According to (2.8) and (2.9), the estimation error $(X_t-\hat{X}_t)$ is of the same distribution with $O_{t-S_i(\beta)}$ for $t\in[D_i(\beta),D_{i+1}(\beta))$. We will use $(X_t-\hat{X}_t)$ and $O_{t-S_i(\beta)}$ interchangeably for $t\in[D_i(\beta),D_{i+1}(\beta))$. In order to prove Lemma 2.1, we need to consider the following two cases:

Case 1: If $|X_{D_i(\beta)}-\hat{X}_{D_i(\beta)}|=|O_{Y_i}|\geq v(\beta)$, then (4.55) tells us $S_{i+1}(\beta)=D_i(\beta)$. Hence,

$$D_i(\beta)=S_i(\beta)+Y_i,\tag{2.141}$$

$$D_{i+1}(\beta)=S_{i+1}(\beta)+Y_{i+1}=D_i(\beta)+Y_{i+1}.\tag{2.142}$$

Using these and the fact that the Y_i's are independent of the OU process, we can obtain

$$\mathbb{E}\left[D_{i+1}(\beta)-D_i(\beta)\middle|O_{Y_i},|O_{Y_i}|\geq v(\beta)\right]=\mathbb{E}[Y_{i+1}],\tag{2.143}$$

and

$$
\begin{aligned}
&\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}(X_t-\hat{X}_t)^2dt\middle|O_{Y_i},|O_{Y_i}|\geq v(\beta)\right]\\
=&\mathbb{E}\left[\int_{Y_i}^{Y_i+Y_{i+1}}O_s^2ds\middle|O_{Y_i},|O_{Y_i}|\geq v(\beta)\right]\\
\overset{(a)}{=}&\frac{\sigma^2}{2\theta}\mathbb{E}[Y_{i+1}]+\gamma O_{Y_i}^2-\frac{\sigma^2}{4\theta^2}\mathbb{E}\left[1-e^{-2\theta Y_{i+1}}\right]\\
=&\mathsf{mse}_\infty[\mathbb{E}(Y_{i+1})-\gamma]+O_{Y_i}^2\gamma,
\end{aligned}
\tag{2.144}
$$

where Step (a) follows from the proof of (2.140).

Case 2: If $|X_{D_i(\beta)}-\hat{X}_{D_i(\beta)}|=|O_{Y_i}|<v(\beta)$, then (4.55) tells us that, almost surely,

$$
|X_{S_{i+1}(\beta)}-\hat{X}_{S_{i+1}(\beta)}|=v(\beta). \tag{2.145}
$$

Let us consider the following equation:

$$
\begin{aligned}
&\mathbb{E}\left[D_{i+1}(\beta)-D_i(\beta)\middle|O_{Y_i},|O_{Y_i}|<v(\beta)\right]\\
=&\mathbb{E}\Big[(D_{i+1}(\beta)-S_{i+1}(\beta))+(S_{i+1}(\beta)-S_i(\beta))\\
&\qquad-(D_i(\beta)-S_i(\beta))\ \Big|O_{Y_i},|O_{Y_i}|<v(\beta)\Big].
\end{aligned}
\tag{2.146}
$$

Because $D_{i+1}(\beta)=S_{i+1}(\beta)+Y_{i+1}$, the remaining task is to find $\mathbb{E}\left[S_{i+1}(\beta)-S_i(\beta)\middle|O_{Y_i},|O_{Y_i}|<v(\beta)\right]$, and $\mathbb{E}\left[D_i(\beta)-S_i(\beta)\middle|O_{Y_i},|O_{Y_i}|<v(\beta)\right]$ to compute (4.208). By invoking Lemma 4.7, we can obtain

$$
\mathbb{E}\left[S_{i+1}(\beta)-S_i(\beta)\middle|O_{Y_i},|O_{Y_i}|<v(\beta)\right]=R_1(v(\beta)), \tag{2.147}
$$

$$
\mathbb{E}\left[D_i(\beta)-S_i(\beta)\middle|O_{Y_i},|O_{Y_i}|<v(\beta)\right]=R_1(|O_{Y_i}|), \tag{2.148}
$$

Substituting (4.206), (4.207), and $D_{i+1}(\beta)=S_{i+1}(\beta)+Y_{i+1}$ in (4.208), we get that

$$
\begin{aligned}
&\mathbb{E}\left[D_{i+1}(\beta)-D_i(\beta)\middle|O_{Y_i},|O_{Y_i}|<v(\beta)\right]\\
=&\mathbb{E}[Y_{i+1}]+R_1(v(\beta))-R_1(|O_{Y_i}|).
\end{aligned}
\tag{2.149}
$$

In addition, let us consider the following equation:

$$
\begin{aligned}
&\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}(X_t-\hat{X}_t)^2dt\middle|O_{Y_i},|O_{Y_i}|<v(\beta)\right]\\
=&\mathbb{E}\left[\int_{S_{i+1}(\beta)}^{D_{i+1}(\beta)}(X_t-\hat{X}_t)^2dt+\int_{S_i(\beta)}^{S_{i+1}(\beta)}(X_t-\hat{X}_t)^2dt\right.\\
&\qquad\left.-\int_{S_i(\beta)}^{D_i(\beta)}(X_t-\hat{X}_t)^2dt\middle|O_{Y_i},|O_{Y_i}|<v(\beta)\right] \qquad (2.150)
\end{aligned}
$$

Next, we need to compute the expectations in (2.150). By invoking Lemma 4.7 again, we can obtain

$$
\begin{aligned}
&\mathbb{E}\left[\int_{S_i(\beta)}^{S_{i+1}(\beta)}(X_t-\hat{X}_t)^2dt\middle|O_{Y_i},|O_{Y_i}|<v(\beta)\right]\\
=&\mathbb{E}\left[\int_0^{Y_i+Z_i}O_s^2ds\middle|O_{Y_i},|O_{Y_i}|<v(\beta)\right]=R_2(v(\beta)), \qquad (2.151)\\
&\mathbb{E}\left[\int_{S_i(\beta)}^{D_i(\beta)}(X_t-\hat{X}_t)^2dt\middle|O_{Y_i},|O_{Y_i}|<v(\beta)\right]\\
=&\mathbb{E}\left[\int_0^{Y_i}O_s^2ds\middle|O_{Y_i},|O_{Y_i}|<v(\beta)\right]=R_2(|O_{Y_i}|). \qquad (2.152)
\end{aligned}
$$

By substituting (4.210), (4.211), and (2.140) in (2.150), we have

$$
\begin{aligned}
&\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}(X_t-\hat{X}_t)^2dt\middle|O_{Y_i},|O_{Y_i}|<v(\beta)\right]\\
=&\mathsf{mse}_\infty[\mathbb{E}(Y_{i+1})-\gamma]+v^2(\beta)\gamma+R_2(v(\beta))-R_2(|O_{Y_i}|). \qquad (2.153)
\end{aligned}
$$

By combining (2.143) and (2.149) of the two cases, yields

$$
\begin{aligned}
&\mathbb{E}\left[D_{i+1}(\beta)-D_i(\beta)\middle|O_{Y_i}\right]\\
=&\mathbb{E}[Y_{i+1}]+\max\{R_1(v(\beta))-R_1(|O_{Y_i}|),0\}. \qquad (2.154)
\end{aligned}
$$

Similarly, by combining (4.200) and (4.212) of the two cases, yields

$$\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}(X_t-\hat{X}_t)^2dt\middle|O_{Y_i}\right]$$
$$=\mathsf{mse}_\infty[\mathbb{E}(Y_{i+1})-\gamma]+\max\{v^2(\beta),O_{Y_i}^2\}\gamma+\max\{R_2(v(\beta))-R_2(|O_{Y_i}|),0\}. \tag{2.155}$$

Finally, by taking the expectation over O_{Y_i} in (4.214) and (4.217) and using the fact that $R_1(\cdot)$ and $R_2(\cdot)$ are even functions, Lemma 2.1 is proven.

2.J Proof of Lemma 2.8

Because the Y_i's are *i.i.d.*, (2.68) is determined by the control decision Z_i and the information (O_{Y_i}, Y_i). Hence, (O_{Y_i}, Y_i) is a *sufficient statistic* for determining Z_i in (2.65). Therefore, there exists an optimal policy $(Z_0, Z_1, \ldots)$ to (2.65), in which Z_i is determined based on only (O_{Y_i}, Y_i). By this, (2.65) is decomposed into a sequence of per-sample MDPs, given by (4.126). This completes the proof.

2.K Proof of Equation (2.77)

Define $S(v) = H'(v)$. Then, (4.130) becomes

$$S'(v)-\frac{2\theta}{\sigma^2}vS(v)=\frac{2}{\sigma^2}(v^2-\beta). \tag{2.156}$$

Equation (2.156) can be solved by using the integrating factor method [94, Sec. I.5], which applies to any ODE of the form

$$S'(v)+a(v)S(v)=b(v). \tag{2.157}$$

In the case of (2.156),

$$a(v)=-\frac{2\theta}{\sigma^2}v,\ b(v)=\frac{2}{\sigma^2}(v^2-\beta). \tag{2.158}$$

The integrating factor of (2.156) is

$$M(v) = e^{\int a(v)dv} = e^{-\frac{\theta}{\sigma^2}v^2}. \tag{2.159}$$

Multiplying $e^{-\frac{\theta}{\sigma^2}v^2}$ on both sides of (2.156) and transforming the left-hand side into a total derivative, yields

$$\left[S(v)e^{-\frac{\theta}{\sigma^2}v^2}\right]' = b(v)e^{-\frac{\theta}{\sigma^2}v^2}. \tag{2.160}$$

Taking the integration on both sides of (2.160), yields

$$\begin{aligned} S(v)e^{-\frac{\theta}{\sigma^2}v^2} &= \int \frac{2}{\sigma^2}(v^2-\beta)e^{-\frac{\theta}{\sigma^2}v^2}dv \\ &= \int \frac{2}{\sigma^2}e^{-\frac{\theta}{\sigma^2}v^2}v^2dv - \int \frac{2}{\sigma^2}\beta e^{-\frac{\theta}{\sigma^2}v^2}dv. \end{aligned} \tag{2.161}$$

The indefinite integrals in (2.161) are given by [99, Sec. 15.3.1, (Eq. 36)]

$$\int \frac{2}{\sigma^2}e^{-\frac{\theta}{\sigma^2}v^2}v^2dv = \frac{\sqrt{\pi}\sigma}{2\theta^{\frac{3}{2}}}\mathrm{erf}\left(\frac{\sqrt{\theta}}{\sigma}v\right) - \frac{v}{\theta}e^{-\frac{\theta}{\sigma^2}v^2} + C_1, \tag{2.162}$$

$$\int \frac{2}{\sigma^2}\beta e^{-\frac{\theta}{\sigma^2}v^2}dv = \frac{\sqrt{\pi}\beta}{\sigma\sqrt{\theta}}\mathrm{erf}\left(\frac{\sqrt{\theta}}{\sigma}v\right) + C_2, \tag{2.163}$$

where $\mathrm{erf}(\cdot)$ is the error function defined in (2.17). Combining (2.161)-(4.134), results in

$$S(v) = \left(\frac{\sqrt{\pi}\sigma}{2\theta^{\frac{3}{2}}} - \frac{\sqrt{\pi}\beta}{\sigma\sqrt{\theta}}\right)\mathrm{erf}\left(\frac{\sqrt{\theta}}{\sigma}v\right)e^{\frac{\theta}{\sigma^2}v^2} - \frac{v}{\theta} + C_3e^{\frac{\theta}{\sigma^2}v^2}, \tag{2.164}$$

where $C_3 = C_1 + C_2$. We need to integrate $S(v)$ in (2.164) again to get $H(v)$

$$\begin{aligned} H(v) &= \int S(v)dv \\ &= \int \left(\frac{\sqrt{\pi}\sigma}{2\theta^{\frac{3}{2}}} - \frac{\sqrt{\pi}\beta}{\sigma\sqrt{\theta}}\right)\mathrm{erf}\left(\frac{\sqrt{\theta}}{\sigma}v\right)e^{\frac{\theta}{\sigma^2}v^2}dv - \int \frac{v}{\theta}dv \\ &\quad + \int C_3e^{\frac{\theta}{\sigma^2}v^2}dv, \end{aligned} \tag{2.165}$$

which requires the following integral [100, Sec. 8.250 (Eq. 1,4)]:

$$\begin{aligned}&\int \operatorname{erf}\left(\frac{\sqrt{\theta}}{\sigma}v\right)e^{\frac{\theta}{\sigma^2}v^2}dv\\ =&\frac{\sigma}{\sqrt{\theta}\sqrt{\pi}}\frac{\theta}{\sigma^2}v^2{}_2F_2\left(1,1;\frac{3}{2},2;\frac{\theta}{\sigma^2}v^2\right)+C.\end{aligned}\tag{2.166}$$

By using (2.166), we can compute the first integral of (2.165)

$$\begin{aligned}&\int\left(\frac{\sqrt{\pi}\sigma}{2\theta^{\frac{3}{2}}}-\frac{\sqrt{\pi}\beta}{\sigma\sqrt{\theta}}\right)\operatorname{erf}\left(\frac{\sqrt{\theta}}{\sigma}v\right)e^{\frac{\theta}{\sigma^2}v^2}dv\\ =&\left(\frac{1}{2\theta}-\frac{\beta}{\sigma^2}\right)v^2\ {}_2F_2\left(1,1;\frac{3}{2},2;\frac{\theta}{\sigma^2}v^2\right)+C_4.\end{aligned}\tag{2.167}$$

The remaining integrals in (2.165) are as follows [100, Sec. 3.478 (Eq. 3)]

$$\int C_3e^{\frac{\theta}{\sigma^2}v^2}dv=C_5\operatorname{erfi}\left(\frac{\sqrt{\theta}}{\sigma}v\right)+C_6,\tag{2.168}$$

$$\int\frac{v}{\theta}dv=-\frac{v^2}{2\theta}+C_7,\tag{2.169}$$

where $\operatorname{erfi}(\cdot)$ is the imaginary error function defined in (4.140). Hence, by substituting (2.167), (2.168), and (2.169) in (2.165), $H(v)$ in (2.77) follows. This completes the proof of (2.77).

2.L Proof of Lemma 2.9

The proof of Lemma 2.9 consists of the following two cases:

Case 1: If $|v|\geq v_*$, (4.129) implies $\tau_*=0$. Hence,

$$\mathbb{E}_v\left[\tau_*\middle||v|\geq v_*\right]=\mathbb{E}_v\left[\int_0^{\tau_*}1ds\middle||v|\geq v_*\right]=0,\tag{2.170}$$

and

$$\mathbb{E}_v\left[\int_0^{\tau_*}{V_s}^2ds\middle||v|\geq v_*\right]=0.\tag{2.171}$$

Because $V_0 = v$, we have

$$\mathbb{E}_v[V_{\tau_*}^2] = \mathbb{E}_v[V_0^2] = v^2. \tag{2.172}$$

By combining (2.170)-(2.172), we get

$$\mathbb{E}_v\left[-\gamma V_{\tau_*}^2 - \int_0^{\tau_*} ({V_s}^2 - \beta) ds \middle| |v| \geq v_*\right] = -\gamma v^2. \tag{2.173}$$

Case 2: If $|v| < v_*$, (2.73) tells us that, almost surely,

$$V_{\tau_*} = v_*. \tag{2.174}$$

Similar to the proof of Lemma 2.1, we can use Lemma 2.5 to obtain

$$\begin{aligned}
&\mathbb{E}_v\left[\tau_* \middle| |v| < v_*\right] \\
=&\mathbb{E}_v\left[\int_0^{\tau_*} 1 ds \middle| |v| < v_*\right] \\
=&R_1(v_*) - R_1(v) \\
=&\frac{v_*^2}{\sigma^2} {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta}{\sigma^2} v_*^2\right) - \frac{v^2}{\sigma^2} {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta}{\sigma^2} v^2\right),
\end{aligned} \tag{2.175}$$

$$\begin{aligned}
&\mathbb{E}_v\left[\int_0^{\tau_*} {V_s}^2 ds \middle| |v| < v_*\right] \\
=&R_2(v_*) - R_2(v) \\
=&-\frac{v_*^2}{2\theta} + \frac{v_*^2}{2\theta}\ {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta}{\sigma^2} v_*^2\right) \\
&+ \frac{v^2}{2\theta} - \frac{v^2}{2\theta}\ {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta}{\sigma^2} v^2\right),
\end{aligned} \tag{2.176}$$

and

$$\mathbb{E}_v\left[V_{\tau_*}^2 \middle| |v| < v_*\right] = v_*^2. \tag{2.177}$$

Combining (2.175)-(2.177), yields

$$\begin{aligned}
&\mathbb{E}_v\left[-\gamma V_{\tau_*}^2 - \int_0^{\tau_*} ({V_s}^2-\beta)ds \middle| |v|<v_*\right]\\
=&-\frac{v^2}{2\theta}+\left(\frac{1}{2\theta}-\frac{\beta}{\sigma^2}\right){}_2F_2\big(1,1;\frac{3}{2},2;\frac{\theta}{\sigma^2}v^2\big)v^2\\
&+\frac{1}{2\theta}\mathbb{E}\left(e^{-2\theta Y_i}\right)v_*^2-\left(\frac{1}{2\theta}-\frac{\beta}{\sigma^2}\right){}_2F_2\big(1,1;\frac{3}{2},2;\frac{\theta}{\sigma^2}v_*^2\big)v_*^2. \qquad (2.178)
\end{aligned}$$

By combining (2.173) and (2.178), Lemma 2.9 is proven.

2.M Proof of Lemma 2.10

The proof of Lemma 2.10 consists of the following two cases:

Case 1: If $|v| \geq v_*$, (2.81) tells us that

$$H(v) = -\gamma v^2. \qquad (2.179)$$

Hence, Lemma 2.10 holds in *Case 1*.

Case 2: $|v| < v_*$. Because $H(v)$ is an even function and $H(v) = -\gamma v^2$ holds at $v = \pm v_*$, to prove $H(v) \geq -\gamma v^2$ for $|v| < v_*$, it is sufficient to show that for all $v \in [0, v_*)$

$$H'(v) < [-\gamma v^2]' = -2\gamma v. \qquad (2.180)$$

Hence, the remaining task is to prove that (2.180) holds for $v \in [0, v_*)$.

After some manipulations, we can obtain from (2.85) that

$$(\mathsf{mse}_\infty - \beta)\, G\left(\frac{\sqrt{\theta}}{\sigma}v_*\right) = \mathsf{mse}_\infty \mathbb{E}(e^{-2\theta Y_i}). \qquad (2.181)$$

Because $G(\cdot) > 0$ is an increasing function, it holds for all $v \in [0, v_*)$ that

$$\begin{aligned}
(\mathsf{mse}_\infty - \beta)\, G\left(\frac{\sqrt{\theta}}{\sigma}v\right) <& (\mathsf{mse}_\infty - \beta)\, G\left(\frac{\sqrt{\theta}}{\sigma}v_*\right)\\
=&\mathsf{mse}_\infty \mathbb{E}(e^{-2\theta Y_i}). \qquad (2.182)
\end{aligned}$$

One can obtain (2.180) from (4.144) and (2.182). Hence, Lemma 2.10 also holds in *Case 2*. This completes the proof.

2.N Proof of Lemma 2.11

We need the following lemma in the proof of Lemma 2.11:

Lemma 2.12 $(1-2x^2)G(x) \le 1$ *for all* $x \ge 0$.

Proof 2.13 *Because* $G(0) = 1$, *it suffices to show that for all* $x > 0$

$$[(1-2x^2)G(x)]' \le 0. \tag{2.183}$$

We have

$$[(1-2x^2)G(x)]' = -\frac{1}{x^2}e^{x^2}\int_0^x e^{-t^2}dt + \frac{1}{x} - 4x^2 e^{x^2}\int_0^x e^{-t^2}dt - 2x. \tag{2.184}$$

Because e^{-t^2} *is decreasing on* $t \in [0,\infty)$, *for all* $x > 0$

$$\int_0^x e^{-t^2}dt \ge \int_0^x e^{-x^2}dt = xe^{-x^2}. \tag{2.185}$$

Hence,

$$-\frac{1}{x^2}e^{x^2}\int_0^x e^{-t^2}dt + \frac{1}{x} \le 0. \tag{2.186}$$

Substituting (2.186) *into* (2.184), (2.183) *follows. This completes the proof.*

Now we are ready to prove Lemma 2.11.

Proof 2.14 (Proof of Lemma 2.11) *The function* $H(v)$ *is continuously differentiable on* $\mathbb{R}$. *In addition,* $H''(v)$ *is continuous everywhere but at* $v = \pm v_*$. *Since the Lebesgue measure of those time* t *for which* $V_t = \pm v_*$ *is zero, the values* $H''(\pm v_*)$ *can be chosen in the sequel* *arbitrarily. By using Itô's formula [101, Theorem 7.13], we obtain that almost surely*

$$
\begin{aligned}
&H(V_t) - H(v) \\
=&\int_0^t \frac{\sigma^2}{2}\left[H''(V_r) - \theta V_r H'(V_r) - (V_r^2 - \beta)\right] dr + \int_0^t \sigma H'(V_r) dW_r. \qquad (2.187)
\end{aligned}
$$

For all $t \geq 0$ and all $v \in \mathbb{R}$, we can show that

$$
\mathbb{E}_v \left\{ \int_0^t \left[\sigma H'(V_r)\right]^2 dr \right\} < \infty.
$$

This and [101, Theorem 7.11] imply that $\int_0^t \sigma H'(V_r) dW_r$ is a martingale and

$$
\mathbb{E}_v \left[\int_0^t \sigma H'(V_r) dW_r \right] = 0, \ \forall \ t \geq 0. \qquad (2.188)
$$

Hence,

$$
\begin{aligned}
&\mathbb{E}_v \left[H(V_t) - H(v)\right] \\
=&\mathbb{E}_v \left[\int_0^t \frac{\sigma^2}{2}\left[H''(V_r) - \theta V_r H'(V_r) - (V_r^2 - \beta)\right] dr \right]. \qquad (2.189)
\end{aligned}
$$

Next, we show that for all $v \in \mathbb{R}$

$$
\frac{\sigma^2}{2} H''(v) - \theta v H'(v) - (v^2 - \beta) \leq 0. \qquad (2.190)
$$

Let us consider the following two cases:

Case 1: *If $|v| < v_*$, then* (2.74) *implies*

$$
\frac{\sigma^2}{2} H''(v) - \theta v H'(v) - (v^2 - \beta) = 0. \qquad (2.191)
$$

Case 2: $|v| > v_*$. *In this case,* $H(v) = -\gamma v^2$. *Hence,*

$$\begin{aligned}
&\frac{\sigma^2}{2}H''(v) - \theta v H'(v) \\
=&\frac{\sigma^2}{2}(-2\gamma) - \theta v(-2\gamma v) \\
=& - \sigma^2\gamma + 2\theta\gamma v^2 \\
=& - \mathsf{mse}_{Y_i} + \mathbb{E}[1 - e^{-2\theta Y_i}]v^2.
\end{aligned} \tag{2.192}$$

Substituting (2.192) *into* (2.190), *yields*

$$\mathbb{E}[e^{-2\theta Y_i}]v^2 \geq \beta - \mathsf{mse}_{Y_i}. \tag{2.193}$$

To prove (2.193), *since* $|v| > v_*$, *it suffices to show that*

$$\mathbb{E}[e^{-2\theta Y_i}]v_*^2 \geq \beta - \mathsf{mse}_{Y_i}, \tag{2.194}$$

which is equivalent to

$$(\mathsf{mse}_\infty - \mathsf{mse}_{Y_i})\frac{v_*^2}{\mathsf{mse}_\infty} \geq (\mathsf{mse}_\infty - \mathsf{mse}_{Y_i}) - (\mathsf{mse}_\infty - \beta). \tag{2.195}$$

We now prove (2.195). *By Lemma 2.12, we get*

$$\left(1 - \frac{v_*^2 2\theta}{\sigma^2}\right) G\left(\frac{\sqrt{\theta}}{\sigma}v_*\right) \leq 1. \tag{2.196}$$

Hence,

$$\left(1 - \frac{v_*^2}{\mathsf{mse}_\infty}\right) G\left(\frac{\sqrt{\theta}}{\sigma}v_*\right) \leq 1. \tag{2.197}$$

By substituting (2.85) *into* (2.198), *we obtain*

$$\begin{aligned}&(\mathsf{mse}_\infty - \mathsf{mse}_{Y_i})\left(1-\frac{v_*^2}{\mathsf{mse}_\infty}\right) G\left(\frac{\sqrt{\theta}}{\sigma}v_*\right)\\ \leq\, &(\mathsf{mse}_\infty - \beta)\, G\left(\frac{\sqrt{\theta}}{\sigma}v_*\right).\end{aligned} \tag{2.198}$$

Because $G(x) > 0$ *for all* $x > 0$,

$$(\mathsf{mse}_\infty - \mathsf{mse}_{Y_i})\left(1-\frac{v_*^2}{\mathsf{mse}_\infty}\right) \leq \mathsf{mse}_\infty - \beta, \tag{2.199}$$

which implies (2.195). *Hence,* (2.190) *holds in both cases. Thus,* $\mathbb{E}_v\left[H(V_t) - H(v)\right] \leq 0$ *holds for all* $t \geq 0$ *and* $v \in \mathbb{R}$. *This completes the proof.*

2.O Proof of Theorem 2.6

According to [91, Prop. 6.2.5], if we can find $\pi^\star = (Z_1, Z_2, \ldots)$ and $\lambda^\star$ that satisfy the following conditions:

$$\pi^\star \in \Pi_1, \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\left[Y_i + Z_i\right] - \frac{1}{f_{\max}} \geq 0, \tag{2.200}$$

$$\lambda^\star \geq 0, \tag{2.201}$$

$$L(\pi^\star; \lambda^\star) = \inf_{\pi\in\Pi_1} L(\pi; \lambda^\star), \tag{2.202}$$

$$\lambda^\star\left\{\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\left[Y_i + Z_i\right] - \frac{1}{f_{\max}}\right\} = 0, \tag{2.203}$$

then $\pi^\star$ is an optimal solution to (2.62) and $\lambda^\star$ is a geometric multiplier [91] for (2.62). Further, if we can find such $\pi^\star$ and $\lambda^\star$, then the duality gap between (2.62) and (2.66) must be zero, because otherwise there is no geometric multiplier [91, Prop. 6.2.3(b)]. The remaining task is to find $\pi^\star$ and $\lambda^\star$ that satisfy (2.200)-(2.203).

According to Theorem 4.7 and Corollary 4.2, a solution $\pi^\star = (Z_0(\beta), Z_1(\beta), \ldots)$ to (2.202) is given by (4.159), where $\beta = \mathsf{mse}_{\text{opt}} + \lambda^\star$. In addition, because the Y_i's are *i.i.d.*, the $Z_i(\beta)$'s in policy $\pi^\star$ are *i.i.d.* Using (2.200), (2.201), and (2.203), the value of $\lambda^\star$ can be obtained by considering two cases: If $\lambda^\star > 0$, because the $Z_i(\beta)$'s are *i.i.d.*, we have from (2.203) that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\left[Y_i + Z_i(\beta)\right] = \mathbb{E}\left[Y_i + Z_i(\beta)\right] = \frac{1}{f_{\max}}. \tag{2.204}$$

If $\lambda^\star = 0$, then (2.200) implies

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\left[Y_i + Z_i(\beta)\right] = \mathbb{E}\left[Y_i + Z_i(\beta)\right] \geq \frac{1}{f_{\max}}. \tag{2.205}$$

Next, we compute $\mathsf{mse}_{\text{opt}}$ and $\beta = \mathsf{mse}_{\text{opt}} + \lambda^\star$. To compute $\mathsf{mse}_{\text{opt}}$, we substitute policy $\pi^\star$ into (2.61), which yields

$$\begin{aligned}\mathsf{mse}_{\text{opt}} = & \lim_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i(\beta)+Y_{i+1}} O_s^2 ds\right]}{\sum_{i=0}^{n-1} \mathbb{E}\left[Y_i+Z_i(\beta)\right]} \\ = & \frac{\mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i(\beta)+Y_{i+1}} O_s^2 ds\right]}{\mathbb{E}[Y_i + Z_i(\beta)]},\end{aligned} \tag{2.206}$$

where in the last equation we have used that the $Z_i(\beta)$'s are *i.i.d.* Hence, the value of $\beta = \mathsf{mse}_{\text{opt}} + \lambda^\star$ can be obtained by considering the following two cases:

Case 1: If $\lambda^\star = 0$, then (2.205) and (2.206) imply that

$$\mathbb{E}\left[Y_i + Z_i(\beta)\right] \geq \frac{1}{f_{\max}}, \tag{2.207}$$

$$\beta = \mathsf{mse}_{\text{opt}} = \frac{\mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i(\beta)+Y_{i+1}} O_s^2 ds\right]}{\mathbb{E}\left[Y_i+Z_i(\beta)\right]}. \tag{2.208}$$

Notice that (2.208) can rewritten as (2.36), which is a fixed-point equation on β. According to Lemma 2.2, one root of (2.36) is in the set $(\mathsf{mse}_{Y_i}, \mathsf{mse}_\infty)$, which is also the unique root of (2.89); we denote this root as β_1. We choose $\pi^\star = (Z_0(\beta_1), Z_1(\beta_1)...)$, where $Z_i(\cdot)$ is given by (2.88). In addition, $\lambda^\star$ must be 0 in *Case 1*. Because $\lambda^\star = \beta_1 - \mathsf{mse}_{\text{opt}}$, we get $\mathsf{mse}_{\text{opt}} = \beta_1$, which is required in (2.208). *Case 1* occurs if the root β_1 of (2.208) satisfies (2.207). We note that $\beta = \mathsf{mse}_\infty$ is another root of (2.208), but we do not pick policy $\pi^\star$ based on this root.

Case 2: If $\lambda^\star > 0$, then (2.204) and (2.206) imply that

$$\mathbb{E}\left[Y_i + Z_i(\beta)\right] = \frac{1}{f_{\max}}, \tag{2.209}$$

$$\beta > \mathsf{mse}_{\text{opt}} = \frac{\mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i(\beta)+Y_{i+1}} O_s^2 ds\right]}{\mathbb{E}\left[Y_i + Z_i(\beta)\right]}. \tag{2.210}$$

When the root β_1 of (2.208) does not satisfy (2.207), Lemma 2.3 tells us that (2.209) has a unique root in the set $[\mathsf{mse}_{Y_i}, \mathsf{mse}_\infty)$, which is denoted by β_2. We choose $\pi^\star = (Z_0(\beta_2), Z_1(\beta_2)...)$, where $Z_i(\cdot)$ is given by (2.88). Further, we choose $\lambda^\star = \beta_2 - \mathsf{mse}_{\text{opt}}$.

Theorem 4.7, together with the fact that $\beta_1, \beta_2 \in [\mathsf{mse}_{Y_i}, \mathsf{mse}_\infty)$ and the arguments above, implies that the selected $\pi^\star$ and $\lambda^\star$ satisfy (2.200)-(2.203). By [91, Prop. 6.2.3(b)], the duality gap between (2.62) and (2.66) is zero. A solution to (2.62) and (2.66) is $\pi^\star$. This completes the proof.

Chapter 3

Performance Bounds for Remote Estimation with Noisy Samples

3.1 Introduction

Many real-time applications, such as state estimation, tracking, and decision-making require fresh and timely updates about the system state. In recent years, to measure the freshness of state updates, the concept of AoI has received significant attention from the research community due to its extensive importance in real-time systems [1, 2, 102]. AoI is expressed as a time difference between the current time and the generation time of the latest received sample.

In practice, the system states are usually in the form of a signal X_t, such as interest rate, currency exchange rate, price of the stock market, and the trajectory of a flying UAV. These real-time signals are random, sometimes the variations are small and later may become huge. Hence, the time difference is not sufficient to distinguish the variation of the signal state and the update policy that minimizes AoI does not produce a smaller estimation error. The problem of sampling an Ornstein-Uhlenbeck (OU) process is recently addressed in [34] and another problem of sampling a Wiener process in [55]. However, the optimal sampling policy provided in [34] is only for the stable scenario. In practice, real-time applications of OU processes consider both stable and unstable cases [103]. Therefore, a sampling problem that considers only the stable scenario is insufficient for practical and more dynamic systems, and a generalization of this problem that considers both stable and unstable cases is necessary.

Moreover, a real-time system often consists of noise along with the signal process. Therefore, the analysis based on noisy observation of samples to minimize signal estimation error is practically much more important in real-time networked control and communication systems. In this chapter, we consider noisy samples of OU process and compute the mse from which we establish estimation performance bounds of mse.

The OU process is defined as the solution to the following stochastic differential equation (SDE) [68, 69]

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t, \tag{3.1}$$

where μ, θ, and $\sigma > 0$ are parameters and W_t represents a Wiener process. In case of stable OU process, $\theta > 0$ [34]. In (3.1), if $\theta \to 0$, and $\sigma = 1$, X_t reduces to a Wiener process. If $\theta < 0$, then X_t becomes an unstable OU process. Examples and properties of OU processes are explained in [34].

We consider the samples of the Gauss-Markov process pass through a channel in a first-come, first-serve (FCFS) strategy. A remotely located estimator utilizes these causally received samples to make an estimate $\hat{X}_t$ of X_t. First, we obtain a lower bound of mse in the absence of any additional noise in the system. Second, our goal is to find the expression of mse with the presence of noise in the system. This analysis provides an upper bound of mse when the estimator receives noisy samples. We summarize the contributions of this paper as follows:

- The optimal sampling problem in the absence of noise is formulated and the solved optimal sampling policy is a threshold policy on *instantaneous estimation error*. The structure of the thresholds $v(\beta)$ of a parameter β are different for the three cases: $\theta > 0$ (Stable OU process), $\theta = 0$ (Wiener process), and $\theta < 0$ (Unstable OU process). The value of β is equal to the optimum value of the time-average expected estimation error. The computation of β remains the same irrespective of the signal models.

- Further, we consider noisy samples and obtain an explicit expression for mse. From the expression, we establish a performance upper bound of mse.

- Our results hold for general *i.i.d.* transmission time distributions of the queueing server with a finite mean.

3.2 Model

This section describes the single-source, single-channel model as shown in Figure 3.1.

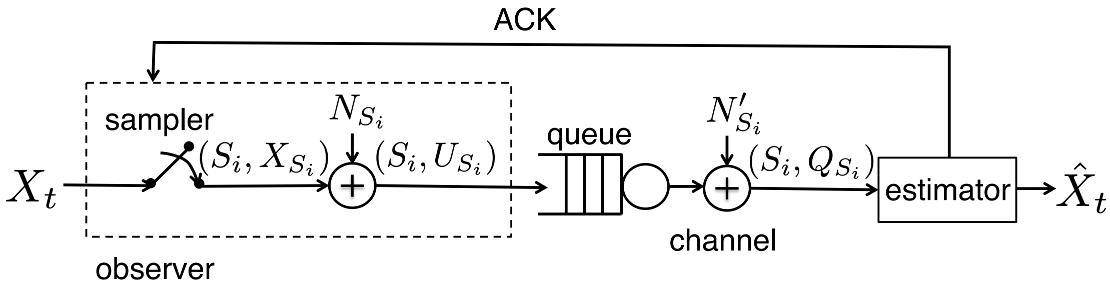


Figure 3.1: A single-source, single-channel remote estimation system consisting of a noisy sample of X_t over a noisy channel.

3.2.1 System Model

We consider a continuous-time remote estimation system that is illustrated in Fig. 3.1, where an observer takes samples from an Gauss-Markov process X_t. After sampling, additional noises from the sampler and the channel are added to the samples. Then, the noisy samples are sent to the estimator. The channel is modeled as a single-server FIFO queue with *i.i.d.* service times. The samples undergo random service times in the channel due to fading, interference, congestions, etc. We also consider that at a time, only one sample can be delivered through the channel.

The operation of the system starts at time instant $t = 0$. The generation time of the i-th sample is S_i, which satisfy $S_i \leq S_{i+1}$ for all i. Then, i-th sample undergoes a random service time Y_i, and is delivered to the estimator at time D_i, where $S_i + Y_i \leq D_i$, $D_i + Y_{i+1} \leq D_{i+1}$, and $0 < \mathbb{E}[Y_i] < \infty$ hold for all i. The i-th sample packet (S_i, X_{S_i}) contains the sample value X_{S_i} and its sampling time S_i. Suppose that after sampling, noise N_{S_i} is being added to the sample X_{S_i} and the noisy observation of the sample X_{S_i} is denoted by U_{S_i}. Hence,

$$U_{S_i} = X_{S_i} + N_{S_i}, \tag{3.2}$$

where N_{S_i} is the additive noise with zero mean and variance b_1. Each sample packet (S_i, U_{S_i}) contains the sampling time S_i and the noisy sample U_{S_i}. If channel noise N'_{S_i} with zero mean and variance b_2 is added to the sample during its transmission through the channel, then the sample value becomes

$$Q_{S_i} = U_{S_i} + N'_{S_i}. \tag{3.3}$$

Initially, at $t = 0$, the state of the system is assumed to hold $S_0 = 0$, and $D_0 = Y_0$. The initial state of the Gauss-Markov process X_0 is a finite constant. The process parameters μ, θ, and σ in (3.1) are known at both the sampler and estimator.

Let, the idle/busy state of the server at time t is denoted by $I_t \in \{0, 1\}$. We also assume that an acknowledgement is immediately sent back to the sampler whenever a sample is delivered and this operation has zero delay. By this assumption, the sampler is aware of the idle/busy state of the server and the available information at time t can be given by $\{X_s, I_s : 0 \leq s \leq t\}$.

3.2.2 Sampling Policies

The sampling time S_i is a finite stopping time with respect to the filtration $\{\mathcal{F}_t^+, t \geq 0\}$ (a non-decreasing and right-continuous family of σ-fields) of the information that is available at the sampler such that [76]

$$\{S_i \leq t\} \in \mathcal{F}_t^+, \forall t \geq 0. \tag{3.4}$$

Let $\pi = (S_1, S_2, ...)$ denote a sampling policy and Π denote the set of *causal* sampling policies that satisfy two conditions: (i) Each sampling policy $\pi \in \Pi$ satisfies (4.87) for all i. (ii) The sequence of inter-sampling times $\{T_i = S_{i+1} - S_i, i = 0, 1, \ldots\}$ forms a *regenerative process* [34, Section IIB]: An increasing sequence $0 \leq l_1 < l_2 < \ldots$ of almost surely finite random integers exists such that the post-l_k process $\{T_{l_k+i}, i = 0, 1, \ldots\}$ is independent of the pre-l_k process $\{T_i, i = 0, 1, \ldots, l_k - 1\}$ and has same distribution as the post-l_0 process $\{T_{l_0+i}, i = 0, 1, \ldots\}$; We further assume that $\mathbb{E}[l_{k+1} - l_k] < \infty$, $\mathbb{E}[S_{l_1}] < \infty$, and $0 < \mathbb{E}[S_{l_{k+1}} - S_{l_k}] < \infty,\ k = 1, 2, \ldots$

3.2.3 MMSE Estimator

In this section, we provide the MMSE estimator for noisy samples of the OU process.

By using the expression of OU process for stable scenario [83, Eq. (3)] and the strong Markov property of the OU process [74, Eq. (4.3.27)], a solution to (3.1) for $t \in [S_i, \infty)$ given by the following three cases:

$$X_t = \begin{cases} X_{S_i} e^{-\theta(t-S_i)} + \mu\big[1 - e^{-\theta(t-S_i)}\big] + \frac{\sigma}{\sqrt{2\theta}} e^{-\theta(t-S_i)} W_{e^{2\theta(t-S_i)}-1}, & \text{if } \theta > 0, \\ \sigma W_t, & \text{if } \theta = 0, \\ X_{S_i} e^{-\theta(t-S_i)} + \mu\big[1 - e^{-\theta(t-S_i)}\big] + \frac{\sigma}{\sqrt{-2\theta}} e^{-\theta(t-S_i)} W_{1-e^{2\theta(t-S_i)}}, & \text{if } \theta < 0. \end{cases} \tag{3.5}$$

The estimator uses causally received samples to formulate an estimate $\hat{X}_t$ of the real-time signal value X_t at any time $t \geq 0$. The available information at the estimator has two parts: (i) $M_t = \{(S_i, Q_{S_i}, D_i) : D_i \leq t\}$, which contains the sampling time S_i, noisy sample value Q_{S_i}, and delivery time D_i of the samples that have been delivered by time t and (ii) no sample has been received after the last delivery time $\max\{D_i : D_i \leq t\}$. Similar to [30, 34, 55, 84], we assume that the estimator neglects the second part of the information. Then, as shown in [38], the MMSE estimator for $t \in [D_i, D_{i+1}),\ i = 0, 1, 2, \ldots$ for all of the cases in (3.5) is given as follows

$$\begin{aligned} \hat{X}_t =& \mathbb{E}[X_t | M_t] \\ =& Q_{S_i} e^{-\theta(t-S_i)} + \mu\big[1 - e^{-\theta(t-S_i)}\big]. \end{aligned} \tag{3.6}$$

3.3 Main Results

We evaluate the performance of remote estimation by the time-average mean square error which is expressed as follows:

$$\mathsf{mse} = \limsup_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\int_0^T (X_t - \hat{X}_t)^2 dt\right]. \tag{3.7}$$

A lower bound of (3.7) can be obtained when the additive noises are not considered ($N_{S_i} = 0, N'_{S_i} = 0$). On the other hand, an upper bound can be found by taking both the noises into account. Moreover, we formulate the following optimal sampling problem that minimizes the time-average mean-squared estimation error over an infinite time-horizon when no noise is considered.

$$\mathsf{mse}_{\text{opt-wn}} = \min_{\pi\in\Pi} \ \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T (X_t - \hat{X}_t)^2 dt\right], \tag{3.8}$$

where $\mathsf{mse}_{\text{opt-wn}}$ is the optimum value of (3.8) without noise.

3.3.1 Lower Bounds for mse

Let us consider a Gauss-Markov process with initial state $O_0 = 0$ and parameter $\mu = 0$, which can be expressed as

$$O_t = \begin{cases} \frac{\sigma}{\sqrt{2\theta}} e^{-\theta t} W_{e^{2\theta t}-1}, & \text{if } \theta > 0, \\ \sigma W_t, & \text{if } \theta = 0, \\ \frac{\sigma}{\sqrt{-2\theta}} e^{-\theta t} W_{1-e^{2\theta t}}, & \text{if } \theta < 0. \end{cases} \tag{3.9}$$

Before presenting the optimal sampler without noise, let us define the following parameter:

$$\mathsf{mse}_{Y_i} = \begin{cases} \frac{\sigma^2}{2\theta}\mathbb{E}[1 - e^{-2\theta Y_i}], & \text{if } \theta \neq 0, \\ \sigma^2 \mathbb{E}[Y_i], & \text{if } \theta = 0, \end{cases} \tag{3.10}$$

where mse_{Y_i} is the lower bound of mse. We will also need to use the following two functions

$$G(x) = \frac{\sqrt{\pi}}{2}\frac{e^{x^2}}{x}\operatorname{erf}(x), \ x \in [0, \infty), \tag{3.11}$$

$$K(x) = \frac{\sqrt{\pi}}{2}\frac{e^{-x^2}}{x}\operatorname{erfi}(x), \ x \in [0, \infty), \tag{3.12}$$

where if $x = 0$, both $G(x)$ and $K(x)$ are defined as their right limits $G(0) = \lim_{x\to 0^+} G(x) = 1$, and $K(0) = \lim_{x\to 0^+} K(x) = 1$. Furthermore, $\operatorname{erf}(\cdot)$ and $\operatorname{erfi}(\cdot)$ are the error function and imaginary error function respectively, defined as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2} dt, \operatorname{erfi}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{t^2} dt. \tag{3.13}$$

Note that $G(x)$ is strictly increasing on $x \in [0, \infty)$ [34], whereas $K(x)$ is strictly decreasing on $x \in [0, \infty)$. Hence, their inverses $G^{-1}(\cdot)$ and $K^{-1}(\cdot)$ are properly defined.

First, we consider that the system has no noise, i.e., $N_{S_i} = 0$ and $N'_{S_i} = 0$. Therefore, from (3.2) and (3.3), we get, $X_{S_i} = U_{S_i} = Q_{S_i}$. Then, the following theorem illustrates that the optimal sampling policy is a threshold policy and the threshold is found for all the three cases of the Gauss-Markov process parameter θ.

Theorem 3.1 *If the Y_i's are i.i.d. with $0 < \mathbb{E}[Y_i] < \infty$, then $(S_1(\beta), S_2(\beta), \ldots)$ with a parameter β is an optimal solution to (3.8), where*

$$S_{i+1}(\beta) = \inf\left\{t \geq D_i(\beta) : \left|X_t - \hat{X}_t\right| \geq v(\beta)\right\}, \tag{3.14}$$

$D_i(\beta) = S_i(\beta) + Y_i$, and $v(\beta)$ is given by

$$v(\beta) = \begin{cases} \frac{\sigma}{\sqrt{\theta}} G^{-1}\left(\frac{\frac{\sigma^2}{2\theta} - \mathsf{mse}_{Y_i}}{\frac{\sigma^2}{2\theta} - \beta}\right), & \text{if } \theta > 0, \\ \sqrt{3(\beta - \mathbb{E}[Y_i])}, & \text{if } \theta = 0, \\ \frac{\sigma}{\sqrt{-\theta}} K^{-1}\left(\frac{\frac{\sigma^2}{2\theta} - \mathsf{mse}_{Y_i}}{\frac{\sigma^2}{2\theta} - \beta}\right), & \text{if } \theta < 0, \end{cases} \tag{3.15}$$

where $G^{-1}(\cdot)$ is the inverse function of $G(\cdot)$ in (3.11), $K^{-1}(\cdot)$ is the inverse function of $K(\cdot)$ in (3.12), and β is the unique root of

$$\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} (X_t - \hat{X}_t)^2 dt\right] - \beta \mathbb{E}[D_{i+1}(\beta) - D_i(\beta)] = 0. \tag{3.16}$$

The optimal objective value to (3.8) is then given by

$$\mathsf{mse}_{\text{opt-wn}} = \frac{\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} (X_t - \hat{X}_t)^2 dt\right]}{\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)]}. \tag{3.17}$$

In [34], it is proved that the optimal sampling policy for stable OU process, i.e., when $\theta > 0$ is a threshold policy. The threshold obtained in [34] coincides with $v(\beta)$ in (3.15) for the case of $\theta > 0$. For $\theta = 0$, the threshold is obtained for $\sigma = 1$ which represents a Wiener process [35]. For $\theta < 0$, the proof procedure works in the same way as explained in [34] for stable OU processes. The threshold $v(\beta)$ is obtained by solving similar free boundary problems explained in [34] and the optimality of (4.57) for $\theta < 0$ is thus guaranteed. However, the threshold structure is different for all the three cases in Theorem 3.1. The function $K(x)$ in (3.12) is related to the function $G(x)$ in (3.11) as follows

$$K(x) = G(jx), \tag{3.18}$$

where j is the imaginary number represented by $j = \sqrt{-1}$. Therefore, the threshold $v(\beta)$ for $\theta < 0$ can be expressed by the following equation as well:

$$v(\beta) = j^{-1}\frac{\sigma}{\sqrt{-\theta}}G^{-1}\left(\frac{\frac{\sigma^2}{2\theta} - \mathsf{mse}_{Y_i}}{\frac{\sigma^2}{2\theta} - \beta}\right). \tag{3.19}$$

Though the threshold functions $v(\beta)$ varies with signal structure, the computation of the parameter β remains the same for all cases and the uniqueness of the root of (4.56) is proved in [34]. The decision of taking a new sample defined in (4.55) works in the same way as explained in [34].

3.3.2 Upper Bounds for mse

Suppose that the additive noise in the sampler and channel exist in the system, i.e., $N_{S_i} \neq 0, N'_{S_i} \neq 0$. Moreover, the sampler follows the sampling strategy obtained in (3.14). Because the noises N_{S_i} and N'_{S_i} are independent of the sampling times and the observed Gauss-Markov process, by utilizing (3.2), (3.3), (3.5), (3.6), and (3.17), the mse at the estimator which is an upper bound of (3.7) can be expressed as

$$\begin{aligned}
\mathsf{mse} =& \frac{\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}(X_t - \hat{X}_t)^2 dt\right]}{\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)]} \\
=& \frac{\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}(O_{t-S_i} - (N_{S_i} + N'_{S_i})e^{-\theta(t-S_i)})^2 dt\right]}{\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)]} \\
=& \frac{\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}O_{t-S_i}^2 dt\right]}{\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)]} + \frac{\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}(N_{S_i} + N'_{S_i})^2 e^{-2\theta(t-S_i)} dt\right]}{\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)]},
\end{aligned} \tag{3.20}$$

where (3.20) follows due to the fact that the Gauss-Markov process O_t has initial state $O_0 = 0$ and the noises N_{S_i} and N'_{S_i} with zero mean are independent of the observed Gauss-Markov process and sampling times.

To compute (3.20), the first fractional term remains the same as the $\mathsf{mse}_{\text{opt-wn}}$ in (4.57) with $N_{S_i} = 0$ and $N'_{S_i} = 0$. For stable OU processes, the associated $\mathsf{mse}_{\text{opt-wn}}$ is computed in [34, Lemma 1]. The expression of O^2_{t-Si} is the same for both stable and unstable OU processes. Therefore, the solution for (3.20) holds for all three cases in (3.9). For computing the second term, as N_{S_i} and N'_{S_i} are independent of the observed Gauss-Markov process and the sampling times, the numerator of the second fractional term in (3.20) can be written as:

$$\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} (N_{S_i} + N'_{S_i})^2 e^{-2\theta(t-S_i)} dt\right] = \mathbb{E}[(N_{S_i} + N'_{S_i})^2]\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} e^{-2\theta(t-S_i)} dt\right]. \quad (3.21)$$

Then, we have the following lemma for the last term in (3.21).

Lemma 3.1 *It holds that*

$$\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} e^{-2\theta(t-S_i)} dt\right] = \frac{1}{2\theta}\mathbb{E}\left[e^{-2\theta Y_i}\left\{1 - \min\left(1, \frac{{}_1F_1\left(1, \frac{1}{2}, \frac{\theta}{\sigma^2}O^2_{Y_i}\right)}{{}_1F_1\left(1, \frac{1}{2}, \frac{\theta}{\sigma^2}v^2(\beta)\right)}\right)\mathbb{E}[e^{-2\theta Y_{i+1}}]\right\}\right]. \quad (3.22)$$

Proof 3.1 *See Appendix 3.A.*

By using Lemma 3.1 and the expressions obtained in [34, Lemma 1], all the associated expectations in (3.20) can be obtained by Monte Carlo simulations of scalar random variables O_{Y_i} and Y_i, which does not require to directly simulate the entire random process $\{O_t, t \geq 0\}$.

3.4 Numerical Analysis

Figure 3.2 illustrates the MSE of *i.i.d* normalized log-normal service time, where $Y_i = e^{\alpha X_i}/\mathbb{E}[e^{\alpha X_i}]$, and $\alpha > 0$ is the scale parameter of log-normal distribution. The $(X_1, X_2, \dots)$ are *i.i.d.* Gaussian random variables, where $\mathbb{E}[X_i] = 0$ and $\text{Var}(X_i) = 1$. The maximum throughput of the queue is 1 as $\mathbb{E}[Y_i] = 1$. Both of the noises N_{S_i} and N'_{S_i} are considered to have 0 mean and variance 0.1. With the growth of the scale parameter α, the tail of

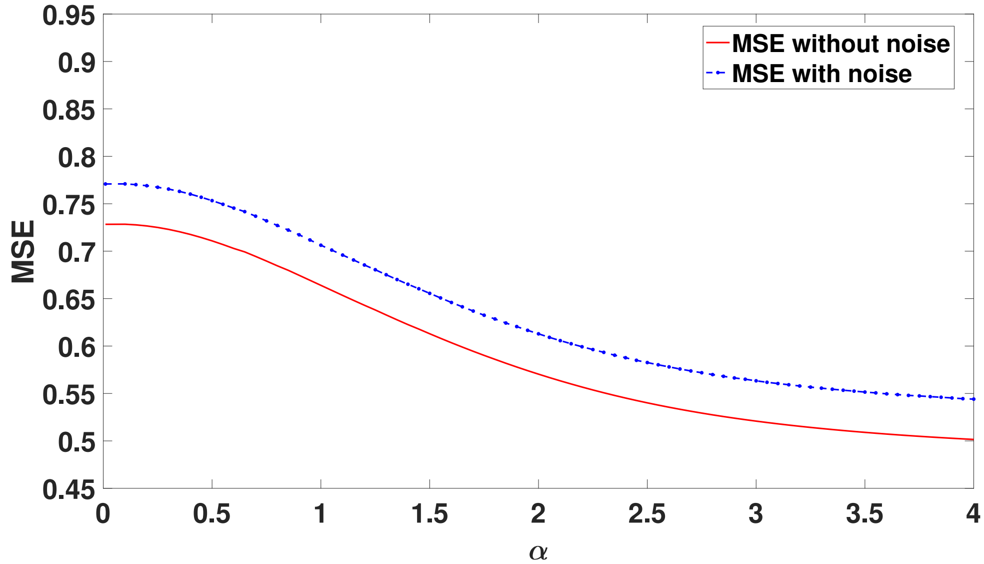


Figure 3.2: MSE vs. the scale parameter α of $i.i.d.$ normalized log-normal service time distribution with $\mathbb{E}[Y_i] = 1$, where the parameters of the Gauss-Markov process are $\sigma = 1$ and $\theta = 0.5$.

the log-normal distribution becomes heavier. The MSE with noise curve shows performance degradation as the additional term due to noise added with the mse without noise.

3.5 Conclusion

In this chapter, we have studied noisy samples over a noisy channel. We have obtained the performance upper and lower bounds of mse. The additional term added in the upper bound of mse due to noise is found. An optimal sampler design for noisy samples of Gauss-Markov processes will be another interesting future research direction.

Appendix

3.A Proof of Lemma 3.1

In order to prove Lemma 3.1, we need to consider the following two cases:

Case 1: If $|X_{D_i(\beta)} - \hat{X}_{D_i(\beta)}| = |O_{Y_i}| \geq v(\beta)$, then $S_{i+1}(\beta) = D_i(\beta)$. Hence,

$$D_i(\beta) = S_i(\beta) + Y_i, \tag{3.23}$$

$$D_{i+1}(\beta) = S_{i+1}(\beta) + Y_{i+1} = D_i(\beta) + Y_{i+1}. \tag{3.24}$$

Let us consider the following equation:

$$\begin{aligned}
&\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} e^{-2\theta(t-S_i)}dt \Big| O_{Y_i} = q, Y_i = y, |O_{Y_i}| \geq v(\beta)\right] \\
=&\mathbb{E}\left[\int_{Y_i}^{Y_i+Y_{i+1}} e^{-2\theta s}ds \Big| O_{Y_i} = q, Y_i = y, |O_{Y_i}| \geq v(\beta)\right] \\
=&\mathbb{E}\left[\frac{1}{2\theta}e^{-2\theta y}(1 - e^{-2\theta Y_{i+1}}) \Big| O_{Y_i} = q, Y_i = y, |O_{Y_i}| \geq v(\beta)\right] \\
=&\frac{1}{2\theta}e^{-2\theta y}\mathbb{E}\left[1 - e^{-2\theta Y_{i+1}} \Big| O_{Y_i} = q, Y_i = y, |O_{Y_i}| \geq v(\beta)\right] \\
=&\frac{1}{2\theta}e^{-2\theta y}\big\{1 - \mathbb{E}[e^{-2\theta Y_{i+1}}]\big\},
\end{aligned} \tag{3.25}$$

where (3.25) holds due to the fact that Y_{i+1} is independent of O_{Y_i} and Y_i.

Case 2: If $|X_{D_i(\beta)} - \hat{X}_{D_i(\beta)}| = |O_{Y_i}| < v(\beta)$, then

$$\begin{aligned}
&\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} e^{-2\theta(t-S_i)}dt \middle| O_{Y_i}=q, Y_i=y, |O_{Y_i}|<v(\beta)\right]\\
=&\mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i+Y_{i+1}} e^{-2\theta s}ds \middle| O_{Y_i}=q, Y_i=y, |O_{Y_i}|<v(\beta)\right]\\
=&\mathbb{E}\left[\frac{1}{2\theta}e^{-2\theta y}(1-e^{-2\theta Z_i}e^{-2\theta Y_{i+1}}) \middle| O_{Y_i}=q, Y_i=y\right]\\
=&\frac{1}{2\theta}e^{-2\theta y}\mathbb{E}\left[1-e^{-2\theta Z_i}e^{-2\theta Y_{i+1}} \middle| O_{Y_i}=q, Y_i=y\right]\\
=&\frac{1}{2\theta}e^{-2\theta y}\left\{1-\mathbb{E}\left[e^{-2\theta Z_i}\middle| O_{Y_i}=q, Y_i=y\right]\mathbb{E}[e^{-2\theta Y_{i+1}}]\right\}\\
=&\frac{1}{2\theta}e^{-2\theta y}\left\{1-\mathbb{E}[e^{-2\theta Z_i}|O_{Y_i}=q]\mathbb{E}[e^{-2\theta Y_{i+1}}]\right\},
\end{aligned} \tag{3.26}$$

where the last equation in (3.26) holds because Z_i is conditionally independent of Y_i given O_{Y_i}. Next, we need to compute $\mathbb{E}[e^{-2\theta Z_i}|O_{Y_i}=q]$, where Z_i is a hitting time of the time-shifted OU process O_{t+Y_i} given as

$$\begin{aligned}
&Z_i =\\
&\inf\{t: O_{t+Y_i} \notin (-v(\beta), v(\beta)) | O_{Y_i}=q \in (-v(\beta), v(\beta))\}.
\end{aligned} \tag{3.27}$$

By using the characteristic function of the hitting time of the OU process in [104, Eq. 15a], we get that

$$\mathbb{E}[e^{-2\theta Z_i}|O_{Y_i}=q] = \frac{{}_1F_1\left(1, \frac{1}{2}, \frac{\theta}{\sigma^2}q^2\right)}{{}_1F_1\left(1, \frac{1}{2}, \frac{\theta}{\sigma^2}v^2(\beta)\right)}. \tag{3.28}$$

Therefore, (3.26) becomes

$$\begin{aligned}
&\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} e^{-2\theta(t-S_i)}dt \middle| O_{Y_i}=q, Y_i=y, |O_{Y_i}|<v(\beta)\right]\\
=&\frac{1}{2\theta}e^{-2\theta y}\left\{1-\frac{{}_1F_1\left(1, \frac{1}{2}, \frac{\theta}{\sigma^2}q^2\right)}{{}_1F_1\left(1, \frac{1}{2}, \frac{\theta}{\sigma^2}v^2(\beta)\right)}\mathbb{E}[e^{-2\theta Y_{i+1}}]\right\}.
\end{aligned} \tag{3.29}$$

By combining (3.25) and (3.29), we get that

$$\begin{aligned}
&\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} e^{-2\theta(t-S_i)}dt \middle| O_{Y_i}=q, Y_i=y\right] \\
=&\frac{1}{2\theta}e^{-2\theta y}\left[1-\min\left\{1, \frac{{}_1F_1\left(1,\frac{1}{2},\frac{\theta}{\sigma^2}q^2\right)}{{}_1F_1\left(1,\frac{1}{2},\frac{\theta}{\sigma^2}v^2(\beta)\right)}\right\}\mathbb{E}[e^{-2\theta Y_{i+1}}]\right].
\end{aligned} \tag{3.30}$$

Finally, by taking the expectation over O_{Y_i} and Y_i in (3.30), Lemma 3.1 is proven.

Chapter 4

Remote Estimation of Multiple Gauss-Markov Processes over Multiple Channels

4.1 Introduction

Due to the prevalence of networked control and cyber-physical systems, real-time estimation of the states of remote systems has become increasingly important for next-generation networks. For instance, a timely and accurate estimate of the trajectories of nearby vehicles and pedestrians is imperative in autonomous driving, and real-time knowledge about the movements of surgical robots is essential for remote surgery. In these examples, real-time system state estimation is of paramount importance to the performance of these networked systems. Other notable applications of remote state estimation include UAV navigation, factory automation, environment monitoring, and augmented/virtual reality.

To assess the freshness of system state information, the metric AoI has drawn significant attention in recent years, e.g., [1], [2]. AoI is defined as the time difference between the current time and the generation time of the freshest received state sample. Besides AoI, nonlinear functions of the AoI have been introduced in [8], [64], [4] and illustrated to be useful as a metric of information freshness in sampling, estimation, and control [26], [4].

In many applications, the system state of interest is in the form of a signal X_t, which may vary quickly at time t and change slowly at a later time $t+\tau$ (even if the system state X_t is Markovian and time-homogeneous). AoI, as a metric of the time difference, cannot precisely characterize how fast the signal X_t varies at different time instants. To achieve more accurate system state estimation, it is important to consider *signal-aware remote estimation*, where the signal sampling and transmission scheduling decisions are made using the historical *realization* of the signal process X_t. Signal-aware remote estimation can achieve better performance than *AoI-based, signal-agnostic remote estimation*, where the sampling and scheduling decisions are made using the probabilistic distribution of the signal process X_t, and the mean-squared estimation error can be expressed as a function of the AoI. The connection between signal-aware remote estimation and AoI minimization was first revealed

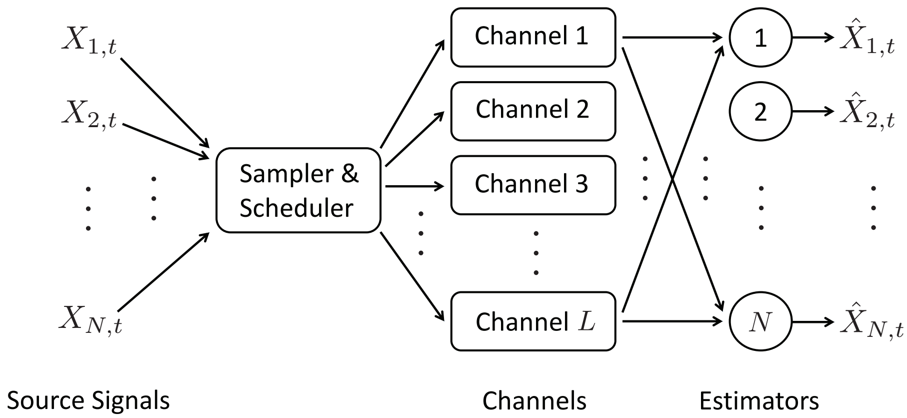


Figure 4.1: A multi-source, multi-channel remote estimation system.

in a problem of sampling a Wiener process [105]. Subsequently, it was generalized to the case of (stable) Ornstein-Uhlenbeck (OU) process in [34].

In many remote estimation and networked control systems, multiple sensors send their measurements (i.e., signal samples) to the destined estimators. For example, tire pressure, speed, and acceleration sensors in a self-driving vehicle send their data samples to the controller and nearby vehicles to make safe maneuvers [1]. In this paper, we consider a remote estimation system with N source-estimator pairs and L channels, as illustrated in Figure 2.1. Each source n is a continuous-time Gauss-Markov process $X_{n,t}$, defined as the solution of a Stochastic Differential Equation (SDE)

$$dX_{n,t} = \theta_n(\mu_n - X_{n,t})dt + \sigma_n dW_{n,t}, \tag{4.1}$$

where θ_n, μ_n, and $\sigma_n > 0$ are the parameters of the Gauss-Markov process, and the $W_{n,t}$'s are independent Wiener processes. If $\theta_n > 0$, $X_{n,t}$ is a stable Ornstein-Uhlenbeck (OU) process, which is the only nontrivial continuous-time process that is stationary, Gaussian, and Markovian [69]. If $\theta_n = 0$, then $X_{n,t} = \sigma_n W_{n,t}$ is a scaled Wiener process [101]. If $\theta_n < 0$, we call $X_{n,t}$ an unstable Ornstein-Uhlenbeck (OU) process, because $\lim_{t\to\infty} \mathbb{E}[X_{n,t}^2] = \infty$ in this case. These Gauss-Markov processes can be used to model random walks [106], interest rates [107], commodity prices [70], robotic swarms [72], biological processes [108], control systems (e.g., the transfer of liquids or gases in and out of a tank) [27], state exploration in deep reinforcement learning [109], and etc. A centralized sampler and scheduler decides when to take samples from the N Gauss-Markov processes and send the samples over L channels to remote estimators. At any time, at most L sources can send samples over the channels. The samples experience *i.i.d.* random transmission times over the channels due to interference, fading, etc. The n-th estimator uses causally received samples to reconstruct an estimate $\hat{X}_{n,t}$ of the real-time source value $X_{n,t}$.

Our objective is to find a sampling and transmission scheduling policy that minimizes the weighted sum of the time-average expected estimation errors of these Gauss-Markov sources. We develop a Whittle index policy to solve this problem. The technical contributions of this work are summarized as follows:

- We study the optimal sampling and transmission scheduling problem for the remote estimation of multiple continuous Gauss-Markov processes over parallel channels with *i.i.d.* random transmission times. This problem is a continuous-time Restless Multi-armed Bandit (RMAB) problem with a continuous state space, for which it is typically quite challenging to show indexability or to evaluate the Whittle index efficiently. We are able to prove indexability (see Theorem 4.1) and derive an exact expression for the Whittle index (Theorem 4.2 and Lemma 5.2). These results generalize prior studies on the remote estimation of a single Gauss-Markov process [34,55,110] to the multi-source, multi-channel case. To the best of our knowledge, such results for multi-source remote estimation of Gauss-Markov processes were unknown before. Among the technical tools used to prove these results are Shiryaev's free boundary method [74] for solving optimal stopping problems and Dynkin's formula [75] for evaluating expectations involving stopping times.

- We further investigate signal-agnostic remote estimation. In this context, the optimal sampling and scheduling problem becomes a multi-source AoI minimization problem over parallel channels with *i.i.d.* random transmission times. We establish the indexability property and derive a precise expression of the Whittle index (Theorems 4.3-4.4 and Lemma 4.3). Technically, these results carry forth and expand upon prior findings on Whittle index based AoI minimization [16, 41, 42] in the following manner: In [16, 41, 42], the transmission time remains constant, resulting in the optimality of the zero-wait sampling policy defined in [111, 112]. Consequently, the Whittle index derived in that case consistently maintains a non-negative value. In contrast, our results take into account scenarios involving *i.i.d.* random transmission times. In such instances, the optimality of the zero-wait sampling policy is not guaranteed, leading to the possibility of both positive and negative values for the Whittle index.

- Our results unite two important theoretical frameworks for remote estimation and AoI minimization: threshold-based sampling [4, 34, 55, 110] and Whittle index-based scheduling [16, 41, 42]. In the single-source, single-channel scenario, we demonstrate that the optimal solution to the sampling and scheduling problem can be expressed as both a threshold-based sampling strategy ([34, 55, 110]) and a Whittle index-based scheduling policy (see Theorems 4.5, 4.6). Particularly noteworthy is that the Whittle index is equal to zero at time t if and only if two conditions are satisfied: (i) the channel must be idle at time t, and (ii) the threshold condition is precisely met at time t. Moreover, the methodology used for deriving threshold-based sampling in the single-source, single-channel scenario plays a pivotal role in establishing indexability and evaluating the Whittle index in the more complex multi-source, multi-channel scenario.

- Our numerical results show that the proposed policy performs better than the signal-agnostic AoI-based Whittle index policy and the Maximum-Age-First, Zero-Wait (MAF-ZW) policy. The performance gain of the proposed policy is high when some of the Gauss-Markov processes are highly unstable.

4.2 Model

This section describes the multi-source, multi-channel system depicted in Figure 4.1

4.2.1 System Model

Consider a remote estimation system with N source-estimator pairs and L channels, which is shown in Figure 4.1. Each source n is a continuous-time Gauss-Markov process $X_{n,t}$, as defined in (4.1). The sources are independent of each other and the parameters θ_n,

μ_n, and σ_n may vary across the sources. Hence, the N sources could be a mixing of scaled Wiener processes, stable OU processes, and unstable OU processes. A centralized sampler and transmission scheduler chooses when to take samples from the sources and transmit the samples over the channels to the associated remote estimators. At any given time, each source can be served by no more than one channel. In other words, if there are multiple samples from the same source waiting to be transmitted, only one of these samples can be transmitted over a single channel simultaneously. Sample transmissions are *non-preemptive*, i.e., once a channel starts to send a sample, it must finish transmitting the current sample before switching to serve another sample. Whenever a sample is delivered to the associated estimator, an acknowledgment (ACK) is immediately sent back to the scheduler.

The operation of the system starts at time $t = 0$. Let $S_{n,i}$ be the generation time of the i-th sample of source n, which satisfies $S_{n,i} \leq S_{n,i+1}$. This sample is submitted to a channel at time $G_{n,i}$, undergoes a random transmission time $Y_{n,i}$, and is delivered to the estimator n at time $D_{n,i}$, where $S_{n,i} \leq G_{n,i}$, and $G_{n,i} + Y_{n,i} = D_{n,i}$. Because (i) each source can be served by at most one channel at a time and (ii) the sample transmissions are non-preemptive, $D_{n,i} \leq G_{n,i+1}$. The sample transmission times $Y_{n,i}$'s are *i.i.d.* across samples and channels with mean $0 < \mathbb{E}[Y_{n,i}] < \infty$. In addition, we assume that the $Y_{n,i}$'s are independent of the Gauss-Markov processes $X_{n,t}$. The i-th sample packet $(S_{n,i}, X_{n,S_{n,i}})$ contains the sample value $X_{n,S_{n,i}}$ and its sampling time $S_{n,i}$. Let $U_n(t) = \max_i\{S_{n,i} : D_{n,i} \leq t, i = 1, 2, \ldots\}$ be the generation time of the freshest received sample from source n at time t. The AoI of source n at time t is defined as [1, 2]

$$\Delta_n(t) = t - U_n(t) = t - \max_i\{S_{n,i} : D_{n,i} \leq t, i = 1, 2, \ldots\}. \tag{4.2}$$

Because $D_{n,i} \leq D_{n,i+1}$, $\Delta_n(t)$ can also be expressed as

$$\Delta_n(t) = t - S_{n,i}, \text{if}\, t \in [D_{n,i}, D_{n,i+1}),\ i = 0, 1, \ldots. \tag{4.3}$$

At time $t = 0$, the initial state of the system satisfies $S_{n,0} = 0$, and $D_{n,0} = Y_{n,0}$. The initial value of the Gauss-Markov process $X_{n,0}$ is finite.

4.2.2 MMSE Estimator

At any time $t \geq S_{n,i}$, the Gauss-Markov process $X_{n,t}$ can be expressed as

$$X_{n,t} = \begin{cases} X_{n,S_{n,i}} e^{-\theta_n(t-S_{n,i})} + \mu_n\left[1 - e^{-\theta_n(t-S_{n,i})}\right] + \frac{\sigma_n}{\sqrt{2\theta_n}} W_{n,1-e^{-2\theta_n(t-S_{n,i})}}, & \text{if } \theta_n > 0, \\ \sigma_n W_{n,t}, & \text{if } \theta_n = 0, \\ X_{n,S_{n,i}} e^{-\theta_n(t-S_{n,i})} + \mu_n\left[1 - e^{-\theta_n(t-S_{n,i})}\right] + \frac{\sigma_n}{\sqrt{-2\theta_n}} W_{n,e^{-2\theta_n(t-S_{n,i})}-1}, & \text{if } \theta_n < 0, \end{cases} \quad (4.4)$$

where three expressions are provided for stable OU process ($\theta_n > 0$), scaled Wiener process ($\theta_n = 0$), and unstable OU process ($\theta_n < 0$), respectively. The first two expressions in (4.4) for the stable OU process and the scaled Wiener process were provided in [83]. The third expression in (4.4) for the unstable OU process is proven in Appendix 4.A.

At time t, each estimator n utilizes causally received samples to construct an estimate $\hat{X}_{n,t}$ of the signal value $X_{n,t}$. The information that is available at the estimator contains two parts: (i) $M_{n,t} = \{(S_{n,i}, X_{n,S_{n,i}}, G_{n,i}, D_{n,i}) : D_{n,i} \leq t, i = 1, 2, \ldots\}$, which contains the sampling time $S_{n,i}$, sample value $X_{n,S_{n,i}}$, transmission starting time $G_{n,i}$, and the delivery time $D_{n,i}$ of the samples up to time t and (ii) no sample has been received after the last delivery time $\max_i\{D_{n,i} : D_{n,i} \leq t, i = 1, 2, \ldots\}$. Similar to [30, 34, 55, 84], we assume that the estimator neglects the second part of the information. If $t \in [D_{n,i}, D_{n,i+1})$, the MMSE estimator is given by [34, 110]

$$\begin{aligned} \hat{X}_{n,t} &= \mathbb{E}[X_{n,t} | M_{n,t}] \\ &= \begin{cases} X_{n,S_{n,i}} e^{-\theta_n(t-S_{n,i})} + \mu_n\left[1 - e^{-\theta_n(t-S_{n,i})}\right], & \text{if } \theta_n \neq 0, \\ \sigma_n W_{n,S_{n,i}}, & \text{if } \theta_n = 0. \end{cases} \end{aligned} \quad (4.5)$$

The estimation error $\varepsilon_n(t)$ of source n at time t is given by

$$\varepsilon_n(t) = X_{n,t} - \hat{X}_{n,t}. \quad (4.6)$$

By substituting (4.4) and (4.5) into (4.6), if $t \in [D_{n,i}, D_{n,i+1})$, then

$$\varepsilon_n(t) = \begin{cases} \frac{\sigma_n}{\sqrt{2\theta_n}} W_{n,1-e^{-2\theta_n(t-S_{n,i})}}, & \text{if } \theta_n > 0, \\ \sigma_n(W_{n,t} - W_{n,S_{n,i}}), & \text{if } \theta_n = 0, \\ \frac{\sigma_n}{\sqrt{-2\theta_n}} W_{n,e^{-2\theta_n(t-S_{n,i})}-1}, & \text{if } \theta_n < 0. \end{cases} \tag{4.7}$$

4.3 Problem Formulation

Let $\pi = (\pi_n)_{n=1}^N$ denote a sampling and scheduling policy, where $\pi_n = ((S_{n,1}, G_{n,1}), (S_{n,2}, G_{n,2}), \ldots)$ contains the sampling and transmission starting time instants of source n. Let π_n denote a sub-sampling and scheduling policy for source n. In *causal* sampling and scheduling policies, each sampling time $S_{n,i}$ is determined based on the up-to-date information that is available at the scheduler, without using any future information. Let Π denote the set of all causal sampling and scheduling policies and let Π_n denote the set of causal sub-sampling and scheduling policies for source n, both of which satisfy (i) each source can be served by at most one channel at a time, and (ii) the sample transmissions are non-preemptive. At any time t, $c_n(t) \in \{0, 1\}$ denotes the channel occupation status of source n. If source n is being served by a channel at time t, then $c_n(t) = 1$; otherwise, $c_n(t) = 0$. Hence, if $t \in [G_{n,i}, D_{n,i})$, then $c_n(t) = 1$. Because there are L channels, $\sum_{n=1}^N c_n(t) \le L$ is required to hold for all $t \ge 0$.

Our objective is to find a causal sampling and scheduling policy for minimizing the weighted sum of the time-average expected estimation errors of the N Gauss-Markov sources. This sampling and scheduling problem is formulated as

$$\inf_{\pi \in \Pi} \quad \limsup_{T \to \infty} \sum_{n=1}^N w_n \mathbb{E}_\pi \left[\frac{1}{T} \int_0^T \varepsilon_n^2(t) dt \right] \tag{4.8}$$

$$\text{s.t.} \quad \sum_{n=1}^N c_n(t) \le L,\ c_n(t) \in \{0, 1\}, n = 1, 2, \ldots, N, t \in [0, \infty), \tag{4.9}$$

where $w_n > 0$ is the weight of source n. The sampling and scheduling policy π can be simplified by simplifying the sub-sampling and scheduling policy π_n. In Appendix 4.B, we prove that in the optimal policies to (4.8)-(4.9), the sampling time of the i-th sample $S_{n,i}$ and the transmission starting time of the i-th sample $G_{n,i}$ are equal to each other, i.e., $S_{n,i} = G_{n,i}$. Therefore, each sub-policy π_n in π can be simply denoted as $\pi_n = (S_{n,1}, S_{n,2}, \ldots)$.

4.4 Signal-aware Scheduling

Problem (4.8)-(4.9) is a continuous-time Restless Multi-armed Bandit (RMAB) with a continuous state space, where the estimation error $\varepsilon_n(t)$ of source n is the state of the n-th restless bandit and each restless bandit is a Markov Decision Process (MDP) with two actions: active and passive. If a sample of source n is taken and submitted to a channel at time t, we say that bandit n takes an active action at time t; otherwise, bandit n is made passive at time t. If a sample of source n is in service, only the passive action is available for source n.

An efficient approach for solving RMABs is to develop a low-complexity scheduling algorithm by leveraging the Whittle index theory [113, 114]. If all the bandits are indexable and certain technical conditions are satisfied, the Whittle index policy is asymptotically optimal as the number of bandits N and the number of channels L increases to infinity, keeping the ratio L/N constant [113]. In this section, we develop a Whittle index policy for solving problem (8)-(9) in three steps: (i) first, we relax the constraint (4.9) and utilize a Lagrangian dual approach to decompose the original problem into separated per-bandit problems; (ii) next, we prove that the per-bandit problems are indexable; and (iii) finally, we derive closed-form expressions for the Whittle index. Because the RMAB in (4.8)-(4.9) has a continuous state space and requires continuous-time control, demonstrating indexability in Step (ii) and efficiently evaluating the Whittle index in Step (iii) are technically challenging. However, we are able to overcome these challenges.

4.4.1 Restless Multi-armed Bandit: Relaxation and Lagrangian Decomposition

In standard restless multi-armed bandit problems, the channel resource constraint needs to be satisfied with equality. In this paper, we consider a scenario where less than L bandits can be activated at any time t, as indicated by constraint (4.9). Following [115, Section 5.1.1], we introduce L additional *dummy bandits* that will never change state and hence their estimation errors are 0 (i.e., $\varepsilon_0(t) = 0$). When a *dummy bandit* is activated, it occupies one channel, but it does not incur any estimation error. Let $c_0(t) \in \{0, 1, 2, \ldots, L\}$ denotes the number of *dummy bandits* that are activated at time t. By considering *dummy bandits*, the RMAB (4.8)-(4.9) is equivalent to

$$\inf_{\pi\in\Pi} \quad \limsup_{T\to\infty} \sum_{n=1}^{N} w_n \mathbb{E}_\pi \left[\frac{1}{T}\int_0^T \varepsilon_n^2(t)dt\right] \tag{4.10}$$

$$\text{s.t.} \quad \sum_{n=0}^{N} c_n(t) = L, c_0(t) \in \{0, 1, \ldots, L\}, t \in [0, \infty),$$

$$c_n(t) \in \{0, 1\}, n = 1, 2, \ldots, N, t \in [0, \infty), \tag{4.11}$$

which is an RMAB with an equality constraint.

Following the standard relaxation and Lagrangian dual decomposition procedure in the Whittle index theory [114], we relax the constraint (4.11) as

$$\limsup_{T\to\infty} \sum_{n=0}^{N} \mathbb{E}_\pi \left[\frac{1}{T}\int_0^T c_n(t)dt\right] = L. \tag{4.12}$$

The relaxed constraint (4.12) only needs to be satisfied on average, whereas (4.11) is required to hold at any time t. Then, the RMAB (4.10)-(4.11) is reformulated as

$$\inf_{\pi\in\Pi} \limsup_{T\to\infty} \sum_{n=1}^{N} w_n \mathbb{E}_\pi \left[\frac{1}{T}\int_0^T \varepsilon_n^2(t)dt\right] \tag{4.13}$$

$$\text{s.t.} \limsup_{T\to\infty} \sum_{n=0}^{N} \mathbb{E}_\pi \left[\frac{1}{T}\int_0^T c_n(t)dt\right] = L,$$

$$c_0(t) \in \{0, 1, \ldots, L\}, c_n(t) \in \{0, 1\}, n = 1, 2, \ldots, N, t \in [0, \infty). \tag{4.14}$$

Next, we take the Lagrangian dual decomposition of the relaxed problem (4.13)-(4.14), which produces the following problem with a dual variable $\lambda \in \mathbb{R}$, also known as the activation cost [114]:

$$\inf_{\pi\in\Pi} \limsup_{T\to\infty} \mathbb{E}_\pi \left[\frac{1}{T}\int_0^T \sum_{n=1}^{N} w_n \varepsilon_n^2(t) + \lambda\left(\sum_{n=0}^{N} c_n(t) - L\right)dt\right]. \tag{4.15}$$

The term $\frac{1}{T}\int_0^T \sum_{n=0}^{N} \lambda L dt$ in (4.15) does not depend on policy π and hence can be removed. Then, Problem (4.15) can be decomposed into $(N+1)$ separated sub-problems. The sub-problem associated with source n is

$$\bar{m}_{n,\text{opt}} = \inf_{\pi_n \in \Pi_n} \limsup_{T\to\infty} \mathbb{E}_{\pi_n}\left[\frac{1}{T}\int_0^T w_n \varepsilon_n^2(t) + \lambda c_n(t) dt\right], \tag{4.16}$$

where $\bar{m}_{n,\text{opt}}$ is the optimum value of (4.16) and $n = 1, 2, \ldots, N$. On the other hand, the sub-problem associated with the *dummy bandits* is given by

$$\inf_{\pi_0 \in \Pi_0} \limsup_{T\to\infty} \mathbb{E}_{\pi_n}\left[\frac{1}{T}\int_0^T \lambda c_0(t) dt\right], \tag{4.17}$$

where $\pi_0 = \{c_0(t), t \in [0, \infty)\}$ and Π_0 is the set of all causal activation policies π_0.

4.4.2 Indexability

We now establish the indexability of the RMAB in (4.10)-(4.11). Let $\gamma_n(t) \in [0, \infty)$ denote the amount of time that has been used to send the current sample of source n at time t. Here, if no sample from source n is currently in service at time t, then $\gamma_n(t) = 0$; if a sample from source n is currently in service at time t, then $\gamma_n(t) > 0$. Consequently, if $\gamma_n(t) > 0$, the active action is not available for source n at time t.

Let $\Psi_n(\lambda)$ be a set of states $(\varepsilon, \gamma) \in \mathbb{R} \times [0, \infty)$ such that if $\varepsilon_n(t) = \varepsilon$ and $\gamma_n(t) = \gamma$, the optimal solution for (4.16) (or (4.17) when $n = 0$) is to take a passive action at time t.

Definition 4.1 ***(Indexability).*** *[115] Bandit n is said to be* indexable *if, as the activation cost λ increases from $-\infty$ to ∞, the set $\Psi_n(\lambda)$ increases monotonically, i.e., $\lambda_1 \leq \lambda_2$ implies $\Psi_n(\lambda_1) \subseteq \Psi_n(\lambda_2)$. The RMAB* (4.10)-(4.11) *is said to be* indexable *if all* $(N+1)$ *bandits are indexable.*

In general, establishing the indexability of an RMAB can be a challenging task. Because the per-bandit problem (4.16) is a continuous-time MDP with a continuous state space, determining the indexability of (4.16) appears to be quite formidable. In the sequel, we will utilize the techniques developed in our previous work [34] to solve (4.16) precisely and analytically characterize the set $\Psi_n(\lambda)$. This analysis will allow us to demonstrate that (4.16) is indeed indexable.

Define

$$G(x) = \frac{\sqrt{\pi}}{2}\frac{e^{x^2}}{x}\text{erf}(x), \tag{4.18}$$

$$K(x) = \frac{\sqrt{\pi}}{2}\frac{e^{-x^2}}{x}\text{erfi}(x), \tag{4.19}$$

where erf(x) and erfi(x) are the error function and imaginary error function, respectively, determined by [85, Sec. 8.25]

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt, \tag{4.20}$$

$$\text{erfi}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{t^2}dt. \tag{4.21}$$

If $x = 0$, both $G(x)$ and $K(x)$ are defined as their limits $G(0) = \lim_{x\to 0} G(x) = 1$ and $K(0) = \lim_{x\to 0} K(x) = 1$, respectively. Both $G(\cdot)$ and $K(\cdot)$ are even functions. The function $G(x)$ is strictly increasing on $x \in [0, \infty)$ and $G(0) = 1$ [34]. On the other hand, $K(x)$ is strictly decreasing on $x \in [0, \infty)$ and $K(0) = 1$ [110]. Hence, the inverse functions of $G(x)$ and $K(x)$ are well defined on $x \in [0, \infty)$. The relation between these two functions is given by [110]

$$K(x) = G(jx), \tag{4.22}$$

where $j = \sqrt{-1}$ is the unit imaginary number.

Proposition 4.1 *If the $Y_{n,i}$'s are i.i.d. with $0 < \mathbb{E}[Y_{n,i}] < \infty$, then $(S_{n,1}(\beta_n), S_{n,2}(\beta_n), \ldots)$ with a parameter β_n is an optimal solution to* (4.16)*, where*

$$S_{n,i+1}(\beta_n) = \inf_t \left\{t \geq D_{n,i}(\beta_n) : |\varepsilon_n(t)| \geq v_n(\beta_n)\right\}, \tag{4.23}$$

$D_{n,i}(\beta_n) = S_{n,i}(\beta_n) + Y_{n,i}$, $v_n(\beta_n)$ *is defined by*

$$v_n(\beta_n) = \begin{cases} \frac{\sigma_n}{\sqrt{\theta_n}} G^{-1}\left(\frac{w_n \frac{\sigma_n^2}{2\theta_n}\mathbb{E}[e^{-2\theta_n Y_{n,i}}]}{w_n \frac{\sigma_n^2}{2\theta_n} - \beta_n}\right), & \text{if } \theta_n > 0, \\ \frac{1}{\sqrt{w_n}} \sqrt{3(\beta_n - w_n\sigma_n^2\mathbb{E}[Y_{n,i}])}, & \text{if } \theta_n = 0, \\ \frac{\sigma_n}{\sqrt{-\theta_n}} K^{-1}\left(\frac{w_n \frac{\sigma_n^2}{2\theta_n}\mathbb{E}[e^{-2\theta_n Y_{n,i}}]}{w_n \frac{\sigma_n^2}{2\theta_n} - \beta_n}\right), & \text{if } \theta_n < 0, \end{cases} \tag{4.24}$$

$G^{-1}(\cdot)$ *and* $K^{-1}(\cdot)$ *are the inverse functions of* $G(x)$ *in* (4.18) *and* $K(x)$ *in* (4.19), *respectively, defined in the region of* $x \in [0, \infty)$, *and* β_n *is the unique root of*

$$\mathbb{E}\left[\int_{D_{n,i}(\beta_n)}^{D_{n,i+1}(\beta_n)} w_n \varepsilon_n^2(t)dt\right] - \beta_n \mathbb{E}[D_{n,i+1}(\beta_n) - D_{n,i}(\beta_n)] + \lambda \mathbb{E}[Y_{n,i+1}] = 0. \tag{4.25}$$

The optimal objective value to (5.11) *is given by*

$$\bar{m}_{n,\text{opt}} = \frac{\mathbb{E}\left[\int_{D_{n,i}(\beta_n)}^{D_{n,i+1}(\beta_n)} w_n \varepsilon_n^2(t)dt\right] + \lambda \mathbb{E}[Y_{n,i+1}]}{\mathbb{E}[D_{n,i+1}(\beta_n) - D_{n,i}(\beta_n)]}. \tag{4.26}$$

Furthermore, β_n *is exactly the optimal objective value of* (4.16), *i.e.,* $\beta_n = \bar{m}_{n,\text{opt}}$.

Proof 4.1 *Appendix 4.C.*

Proposition 4.1 complements earlier optimal sampling results for the remote estimation of the Wiener process (i.e., the case of $\theta_n = 0$ and $\lambda = 0$) [35] and stable OU process (i.e., $\theta_n > 0$ and $\lambda = 0$) [34], by (i) adding a third case on unstable OU process (i.e., $\theta_n < 0$) and (ii) incorporating an activation cost $\lambda \in \mathbb{R}$.

By using Proposition 4.1, we can analytically characterize the set $\Psi_n(\lambda)$. To that end, we first show that the threshold $v_n(\beta_n)$ in (4.23) is a function of the activation cost λ. For any given λ, β_n is the unique root of equation (4.25). Hence, β_n can be expressed as an implicit function $\beta_n(\lambda)$ of λ, defined by equation (4.25). Moreover, the threshold $v_n(\beta_n)$ can be rewritten as a function $v_n(\beta_n(\lambda))$ of the activation cost λ. According to (4.23) and the definition of set $\Psi_n(\lambda)$, a point $(\varepsilon_n(t), \gamma_n(t)) \in \Psi_n(\lambda)$ if either (i) $\gamma_n(t) > 0$ such that a sample from source n is currently in service at time t, or (ii) $|\varepsilon_n(t)| < v_n(\beta_n(\lambda))$ such that the threshold condition in (4.23) for taking a new sample is not satisfied. By this, an analytical expression of set $\Psi_n(\lambda)$ is derived as

$$\Psi_n(\lambda) = \{(\varepsilon, \gamma) \in \mathbb{R} \times [0, \infty) : \gamma > 0 \,\text{or}\, |\varepsilon| < v_n(\beta_n(\lambda))\}. \tag{4.27}$$

Using (4.27), we can prove the first key result of the present paper:

Theorem 4.1 *The RMAB problem* (4.10)-(4.11) *is indexable.*

Proof sketch. According to Proposition 4.1, for any λ, the optimal solution to (4.16) is a threshold policy. Using this, we can show that the unique root $\beta_n(\lambda)$ of (4.25) is a strictly increasing function of λ. In addition, $v_n(\beta_n)$ in (4.24) is a strictly increasing function of β_n. Hence, $v_n(\beta_n(\lambda))$ is a strictly increasing function of λ. Substituting this into (4.27), if $\lambda_1 \leq \lambda_2$, then $\Psi_n(\lambda_1) \subseteq \Psi_n(\lambda_2)$. For the *dummy bandits*, it is optimal in (4.17) to activate a bandit when $\lambda < 0$. Hence, *dummy bandits* are always indexable. The details are provided in Appendix 4.D. □

4.4.3 Whittle Index Policy

Next, we introduce the definition of the Whittle index.

Definition 4.2 *[114] If bandit n is indexable, then the Whittle index $W_n(\varepsilon, \gamma)$ of bandit n at state (ε, γ) is defined by*

$$W_n(\varepsilon, \gamma) = \inf_{\lambda}\{\lambda \in \mathbb{R} : (\varepsilon, \gamma) \in \Psi_n(\lambda)\}, \tag{4.28}$$

which is the infimum of the activation cost λ for which it is better not to activate bandit n.

Theorem 4.2 *The following assertions are true for the Whittle index $W_n(\varepsilon, \gamma)$ of problem* (4.16) *at state (ε, γ):*

(a) If $\gamma = 0$, then the Whittle index $W_n(\varepsilon, \gamma)$ is presented in the following three cases:

(i) Case 1: *If $\theta_n > 0$ (i.e., $X_{n,t}$ is a stable OU process), then*

$$W_n(\varepsilon, 0) =$$
$$\frac{w_n}{\mathbb{E}[Y_{n,i}]}\left\{\mathbb{E}[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)]\frac{\sigma_n^2}{2\theta_n}\left(1 - \frac{\mathbb{E}[e^{-2\theta_n Y_{n,i}}]}{G\left(\frac{\sqrt{\theta_n}}{\sigma_n}\varepsilon\right)}\right) - \mathbb{E}\left[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds\right]\right\}, \tag{4.29}$$

(ii) Case 2: *If* $\theta_n = 0$ *(i.e.,* $X_{n,t}$ *is a scaled Wiener process), then*

$$W_n(\varepsilon, 0) = \frac{w_n}{\mathbb{E}[Y_{n,i}]}\left\{\mathbb{E}[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)]\left(\frac{\varepsilon^2}{3} + \sigma_n^2 \mathbb{E}[Y_{n,i}]\right) - \mathbb{E}\left[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds\right]\right\}, \tag{4.30}$$

(iii) Case 3: *If* $\theta_n < 0$ *(i.e.,* $X_{n,t}$ *is an unstable OU process), then*

$$W_n(\varepsilon, 0) = \frac{w_n}{\mathbb{E}[Y_{n,i}]}\left\{\mathbb{E}[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)]\frac{\sigma_n^2}{2\theta_n}\left(1 - \frac{\mathbb{E}[e^{-2\theta_n Y_{n,i}}]}{K\left(\frac{\sqrt{-\theta_n}}{\sigma_n}\varepsilon\right)}\right) - \mathbb{E}\left[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds\right]\right\}, \tag{4.31}$$

where $G(\cdot)$ *and* $K(\cdot)$ *are defined in* (4.18) *and* (4.19)*, respectively.*

(b) If $\gamma > 0$*, then*

$$W_n(\varepsilon, \gamma) = -\infty. \tag{4.32}$$

Proof sketch. When $\gamma = 0$, by (4.27), (4.28), and the monotonicity of $v_n(\cdot)$ and $\beta_n(\cdot)$, the Whittle index $W_n(\varepsilon, 0)$ is equal to the unique root λ of equation

$$|\varepsilon| = v_n(\beta_n(\lambda)). \tag{4.33}$$

Hence, $W_n(\varepsilon, 0) = \beta_n^{-1}(v_n^{-1}(|\varepsilon|)$. By substituting (4.24) and (4.25) into (4.33) and using the fact that $G(\cdot)$ and $K(\cdot)$ are even functions, statement (a) in Theorem 4.2 is proven. When $\gamma > 0$, (ε, γ) is always in the set $\Psi_n(\lambda)$ for any $\lambda \in \mathbb{R}$. Hence, by using (4.28), $W_n(\varepsilon, \lambda) = -\infty$. By this, statement (b) in Theorem 4.2 is proven. The details are provided in Appendix 4.E. □

In Theorem 4.2, the delivery time $D_{n,i}(\varepsilon)$ is expressed as a function of ε for the following reason: in the optimal solution to (4.16), the sample delivery time is a function of the activation cost λ. If the activation cost λ in (4.16) is chosen as $\lambda = W_n(\varepsilon, \gamma)$, then the sample delivery time in the optimal solution to (4.16) is a function of ε. We use the notation

$D_{n,i}(\varepsilon)$ to remind us that the expectations $\mathbb{E}[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)]$ and $\mathbb{E}[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds]$ in (4.29)-(4.31) change as ε varies.

In order to compute the Whittle index $W_n(\varepsilon, \gamma)$, we need to calculate the expectations $\mathbb{E}[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)]$ and $\mathbb{E}[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds]$ in (4.29)-(4.31). Because $S_{n,i}(\varepsilon)$ and $D_{n,i}(\varepsilon)$ are stopping times of the process $X_{n,t}$, numerically evaluating these two expectations is nontrivial. This challenge can be addressed by resorting to Lemma 5.2 provided below, which is obtained by using Dynkin's formula [75, Theorem 7.4.1] to simplify expectations involving stopping times.

To that end, let us introduce a Gauss-Markov process $O_{n,t}$ with a zero initial condition $O_{n,0} = 0$ and parameter $\mu_n = 0$, which is expressed as

$$O_{n,t} = \begin{cases} \frac{\sigma_n}{\sqrt{2\theta_n}} W_{n,1-e^{-2\theta_n t}}, & \text{if } \theta_n > 0, \\ \sigma_n W_{n,t}, & \text{if } \theta_n = 0, \\ \frac{\sigma_n}{\sqrt{-2\theta_n}} W_{n,e^{-2\theta_n t}-1}, & \text{if } \theta_n < 0. \end{cases} \tag{4.34}$$

By comparing (4.7) with (4.34), the estimation error process $\varepsilon_n(t)$ has the same distribution with as the time-shifted Gauss-Markov process $O_{n,t-S_{n,i}(\varepsilon)}$, where $t \in [D_{n,i}(\varepsilon), D_{n,i+1}(\varepsilon))$.

Then, we have the following lemma for computing the expectations in (4.29), (4.30), and (4.31).

Lemma 4.1 *In Theorem 4.2, it holds that*

$$\mathbb{E}[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)] = \mathbb{E}\big[R_{n,1}\big(\max\{|\varepsilon|, |O_{n,Y_{n,i}}|\}\big)\big], \tag{4.35}$$

$$\mathbb{E}\left[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds\right] = \mathbb{E}\big[R_{n,2}\big(\max\{|\varepsilon|, |O_{n,Y_{n,i}}|\} + O_{n,Y_{n,i+1}}\big)\big] - \mathbb{E}\big[R_{n,2}\big(O_{n,Y_{n,i}}\big)\big], \tag{4.36}$$

where if $\theta_n \neq 0$, then

$$R_{n,1}(\varepsilon) = \frac{\varepsilon^2}{\sigma_n^2}\, {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta_n}{\sigma_n^2}\varepsilon^2\right), \tag{4.37}$$

$$R_{n,2}(\varepsilon) = -\frac{\varepsilon^2}{2\theta_n} + \frac{\varepsilon^2}{2\theta_n}\, {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta_n}{\sigma_n^2}\varepsilon^2\right); \tag{4.38}$$

if $\theta_n = 0$, *then*

$$R_{n,1}(\varepsilon) = \frac{\varepsilon^2}{\sigma_n^2}, \tag{4.39}$$

$$R_{n,2}(\varepsilon) = \frac{\varepsilon^4}{6\sigma_n^2}. \tag{4.40}$$

Proof 4.2 *See Appendix 5.A.*

In (4.37) and (4.38), we have used the generalized hypergeometric function, which is defined by [86, Eq. 16.2.1]

$${}_pF_q(a_1, a_2, \cdots, a_p; b_1, b_2, \cdots b_q; z) = \sum_{n=0}^{\infty} \frac{(a_1)_n (a_2)_n \cdots (a_p)_n}{(b_1)_n (b_2)_n \cdots (b_p)_n} \frac{z^n}{n!}, \tag{4.41}$$

where

$$(a)_0 = 1, \tag{4.42}$$

$$(a)_n = a(a+1)(a+2)\cdots(a+n-1), \ n \geq 1. \tag{4.43}$$

Lemma 5.2 is more general than Lemma 1 in [34], because Lemma 5.2 holds for all three cases of the Gauss-Markov processes, i.e., $\theta_n > 0$, $\theta_n = 0$, and $\theta_n < 0$, whereas Lemma 1 in [34] was only shown for $\theta_n > 0$. Moreover, (4.35)-(4.36) in Lemma 5.2 are neater than (22)-(23) in Lemma 1 of [34].

The expectations in (4.35) and (4.36) can be evaluated by Monte-Carlo simulations of scalar random variables $O_{n,Y_{n,i}}$ and $O_{n,Y_{n,i+1}}$ which is much easier than directly simulating the entire process $\{\varepsilon_n(t), t \geq 0\}$.

The Whittle index of the *dummy bandits* is derived in the following lemma.

Lemma 4.2 *The Whittle index of the dummy bandits is 0, i.e.,* $W_0(\varepsilon, \gamma) = 0$.

Proof 4.3 *See Appendix 4.G.*

This policy activates the L bandits with the highest Whittle index at any given time t. As stated in Lemma 4.2, each dummy bandit has a Whittle index of $W_0(\varepsilon_0(t),$

Algorithm 6 Whittle Index Policy for Signal-aware Remote Estimation

Initialize the set of passive bandits $A = \{1, 2, \dots, N\}$.
for all time t **do**
  Update $X_{n,t}$ and $\hat{X}_{n,t}$ for all $n = 1, 2, \dots, N$ using (4.4) and (4.5), respectively.
  Update $\varepsilon_n(t)$, $\gamma_n(t)$, and the Whittle index $W_n(\varepsilon_n(t), \gamma_n(t))$ for all $n = 1, 2, \dots, N$ using (4.6) and (4.29), (4.30), (4.31), (4.32), (4.35), and (4.36).
  Update $A = \{n \in \{1, 2, \dots, N\} : \gamma_n(t) = 0\}$ for all $n = 1, 2, \dots, N$.
  Initialize total number of selected bandits $k = 0$.
  for all $l = 1, 2, \dots, L$ **do**
    if channel l is idle and $\max_{n \in A} W_n(\varepsilon_n(t), \gamma_n(t)) \geq 0$ **then**
      $n = \text{argmax}_{n\in A} W_n(\varepsilon_n(t), \gamma_n(t))$.
      Take a sample of bandit n and send it on channel l.
      Update total number of selected bandits $k = k + 1$.
      $A \leftarrow A - \{n\}$.
    end if
  end for
  Select $L - k$ dummy bandits.
end for

$\gamma_0(t)) = 0$. Consequently, if a bandit n (for $n = 1, 2, \dots, N$) possesses a negative Whittle index, denoted as $W_n(\varepsilon_n(t), \gamma_n(t)) < 0$, it will remain inactive. Furthermore, if source n is being served by a channel at time t such that $\gamma_n(t) > 0$, then $W_n(\varepsilon_n(t), \gamma_n(t)) = -\infty$ and no more channel will be scheduled to serve source n.

The Algorithm for solving (4.10)-(4.11) is provided in Algorithm 6. The set A of available bandits is initialized as $A = \{1, 2, \dots, N\}$. At any time t, the set A is updated for the bandits that do not have samples currently under service and the total number of selected bandits for transmission k is initialized as 0. If channel l is idle and $\max_{n\in A} W_n(\varepsilon_n(t), \gamma_n(t)) \geq 0$, then one sample is taken from bandit $n = \text{argmax}_{n\in A} W_n(\varepsilon_n(t), \gamma_n(t))$ and sent over channel; meanwhile, the total number of selected sources k is increase by 1 and bandit n is removed from the set A. Then, $L - k$ dummy bandits are selected for activation. Algorithm 6 can be either used as an event-driven algorithm or be executed on discretized time slots $t = 0, T_s, 2T_s, \dots$. When T_s is sufficiently small, the performance degradation caused by time discretization can be omitted.

Algorithm 7 Whittle Index Policy for Signal-aware Remote Estimation

Initialize the set of passive bandits $A = \{1, 2, \ldots, N\}$.
for all time t **do**
    Update $X_{n,t}$ and $\hat{X}_{n,t}$ for all $n = 1, 2, \ldots, N$ using (4.4) and (4.5), respectively.
    Update $\varepsilon_n(t)$, $\gamma_n(t)$, and the Whittle index $W_n(\varepsilon_n(t), \gamma_n(t))$ for all $n = 1, 2, \ldots, N$ using (4.6) and (4.29), (4.30), (4.31), (4.32), (4.35), and (4.36).
    Update $A = \{n \in \{1, 2, \ldots, N\} : \gamma_n(t) = 0\}$.
    for all $l = 1, 2, \ldots, L$ **do**
        if channel l is idle and $\max_{n \in A} W_n(\varepsilon_n(t), \gamma_n(t)) \geq 0$ **then**
            $n = \text{argmax}_{n \in A} W_n(\varepsilon_n(t), \gamma_n(t))$.
            Take a sample of bandit n and send it on channel l.
            $A \leftarrow A - \{n\}$.
        end if
    end for
end for

Now, we return to the original RMAB (4.8)-(4.9). The Whittle index scheduling policy for solving the original sampling and scheduling problem (4.8)-(4.9) is illustrated in Algorithm 7. Because RMAB (4.8)-(4.9) and the RMAB (4.10)-(4.11) are equivalent to each other, the Whittle index policy in Algorithm 6 and the Whittle index policy in Algorithm 7 are equivalent. Specifically, at any time t, L bandits having the highest non-negative Whittle index $W_n(\varepsilon, \gamma)$ will be activated. Because in the relaxed RMAB (4.13)-(4.14), a bandit n having $W_n(\varepsilon, \gamma) \leq 0$ will never be made active, the *dummy bandits* with $W_0(\varepsilon, \gamma)$ will be made active. As there are L *dummy bandits*, the constraint (4.11) will be satisfied.

In Algorithm 7, the set A of passive bandits is initialized as $A = \{1, 2, \ldots, N\}$. If channel l is idle and $\max_{n \in A} W_n(\varepsilon_n(t), \gamma_n(t)) \geq 0$, then one sample is taken from bandit $n{=}\text{argmax}_{n \in A} W_n(\varepsilon_n(t), \gamma_n(t))$ and sent over channel; meanwhile, bandit n is removed from the set A of passive bandits.

4.5 Signal-agnostic Scheduling

A scheduling policy $\pi \in \Pi$ is called *signal-agnostic* if the policy π is independent of the observed process $\{X_{n,t}, t \geq 0\}_{n=1}^{N}$. Let $\Pi_{\text{agnostic}} \in \Pi$ denote the set of signal-agnostic, causal policies, defined by

$$\Pi_{\text{agnostic}} = \{\pi \in \Pi : \pi \quad \text{is independent of} \quad \{X_{n,t}, t \geq 0\}_{n=1}^{N}\}. \tag{4.44}$$

In a signal-agnostic policy, the mean-squared estimation error of the process $X_{n,t}$ at time t is [35], [34]

$$\mathbb{E}[\varepsilon_n^2(t)] = p_n(\Delta_n(t)) = \begin{cases} \frac{\sigma_n^2}{2\theta_n}(1 - e^{-2\theta_n \Delta_n(t)}), & \text{if } \theta_n \neq 0, \\ \sigma_n^2 \Delta_n(t), & \text{if } \theta_n = 0, \end{cases} \tag{4.45}$$

where $\Delta_n(t)$ is the AoI and $p_n(\cdot)$ is an increasing function defined in (4.45). By using (4.45), for any policy $\pi \in \Pi_{\text{agnostic}}$

$$\mathbb{E}\left[\int_0^T \varepsilon_n^2(t)dt\right] = \mathbb{E}\left[\int_0^T p_n(\Delta_n(t))dt\right]. \tag{4.46}$$

Hence, the signal-agnostic sampling and scheduling problem can be formulated as

$$\inf_{\pi \in \Pi_{\text{agnostic}}} \limsup_{T \to \infty} \sum_{n=1}^{N} w_n \mathbb{E}_\pi \left[\frac{1}{T} \int_0^T p_n(\Delta_n(t))dt\right] \tag{4.47}$$

$$\text{s.t.} \quad \sum_{n=1}^{N} c_n(t) \leq L, c_n(t) \in \{0, 1\}, t \in [0, \infty). \tag{4.48}$$

4.5.1 Restless Multi-armed Bandit: Relaxation and Lagrangian Decomposition

Problem (4.47)-(4.48) is a continuous-time Restless Multi-armed Bandit (RMAB) with a continuous state space, where $\Delta_n(t)$ of source n is modeled as the state of the restless bandit. Following the procedure developed in Section 4.4.1, we consider L additional *dummy bandits* where $c_0(t) \in \{0, 1, 2, \ldots, L\}$ denotes the number of *dummy bandits* that are activated at time t and reformulate (4.47)-(4.48) as

$$\inf_{\pi\in\Pi_{\text{agnostic}}} \limsup_{T\to\infty} \sum_{n=1}^{N} w_n \mathbb{E}_\pi\left[\frac{1}{T}\int_0^T p_n(\Delta_n(t))dt\right] \tag{4.49}$$

$$\text{s.t.} \quad \limsup_{T\to\infty} \sum_{n=0}^{N} \mathbb{E}_\pi\left[\frac{1}{T}\int_0^T c_n(t)dt\right] = L,$$

$$c_0(t) \in \{0,1,\ldots,L\}, c_n(t) \in \{0,1\}, n = 1,2,\ldots,N, t\in[0,\infty). \tag{4.50}$$

which is an RMAB with an equality constraint.

By relaxing constraint (4.50), the RMAB (4.49)-(4.50) is reformulated as

$$\inf_{\pi\in\Pi_{\text{agnostic}}} \limsup_{T\to\infty} \sum_{n=0}^{N} w_n \mathbb{E}_\pi\left[\frac{1}{T}\int_0^T p_n(\Delta_n(t))dt\right] \tag{4.51}$$

$$\text{s.t.} \quad \limsup_{T\to\infty} \sum_{n=0}^{N} \mathbb{E}_\pi\left[\frac{1}{T}\int_0^T c_n(t)dt\right] = L,$$

$$c_0(t) \in \{0,1,\ldots,L\}, c_n(t) \in \{0,1\}, n = 1,2,\ldots,N, t\in[0,\infty). \tag{4.52}$$

Next, we take the Lagrangian dual decomposition of the relaxed problem (4.51)-(4.52), which produces the following problem with a dual variable $\lambda \in \mathbb{R}$:

$$\inf_{\pi\in\Pi_{\text{agnostic}}} \limsup_{T\to\infty} \mathbb{E}_\pi\left[\frac{1}{T}\int_0^T \sum_{n=1}^{N} w_n p_n(\Delta_n(t)) + \lambda\left(\sum_{n=0}^{N} c_n(t) - L\right)dt\right]. \tag{4.53}$$

Then, problem (4.53) can be decomposed into $(N+1)$ separated sub-problems. The sub-problem associated with source n is

$$\bar{m}_{n,\text{age-opt}} = \inf_{\pi_n\in\Pi_{n,\text{agnostic}}} \limsup_{T\to\infty} \mathbb{E}_{\pi_n}\left[\frac{1}{T}\int_0^T w_n p_n(\Delta_n(t)) + \lambda c_n(t)dt\right], \tag{4.54}$$

where $\bar{m}_{n,\text{age-opt}}$ is the optimum value of (4.54), $\pi_n = (S_{n,1}, S_{n,2}, \ldots)$ denotes a sub-scheduling policy for source n, and $\Pi_{n,\text{agnostic}}$ is the set of all causal sub-scheduling policies of source n.

4.5.2 Indexability

An optimal solution to problem (4.54) is provided in the following proposition.

Proposition 4.2 *For signal-agnostic scheduling, if the* $Y_{n,i}$*'s are i.i.d. with* $0 < \mathbb{E}[Y_{n,i}] < \infty$*, then* $(S_{n,1}(\beta_{n,age}), S_{n,2}(\beta_{n,age}), \ldots)$ *with a parameter* $\beta_{n,age}$ *is an optimal solution to* (4.54)*, where*

$$S_{n,i+1}(\beta_{n,age}) = \inf \left\{ t \geq D_{n,i}(\beta_{n,age}) : \mathbb{E}[p_n(\Delta_n(t + Y_{n,i+1}))] \geq \beta_{n,age} \right\}, \tag{4.55}$$

$D_{n,i}(\beta_{n,age}) = S_{n,i}(\beta_{n,age}) + Y_{n,i}$*, and* $\beta_{n,age}$ *is the unique root of*

$$\mathbb{E}\left[\int_{D_{n,i}(\beta_{n,age})}^{D_{n,i+1}(\beta_{n,age})} w_n \varepsilon_n^2(t) dt\right] - \beta_{n,age}\mathbb{E}[D_{n,i+1}(\beta_{n,age}) - D_{n,i}(\beta_{n,age})] + \lambda\mathbb{E}[Y_{n,i+1}] = 0. \tag{4.56}$$

The optimal objective value to (4.54) *is given by*

$$\bar{m}_{n,\text{age-opt}} = \frac{\mathbb{E}\left[\int_{D_{n,i}(\beta_{n,\text{age}})}^{D_{n,i+1}(\beta_{n,\text{age}})} w_n p_n(\Delta_n(t)) dt\right] + \lambda\mathbb{E}[Y_{n,i+1}]}{\mathbb{E}[D_{n,i+1}(\beta_{n,\text{age}}) - D_{n,i}(\beta_{n,\text{age}})]}. \tag{4.57}$$

Furthermore, $\beta_{n,age}$ *is exactly the optimal objective value of* (4.54)*, i.e.,* $\beta_{n,\text{age}} = \bar{m}_{n,\text{age-opt}}$.

Let $\Psi_{n,\text{age}}(\lambda)$ be a set of states $(\delta, \gamma) \in [0, \infty) \times [0, \infty)$ such that if $\Delta_n(t) = \delta$ and $\gamma_n(t) = \gamma$, the optimal solution for (4.54) is to take a passive action at time t.

Definition 4.3 ***(Indexability).*** *[115] Bandit n is said to be indexable if, as the activation cost* λ *increases from* $-\infty$ *to* ∞*, the set* $\Psi_{n,age}(\lambda)$ *increases monotonically, i.e.,* $\lambda_1 \leq \lambda_2$ *implies* $\Psi_{n,age}(\lambda_1) \subseteq \Psi_{n,age}(\lambda_2)$*. The RMAB* (4.47)-(4.48) *is said to be indexable if all* $(N+1)$ *bandits are indexable.*

By using Proposition 4.2, the set $\Psi_{n,\text{age}}(\lambda)$ in Definition 4.3 can be simplified as

$$\Psi_{n,\text{age}}(\lambda) = \{(\delta, \gamma) \in [0, \infty) \times [0, \infty) : \gamma > 0 \,\text{or}\, \mathbb{E}[p_n(\delta + Y_{n,i+1})] < \beta_{n,\text{age}}(\lambda)\}. \tag{4.58}$$

Following the techniques developed in Section 4.4, we can obtain

Theorem 4.3 *If* $p_n(\delta)$ *is a strictly increasing function of* δ*, the RMAB problem* (4.47)-(4.48) *is indexable.*

Proof 4.4 *See Appendix 4.I.*

4.5.3 Whittle Index Policy

Theorem 4.4 *In the RMAB problem (4.47)-(4.48), if $p_n(\delta)$ is a strictly increasing function of δ, the $Y_{n,i}$'s are i.i.d. with $0 < \mathbb{E}[Y_{n,i}] < \infty$, then the following assertions are true for the Whittle index of source n at state (δ, γ):*

(a) If $\gamma = 0$, then

$$W_{n,age}(\delta, 0) = \frac{w_n}{\mathbb{E}[Y_{n,i}]}\left\{\mathbb{E}[D_{n,i+1}(\delta) - D_{n,i}(\delta)]\mathbb{E}[p_n(\delta + Y_{n,i+1})] - \mathbb{E}\left[\int_{D_{n,i}(\delta)}^{D_{n,i+1}(\delta)} p_n(s)ds\right]\right\}, \tag{4.59}$$

where $D_{n,i}(\delta) = S_{n,i}(\delta) + Y_{n,i}$, and

$$S_{n,i+1}(\delta) = D_{n,i}(\delta) + \max\{\delta - Y_{n,i}, 0\}. \tag{4.60}$$

(b) If $\gamma > 0$, then

$$W_{n,age}(\delta, \gamma) = -\infty. \tag{4.61}$$

Proof 4.5 *See Appendix 4.J.*

The expectations in (4.59) can be easily evaluated using the following lemma:

Lemma 4.3 *In Theorem 4.4, it holds that*

$$\mathbb{E}[D_{n,i+1}(\delta) - D_{n,i}(\delta)] = \mathbb{E}[\max\{\delta, Y_{n,i}\}], \tag{4.62}$$

$$\mathbb{E}\left[\int_{D_{n,i}(\delta)}^{D_{n,i+1}(\delta)} p_n(s)ds\right] = \mathbb{E}[R_{n,3}(\max\{\delta, Y_{n,i}\} + Y_{n,i+1})] - \mathbb{E}[R_{n,3}(Y_{n,i})], \tag{4.63}$$

where

$$R_{n,3}(\delta) = \int_0^\delta p_n(s)ds. \tag{4.64}$$

Algorithm 8 Whittle Index Policy for Signal-aware Remote Estimation

Initialize the set of passive bandits $A = \{1, 2, \ldots, N\}$.
for all time t **do**
  Update $\Delta_n(t)$ for all $n = 1, 2, \ldots, N$ using (4.3), (4.59), (4.61), (4.62), and (4.63).
  Update $A = \{n \in \{1, 2, \ldots, N\} : \gamma_n(t) = 0\}$.
  Initialize total number of selected bandits $k = 0$.
  for all $l = 1, 2, \ldots, L$ **do**
    if channel l is idle and $\max_{n \in A} W_n(\Delta_n(t), \gamma_n(t)) \geq 0$ **then**
      $n = \text{argmax}_{n \in A} W_n(\Delta_n(t), \gamma_n(t))$.
      Take a sample of bandit n and send it on channel l.
      Update total number of selected bandits $k = k + 1$.
      $A \leftarrow A - \{n\}$.
    end if
  end for
  Select $L - k$ dummy bandits.
end for

Proof 4.6 *See Appendix 4.K.*

Theorems 4.3-4.4 and Lemma 4.3 hold for all increasing functions $p_n(\delta)$ of the AoI δ, not necessarily the mean-square estimation error function in (4.45). The Algorithms for solving RMAB (4.49)-(4.50) is provided in Algorithm 8 and for solving the original RMAB is provided in (4.47)-(4.48) Algorithm 9.

Theorems 4.3-4.4, Lemma 4.3, and Algorithms 8-9 generalize prior studies on AoI-based Whittle index policies, e.g., [16, 41, 42]. More specifically, the Whittle index policies detailed in [16, 41, 42] were derived for the scenario of constant transmission times where the zero-wait sampling policy [111, 112] is an optimal solution for the sub-problem (4.54), and the resulting Whittle index always maintains a non-negative value. In contrast, our current study accommodates scenarios involving *i.i.d.* random transmission times. In such cases, the optimality of zero-wait sampling is not assured for sub-problem (4.54), resulting in the potential for both positive and negative values for the Whittle index derived in Theorem 4.4.

Algorithm 9 Whittle Index Policy for Signal-agnostic Remote Estimation

Initialize the set A of passive bandits $A = \{1, 2, \dots, N\}$.
for all time t **do**
  Update $\Delta_n(t)$, $\gamma_n(t)$, and the Whittle index $W_{n,\text{age}}(\Delta_n(t), \gamma_n(t))$ for all $n = 1, 2, \dots, N$ using (4.3), (4.59), (4.61), (4.62), and (4.63).
  Update $A = \{n \in \{1, 2, \dots, N\} : \gamma_n(t) = 0\}$.
  for all $l = 1, 2, \dots, L$ **do**
    if channel l is idle and $\max_{n \in A} W_{n,\text{age}}(\Delta_n(t), \gamma_n(t)) \geq 0$ **then**
      $n = \text{argmax}_{n \in A} W_{n,\text{age}}(\Delta_n(t), \gamma_n(t))$.
      Take a sample of bandit n and send it on channel l.
      $A \leftarrow A - \{n\}$.
    end if
  end for
end for

4.6 Unity of Threshold-based Sampling and Whittle Index-based Scheduling

We find an interesting relationship between Threshold-based Sampling and Whittle Index-based Scheduling for Remote Estimation and AoI minimization. Specifically, for single-source, single-channel case, both of the policies are equivalent. We provide the details thorughout this section.

4.6.1 Remote Estimation

Let consider the special case $N = L = 1$, where the system has a single source and a single channel. Let $w_1 = 1$, then problem (5.1)-(5.2) reduces to

$$\bar{m}_{1,\text{opt}} = \inf_{\pi \in \Pi} \limsup_{T \to \infty} \mathbb{E}_\pi \left[\frac{1}{T} \int_0^T \varepsilon_1^2(t) dt \right]. \tag{4.65}$$

The single-source, single-channel sampling and scheduling problem (4.65) is a special case of Proposition 4.1 with $n = 1$ and $\lambda = 0$. A threshold-based optimal solution to (4.65) is provided by the following corollary of Proposition 4.1.

Corollary 4.1 *If the $Y_{1,i}$'s are i.i.d. with $0 < \mathbb{E}[Y_{1,i}] < \infty$, then $(S_{1,1}(\beta_1), S_{1,2}(\beta_1), \dots)$ with a parameter β_1 is an optimal solution to (4.65), where*

$$S_{1,i+1}(\beta_1) = \inf_t \{ t \geq D_{1,i}(\beta_1) : |\varepsilon_1(t)| \geq v_1(\beta_1) \}, \tag{4.66}$$

$D_{1,i}(\beta_1) = S_{1,i}(\beta_1) + Y_{1,i}$, $v_1(\beta_1)$ is defined by

$$v_1(\beta_1) = \begin{cases} \frac{\sigma_1}{\sqrt{\theta_1}} G^{-1}\left(\frac{\frac{\sigma_1^2}{2\theta_1}\mathbb{E}[e^{-2\theta Y_{1,i}}]}{\frac{\sigma_1^2}{2\theta_1} - \beta_1}\right), & \text{if } \theta_1 > 0, \\ \sqrt{3(\beta_1 - \sigma_1^2 \mathbb{E}[Y_{1,i}])}, & \text{if } \theta_1 = 0, \\ \frac{\sigma_1}{\sqrt{-\theta_1}} K^{-1}\left(\frac{\frac{\sigma_1^2}{2\theta_1}\mathbb{E}[e^{-2\theta_1 Y_{1,i}}]}{\frac{\sigma_1^2}{2\theta_1} - \beta_1}\right), & \text{if } \theta_1 < 0, \end{cases} \tag{4.67}$$

$G^{-1}(\cdot)$ and $K^{-1}(\cdot)$ are the inverse functions of $G(x)$ in (4.18) and $K(x)$ in (4.19), respectively, for the region $x \in [0, \infty)$, and β_1 is the unique root of

$$\mathbb{E}\left[\int_{D_{1,i}(\beta_1)}^{D_{1,i+1}(\beta_1)} \varepsilon_1^2(t)dt\right] - \beta_1 \mathbb{E}[D_{1,i+1}(\beta_1) - D_{1,i}(\beta_1)] = 0. \tag{4.68}$$

The optimal objective value to (4.65) is given by

$$\bar{m}_{1,\text{opt}} = \frac{\mathbb{E}\left[\int_{D_{1,i}(\beta_1)}^{D_{1,i+1}(\beta_1)} \varepsilon_1^2(t)dt\right]}{\mathbb{E}[D_{1,i+1}(\beta_1) - D_{1,i}(\beta_1)]}. \tag{4.69}$$

Furthermore, β_1 is exactly the optimal objective value of (4.65), i.e., $\beta_1 = \bar{m}_{1,\text{opt}}$.

Corollary 4.1 follows directly from Proposition 4.1. For the cases of the Wiener process ($\theta_1 = 0$) and stable OU process ($\theta_1 > 0$), the threshold-based policy in Corollary 4.1 were earlier reported in [34]. The case of unstable OU process ($\theta_1 < 0$) is new.

It is important to note that the threshold-based policy in Corollary 4.1 and the Whittle index policy in the following theorem are equivalent.

Theorem 4.5 *If the $Y_{1,i}$'s are i.i.d. with $0 < \mathbb{E}[Y_{1,i}] < \infty$, then $(S_{1,1}, S_{1,2}, \ldots)$ is an optimal solution to (4.65), where*

$$S_{1,i+1} = \inf_t \{t \geq S_{1,i} : W_1(\varepsilon_1(t), \gamma_1(t)) \geq 0\}, \tag{4.70}$$

and $W_1(\varepsilon_n(t), \gamma_n(t))$ is the Whittle index of source 1, defined by (4.29), (4.30), (4.31), and (4.32).

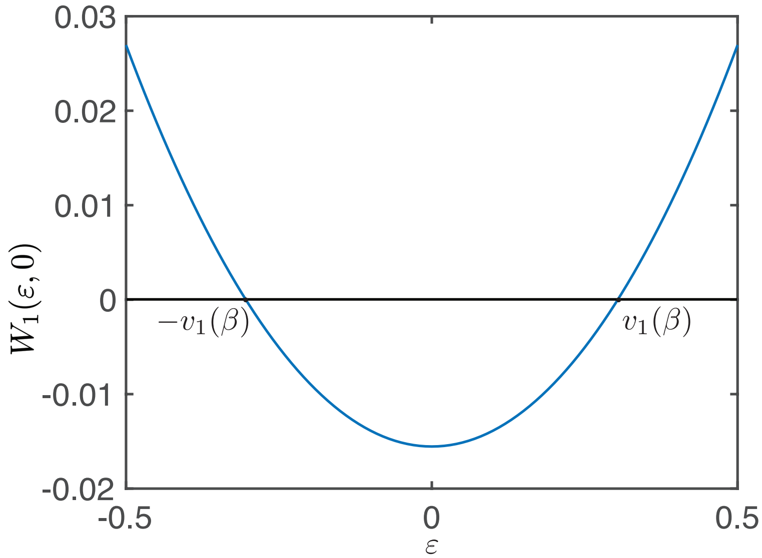


Figure 4.2: Illustration of the Whittle index $W_1(\varepsilon, \gamma)$ and the optimal threshold $v_1(\beta_1)$, where the parameters of the Gauss-Markov process are $\sigma_1 = 1$ and $\theta_1 = 0.1$ and the *i.i.d.* transmission times follow an exponential distribution with mean $\mathbb{E}[Y_{1,i}] = 2$.

Proof sketch. Because (i) Corollary 4.1 provides an optimal solution to (4.65) and (ii) (4.70) is equivalent to the solution in Corollary 4.1, (4.70) is also an optimal solution to (4.65). The details are provided in Appendix 4.H. □

Corollary 4.1 and Theorem 4.5 reveal a unification of threshold-based sampling and scheduling policy developed in [34] and the Whittle index policy developed in this chapter. In particular, if the Whittle index $W_1(\varepsilon_1(t), \gamma_1(t)) = 0$, then (i) the channel is idle at time t and (ii) the instantaneous estimation error $|\varepsilon_1(t)|$ exactly crosses the optimal threshold $v_1(\beta_1)$ at time t. As illustrated in Figure 4.2, $\varepsilon = \pm v_1(\beta_1)$ are the roots of equation $W_1(\varepsilon, 0) = 0$.

The threshold-based sampling and scheduling results outlined in Corollary 4.1 and [34] are applicable specifically to the single-source, single-channel scenario. Nevertheless, our exploration in Sections 4.4.1-4.4.3 illustrates the methodology for utilizing these findings to establish indexability and evaluate the Whittle index in the multi-source, multi-channel scenario.

4.6.2 AoI Minimization

For single-source, single-channel special case with $w_1 = 1$, problem (4.47)-(4.48) reduces to

$$\bar{m}_{1,\text{age-opt}} = \inf_{\pi \in \Pi_{\text{agnostic}}} \limsup_{T \to \infty} \mathbb{E}_\pi \left[\frac{1}{T} \int_0^T p_1(\Delta_1(t))dt \right] \tag{4.71}$$

Theorem 4.6 *If $p_1(\delta)$ is a strictly increasing function of δ, the $Y_{1,i}$'s are* i.i.d. *with $0 < \mathbb{E}[Y_{1,i}] < \infty$, then $(S_{1,1}, S_{1,2}, \ldots)$ is an optimal solution to* (4.71)*, where*

$$S_{1,i+1} = \inf_t \{t \geq S_{1,i} : W_{1,age}(\Delta_1(t), \gamma_1(t)) \geq 0\}, \tag{4.72}$$

where $W_{1,age}(\Delta_1(t), \gamma_1(t))$ is the Whittle index of source 1, defined by (4.59) *and* (4.61)*.*

In the AoI literature, threshold-based scheduling and Whittle index have been two distinct approaches for AoI minimization. Our study unifies the two approaches: for AoI minimization of a single source, the threshold policy in [4, Theorem 1] and the Whittle index policy based in Theorem 4.6 are equivalent. Specifically, if the Whittle index $W_{1,\text{age}}(\varepsilon_1(t), \gamma_1(t)) = 0$, then (i) the channel is idle at time t and (ii) the expected age-penalty function surpasses the threshold in [4, Theorem 1] at time t.

4.7 Numerical Results

In this section, we compare the following three scheduling policies for multi-source remote estimation:

- Maximum Age First, Zero-Wait (MAF-ZW) policy: Suppose that $N \geq L$. Whenever one channel l becomes free, the MAF-ZW policy will take a sample from the source with the highest AoI among the sources that are currently not served by any channel, and send the sample over channel l.
- Signal-agnostic, Whittle Index policy: The policy that we proposed in Algorithm 9.
- Signal-aware, Whittle Index policy: The policy that we proposed in Algorithm 7.

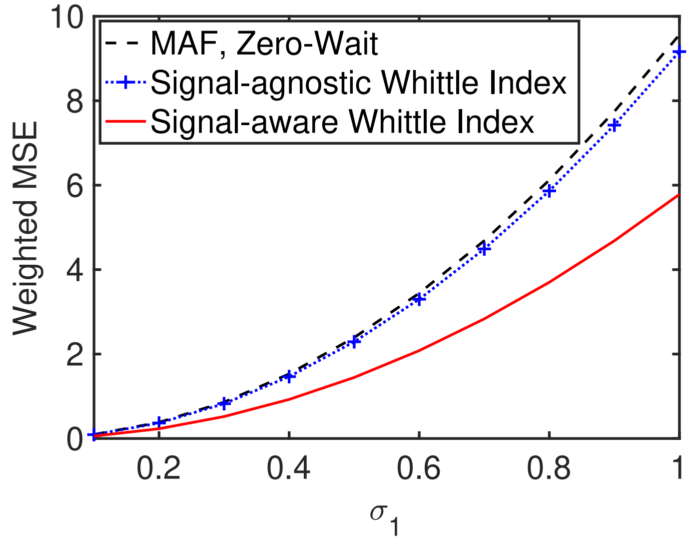


Figure 4.3: Total time-average MSE vs the parameter σ_1 of the Gauss-Markov source 1, where the number of sources is $N = 4$ and the number of channels is $L = 2$. The transmission times are *i.i.d.*, following a normalized log-normal distribution with parameter $\rho = 1.5$, and $\mathbb{E}[Y_{n,i}] = 1$. The other parameters of the Gauss-Markov sources are $\sigma_2 = 0.8, \sigma_3 = 0.9, \sigma_4 = 1$, and $\theta_1 = -0.1, \theta_2 = \theta_3 = \theta_4 = 0.1$.

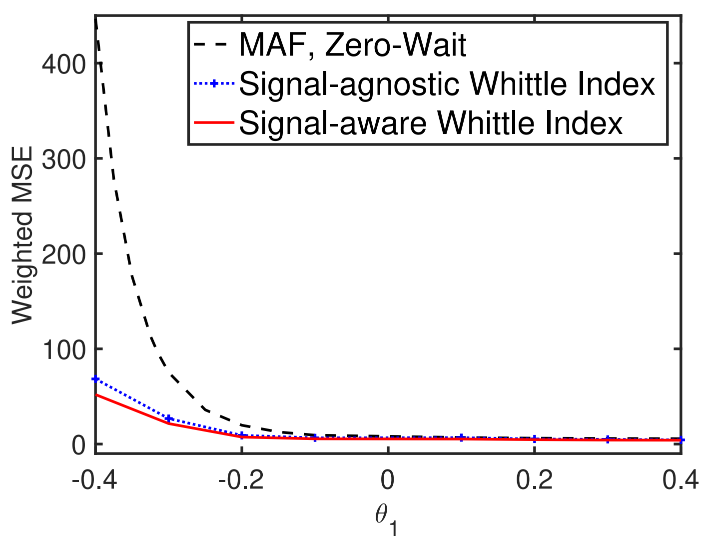


Figure 4.4: Total time-average MSE vs the parameter θ_1 of the Gauss-Markov source 1, where the number of sources is $N = 4$ and the number of channels is $L = 2$. The transmission times are *i.i.d.*, following a normalized log-normal distribution with parameter $\rho = 1.5$, and $\mathbb{E}[Y_{n,i}] = 1$. The other parameters for the Gauss-Markov sources are $\sigma_1 = \sigma_2 = \sigma_3 = \sigma_4 = 1$, and $\theta_2 = 0.2, \theta_3 = 0.3, \theta_4 = 0.1$.

Figure 5.3 depicts the total time-average mean-squared estimation error versus the parameter σ_1 of the Gauss-Markov source 1, where the number of sources is $N = 4$ and the number of channels is $L = 2$. The other parameters of the Gauss-Markov processes are $\sigma_2 = 0.8, \sigma_3 = 0.9, \sigma_4 = 1$, and $\theta_1 = -0.1, \theta_2 = \theta_3 = \theta_4 = 0.1$. The transmission times are *i.i.d.* and follow a normalized log-normal distribution, where $Y_{n,i} = e^{\rho Q_{n,i}}/\mathbb{E}[e^{\rho Q_{n,i}}]$, $\rho > 0$ is the scale parameter of the log-normal distribution, and $(Q_{n,1}, Q_{n,2}, \ldots)$ are *i.i.d.* Gaussian random variables with zero mean and unit variance. In our simulation, $\rho = 1.5$. All sources are given the same weight $w_1 = w_2 = w_3 = w_4 = 1$. In Figure 5.3, the signal-aware Whittle index policy has a smaller total MSE than the signal-agnostic Whittle index policy and the MAF-ZW policy. The total MSE of the signal-aware Whittle index policy achieves up to 1.58 times performance gain over the signal-agnostic Whittle index policy, and up to 1.65 times than the MAF-ZW policy.

Figure 5.5 illustrates the total time-average mean-squared estimation error versus the parameter θ_1 of the Gauss-Markov source 1, where the number of sources is $N = 4$, and the number of channels is $L = 2$. The other parameters of the Gauss-Markov processes are $\theta_2 = 0.2, \theta_3 = 0.3, \theta_4 = 0.1$, and $\sigma_1 = \sigma_2 = \sigma_3 = \sigma_4 = 1$. The transmission time distribution and the weights of the sources are the same as in Figure 5.3. In Figure 5.5, the total MSE of the signal-aware Whittle index policy achieves up to 8.6 times performance gain over the MAF-ZW policy and up to 1.32 times over the signal-agnostic Whittle index policy. When $\theta_1 < 0$, the performance gain of the signal-aware Whittle index policy is much higher than that in the case of $\theta_1 > 0$. This suggests a high performance gain can be achieved if the Gauss-Markov sources are highly unstable. For all three policies, the total MSE decreases, as θ_1 increases.

4.8 Conclusion

In this chapter, we have studied a sampling and scheduling problem in which samples of multiple Gauss-Markov sources are sent to remote estimators that need to monitor the sources in real-time. The formulated sampling and scheduling problem is a restless multi-armed bandit problem, where each bandit process has a continuous state space and requires continuous-time control. We have proved that the problem is indexable and proposed a

Whittle index policy. Analytical expressions of the Whittle index have been obtained. For single-source, single-channel scheduling, we have showed that it is optimal to take a sample at the earliest time when the Whittle index is no less than zero. This result provides a new interpretation of earlier studies on threshold-based sampling policies for the Wiener and Ornstein-Uhlenbeck processes.

Appendix

4.A Proof of (4.4) for $\theta_n < 0$

A solution to (4.1) for initial state $X_{n,0} = 0$ and parameter $\mu_n = 0$ can be written in terms of a stochastic integral as follows

$$X_{n,t} = \sigma_n e^{-\theta_n(t-S_{n,i})} \int_0^{t-S_{n,i}} e^{\theta_n s} dW_s, \tag{4.73}$$

which holds for any value of θ_n. To derive an alternative formula for $\theta_n < 0$, let consider the following well-known lemma:

Lemma 4.4 *Let $Y_{n,t}$ be a Gaussian process with $Y_{n,0} = 0$, $\mathbb{E}[Y_{n,t}] = 0$, and it has independent increments. Then the distribution of $Y_{n,t}$ can be completely determined by its variance function $\mathbb{E}[Y_{n,t}^2]$.*

Define

$$Y_{n,t} = e^{\theta_n(t-S_{n,i})} X_{n,t} = \sigma_n \int_0^{t-S_{n,i}} e^{\theta_n s} dW_s, \tag{4.74}$$

Lemma 4.4 implies that $Y_{n,t}$ in (4.74) is a Gaussian process and its variance function is given by

$$\mathbb{E}[Y_{n,t}^2] = \frac{\sigma_n^2}{2\theta_n}(e^{2\theta_n(t-S_{n,i})} - 1) \tag{4.75}$$

Consider the following process

$$Z_{n,t} = \frac{\sigma_n}{\sqrt{2\theta_n}} W_{n,e^{2\theta_n(t-S_{n,i})}-1}. \tag{4.76}$$

Because Brownian motion $W_{n,t}$ is a Gaussian process with variance function t, $Z_{n,t}$ in (4.76) is a Gaussian process with variance function $\frac{\sigma_n^2}{2\theta_n}(e^{2\theta_n(t-S_{n,i})} - 1)$ which is the same as the variance function of $Y_{n,t}$. From Lemma 4.4, both the processes $Y_{n,t}$ and $Z_{n,t}$ are equal in distribution.

When $\theta_n < 0$, consider $\rho_n = -\theta_n > 0$. In this setting, the variance function of $Y_{n,t}$ can be written as

$$\mathbb{E}[Y_{n,t}^2] = \frac{\sigma_n^2}{2\rho_n}(1 - e^{-2\rho_n(t-S_{n,i})}). \tag{4.77}$$

Let

$$Z'_{n,t} = \frac{\sigma_n}{\sqrt{2\rho_n}} W_{n,1-e^{-2\rho_n(t-S_{n,i})}}, \tag{4.78}$$

which implies $Z'_{n,t}$ has the same variance function as $Y_{n,t}$. Hence, $Z'_{n,t}$ and $Y_{n,t}$ are equal in distribution. By using (4.77) and (4.78) in (4.74), we get that

$$X_{n,t} = \frac{\sigma_n^2}{2\rho_n} e^{\rho_n(t-S_{n,i})} W_{n,1-e^{-2\rho_n(t-S_{n,i})}}, \tag{4.79}$$

from which (4.4) for $\theta_n < 0$ follows. This completes the proof.

4.B Proof of the Simplification of Policy π_n

The sampling and scheduling policy $\pi_n = ((S_{n,1}, G_{n,1}), (S_{n,2}, G_{n,2}), \ldots)$ consists of the sampling time $S_{n,i}$ and the transmission starting time $G_{n,i}$ for each sample i. In policy π_n, sample i can be generated when the server is busy sending another sample, and hence sample i needs to wait for some time before being submitted to the server, i.e., $S_{n,i} < G_{n,i}$. Consider a sampling and scheduling policy $\pi'_n = \{(S_{n,1}, G_{n,1}), \ldots, (S_{n,i-1}, G_{n,i-1}), G_{n,i}, (S_{n,i+1}, G_{n,i+1}), \ldots\}$ such that the generation time and transmission starting time of sample i are equal to each other, i.e., $S_{n,i} = G_{n,i}$. We will show that the MSE of the sampling policy π'_n is smaller than that of the sampling policy π_n.

Note that $\{X_{n,t} : t \in [0, \infty)\}$ does not change according to the sampling policy, and the sample delivery times $\{D_{n,1}, D_{n,2}, \ldots\}$ remain the same in policy π_n and policy π'_n. Hence, the only difference between policies π_n and π'_n is that *the generation time of sample* i. The MMSE estimator under policy π_n is given by (4.5) and the MMSE estimator under policy

π'_n is given by

$$
\begin{aligned}
\hat{X}'_{n,t} =& \mathbb{E}[X_{n,t}|(S_{n,j}, X_{n,S_{n,j}}, G_{n,j}, D_{n,j})_{j\leq i-1}, (G_{n,i}, X_{n,G_{n,i}}, D_{n,i})] \\
=& \begin{cases} \mathbb{E}[X_{n,t}|G_{n,i}, X_{n,G_{n,i}}], & t \in [D_{n,i}, D_{n,i+1}); \\ \mathbb{E}[X_{n,t}|S_{n,j}, X_{n,S_{n,j}}], & t \in [D_{n,j}, D_{n,j+1}),\ j \neq i. \end{cases}
\end{aligned}
\tag{4.80}
$$

Next, we consider another sampling and scheduling policy π''_n in which the samples $(G_{n,i}, X_{n,G_{n,i}})$ and $(S_{n,i}, X_{n,S_{n,i}})$ are both delivered to the estimator at the same time $D_{n,i}$. Clearly, the estimator under policy π''_n has more information than those under policies π_n and π'_n. One can also show that the MMSE estimator under policy π''_n is

$$
\begin{aligned}
\hat{X}''_{n,t} =& \mathbb{E}[X_{n,t}|(S_{n,j}, X_{n,S_{n,j}}, G_{n,j}, D_{n,j})_{j\leq i}, (G_{n,i}, X_{n,G_{n,i}}, D_{n,i})] \\
=& \begin{cases} \mathbb{E}[X_{n,t}|G_{n,i}, X_{n,G_{n,i}}], & t \in [D_{n,i}, D_{n,i+1}); \\ \mathbb{E}[X_{n,t}|S_{n,j}, X_{n,S_{n,j}}], & t \in [D_{n,j}, D_{n,j+1}),\ j \neq i. \end{cases}
\end{aligned}
\tag{4.81}
$$

Notice that, because of the strong Markov property of OU process, the estimator under policy π''_n uses the fresher sample $(G_{n,i}, X_{n,G_{n,i}})$, instead of the stale sample $(S_{n,i}, X_{n,S_{n,i}})$, to construct $\hat{X}''_{n,t}$ during $[D_{n,i}, D_{n,i+1})$. Because the estimator under policy π''_n has more information than that of under policy π_n, one can imagine that policy π''_n has a smaller estimation error than policy π_n, i.e.,

$$
\mathbb{E}\left\{\int_{D_{n,i}}^{D_{n,i+1}} (X_{n,t} - \hat{X}_{n,t})^2 dt\right\} \geq \mathbb{E}\left\{\int_{D_{n,i}}^{D_{n,i+1}} (X_{n,t} - \hat{X}''_{n,t})^2 dt\right\}. \tag{4.82}
$$

To prove (4.82), we invoke the orthogonality principle of the MMSE estimator [98, Prop. V.C.2] under policy π''_n and obtain

$$
\mathbb{E}\left\{\int_{D_{n,i}}^{D_{n,i+1}} 2(X_{n,t} - \hat{X}''_{n,t})(\hat{X}''_{n,t} - \hat{X}_{n,t}) dt\right\} = 0, \tag{4.83}
$$

where we have used the fact that $(G_{n,i}, X_{n,G_{n,i}})$ and $(S_{n,i}, X_{n,S_{n,i}})$ are available by the MMSE estimator under policy π''_n. Next, from (4.83), we can get

$$
\begin{aligned}
&\mathbb{E}\left\{\int_{D_{n,i}}^{D_{n,i+1}} (X_{n,t}-\hat{X}_{n,t})^2 dt\right\}\\
=&\mathbb{E}\left\{\int_{D_{n,i}}^{D_{n,i+1}} (X_{n,t}-\hat{X}''_{n,t})^2 + (\hat{X}''_{n,t}-\hat{X}_{n,t})^2 dt\right\}\\
&+\mathbb{E}\left\{\int_{D_{n,i}}^{D_{n,i+1}} 2(X_{n,t}-\hat{X}''_{n,t})(\hat{X}''_{n,t}-\hat{X}_{n,t}) dt\right\}
\end{aligned}
$$

$$
\begin{aligned}
=&\mathbb{E}\left\{\int_{D_{n,i}}^{D_{n,i+1}} (X_{n,t}-\hat{X}''_{n,t})^2 + (\hat{X}''_{n,t}-\hat{X}_{n,t})^2 dt\right\}\\
\geq&\mathbb{E}\left\{\int_{D_{n,i}}^{D_{n,i+1}} (X_{n,t}-\hat{X}''_{n,t})^2 dt\right\}. \qquad (4.84)
\end{aligned}
$$

In other words, the estimation error of policy π''_n is no greater than that of policy π_n. Furthermore, by comparing (4.80) and (4.81), we can see that the MMSE estimators under policies π''_n and π'_n are exactly the same. Therefore, the estimation error of policy π'_n is no greater than that of policy π_n.

By repeating the above arguments for all samples i satisfying $S_{n,i} < G_{n,i}$, one can show that the sampling policy $\{G_{n,1}, G_{n,2}, \ldots\}$ is better than the sampling policy $\pi = \{(S_{n,1}, G_{n,1}), (S_{n,2}, G_{n,2}), \ldots\}$. This completes the proof.

4.C Proof of Proposition 4.1

In this section, we present the proof of Proposition 4.1 for unstable OU process, i.e., for $\theta_n < 0$. The proofs for stable OU process (i.e., $\theta_n > 0$) and Wiener process (i.e., $\theta_n = 0$) follow the similar steps.

Define the σ-field

$$
\mathcal{N}_{n,t} = \sigma(X_{n,s} : 0 \leq s \leq t), \qquad (4.85)
$$

which is the set of events whose occurrence are determined by the realization of the process $\{X_{n,s} : 0 \leq s \leq t\}$. The right continuous filtration $\{\mathcal{N}^+_{n,t}, t \geq 0\}$ is defined by

$$\mathcal{N}^+_{n,t} = \cup_{s>t} \mathcal{N}_{n,s}. \tag{4.86}$$

In causal sampling policies, each sampling time is a stopping time with respect to the filtration $\{\mathcal{N}^+_{n,t}, t \geq 0\}$, i.e., [76]

$$\{S_{n,i} \leq t\} \in \mathcal{N}^+_{n,t}, \forall t \geq 0. \tag{4.87}$$

Let the sampling and scheduling policy $\pi_n = (S_{n,1}, S_{n,2}, \ldots)$ in (5.11) satisfy two conditions: (i) Each sampling policy $\pi_n \in \Pi_n$ satisfies (4.87) for all i. (ii) The sequence of inter-sampling times $\{T_{n,i} = S_{n,i+1} - S_{n,i}, i = 0, 1, \ldots\}$ forms a *regenerative process* [77, Section 6.1]: There exists an increasing sequence $0 \leq k_1 < k_2 < \ldots$ of almost surely finite random integers such that the post-k_j process $\{T_{n,k_j+i}, i = 0, 1, \ldots\}$ has the same distribution as the post-k_0 process $\{T_{n,k_0+i}, i = 0, 1, \ldots\}$ and is independent of the pre-k_j process $\{T_{n,i}, i = 0, 1, \ldots, k_j - 1\}$; further, we assume that $\mathbb{E}[k_{j+1} - k_j] < \infty$, $\mathbb{E}[S_{n,k_1}] < \infty$, and $0 < \mathbb{E}[S_{n,k_{j+1}} - S_{n,k_j}] < \infty,\ j = 1, 2, \ldots$.

We will prove Proposition 4.1 in three steps: First , we show that it is better not to sample when no channel is free. Second, we decompose the MDP in (5.11) into a series of mutually independent per-sample MDPs. Finally, we solve the per-sample MDP analytically.

The Gauss-Markov process $O_{n,t}$ in (4.34) is the solution to the following SDE

$$dO_{n,t} = -\theta_n O_{n,t} dt + \sigma_n dW_{n,t}. \tag{4.88}$$

In addition, the infinitesimal generator of $O_{n,t}$ is [92, Eq. A1.22]

$$\mathcal{G} = -\theta_n u \frac{\partial}{\partial u} + \frac{\sigma_n^2}{2} \frac{\partial^2}{\partial u^2}. \tag{4.89}$$

In Appendix 4.B, by using the strong Markov property of the Gauss-Markov process and the orthogonality principle of MMSE estimation, we have shown that it is better not to take a sample before the previous sample is delivered. Hence, the sampling time and the transmission starting time are equal to each other. By this, let us consider a sub-class of sampling and scheduling policies $\Pi_{n,1} \subset \Pi_n$ such that each sample is generated and sent out after all previous samples are delivered, i.e.,

$$\Pi_{n,1} = \{\pi_n \in \Pi_n : S_{n,i} = G_{n,i} \geq D_{n,i-1} \text{ for all } i\}.$$

For any policy $\pi_n \in \Pi_{n,1}$, the *information* used for determining $S_{n,i}$ includes: (i) the history of signal values $(X_{n,t} : t \in [0, S_{n,i}])$ and (ii) the service times $(Y_{n,1}, \ldots, Y_{n,i-1})$ of previous samples. Let us define the σ-fields $\mathcal{F}_{n,t} = \sigma(X_{n,s} : s \in [0,t])$ and $\mathcal{F}^+_{n,t} = \cap_{r>t}\mathcal{F}_{n,r}$. Then, $\{\mathcal{F}^+_{n,t}, t \geq 0\}$ is the filtration (i.e., a non-decreasing and right-continuous family of σ-fields) of the Gauss-Markov process $X_{n,t}$. Given the service times $(Y_{n,1}, \ldots, Y_{n,i-1})$ of previous samples, $S_{n,i}$ is a *stopping time* with respect to the filtration $\{\mathcal{F}^+_{n,t}, t \geq 0\}$ of the Gauss-Markov process $X_{n,t}$, that is

$$[\{S_{n,i} \leq t\} | Y_{n,1}, \ldots, Y_{n,i-1}] \in \mathcal{F}^+_{n,t}. \tag{4.90}$$

Hence, the policy space $\Pi_{n,1}$ can be expressed as

$$\begin{aligned}\Pi_{n,1} =& \{S_{n,i} : [\{S_{n,i} \leq t\} | Y_{n,1}, \ldots, Y_{n,i-1}] \in \mathcal{F}^+_{n,t}, \\ & T_{n,i} \text{ is a regenerative process}\}.\end{aligned} \tag{4.91}$$

Let $Z_{n,i} = S_{n,i+1} - D_{n,i} \geq 0$ represent the *waiting time* between the delivery time $D_{n,i}$ of the i-th sample and the generation time $S_{n,i+1}$ of the $(i+1)$-th sample. Then,

$$S_{n,i} = \sum_{j=0}^{i-1} (Y_{n,j} + Z_{n,j}), \tag{4.92}$$

$$D_{n,i} = \sum_{j=0}^{i-1} (Y_{n,j} + Z_{n,j}) + Y_{n,i} \tag{4.93}$$

for each $i = 1, 2, \ldots$. Given $(Y_{n,0}, Y_{n,1}, \ldots)$, $(S_{n,1}, S_{n,2}, \ldots)$ is uniquely determined by $(Z_{n,0}, Z_{n,1}, \ldots)$. Hence, one can also use $\pi = (Z_{n,0},$
$Z_{n,1}, \ldots)$ to represent a sampling and scheduling policy.

By using (4.6), (4.92), (4.93), and the assumption that the inter-sampling times follow a regenerative process, the MDP in (5.11) can be transformed as the following.

$$\bar{m}_{n,\text{opt}} = \inf_{\pi_n \in \Pi_{n,1}} \lim_{t\to\infty} \frac{\sum_{i=1}^{t} \mathbb{E}\left[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}+Y_{n,i+1}} w_n \varepsilon_n^2(s)ds\right] + \lambda \mathbb{E}[Y_{n,i+1}]}{\sum_{i=1}^{t} \mathbb{E}[Y_{n,i}+Z_{n,i}]}. \tag{4.94}$$

In order to solve (4.94), let consider the following MDP with parameter $k \geq 0$:

$$h(k) = \inf_{\pi_n \in \Pi_{n,1}} \lim_{t\to\infty} \sum_{i=1}^{t} \mathbb{E}\left[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}+Y_{n,i+1}} w_n \varepsilon_n^2(s)ds + \lambda Y_{n,i+1} - k(Y_{n,i}+Z_{n,i})\right], \tag{4.95}$$

where $h(k)$ is the optimum value of (4.95). Similar to the Dinkelbach's method [90] for non-linear fractional programming, the following lemma holds for the MDP in (4.94):

Lemma 4.5 *[34], [35] The following assertions are true:*

(a). $\bar{m}_{n,opt} \gtreqless k$ if and only if $h(k) \gtreqless 0$.

(b). If $h(k) = 0$, the solutions to (4.94) *and* (4.95) *are identical.*

Hence, the solution to (4.94) can be obtained by solving (4.95) and finding $k = \bar{m}_{n,\text{opt}}$ for which $h(\bar{m}_{n,\text{opt}}) = 0$.

Define

$$\beta_n = \bar{m}_{n,\text{opt}}. \tag{4.96}$$

In this sequel, we need to introduce the following lemma.

Lemma 4.6 *For any $\beta_n \geq 0$, it holds that*

$$
\begin{aligned}
&\mathbb{E}\Big[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}+Y_{n,i+1}} w_n\varepsilon_n^2(s)ds + \lambda Y_{n,i+1} - \beta_n(Y_{n,i}+Z_{n,i})\Big] \\
=&\mathbb{E}\Big[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}+Y_{n,i+1}} (w_n\varepsilon_n^2(s) - \beta_n)ds + w_n\gamma_n O^2_{Y_{n,i}+Z_{n,i}}\Big] \\
&+ w_n\frac{\sigma_n^2}{2\theta_n}(\mathbb{E}[Y_{n,i}] - \gamma_n) + \lambda\mathbb{E}[Y_{n,i+1}] - \beta_n\mathbb{E}[Y_{n,i}],
\end{aligned} \tag{4.97}
$$

where γ_n is a constant defined as

$$
\gamma_n = \frac{1}{2\theta_n}\mathbb{E}[1 - e^{-2\theta_n Y_{n,i}}]. \tag{4.98}
$$

Proof 4.7 *We can write* (4.97) *as*

$$
\begin{aligned}
&\mathbb{E}\Big[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}+Y_{n,i+1}} w_n\varepsilon_n^2(s)ds + \lambda Y_{n,i+1} - \beta_n(Y_{n,i}+Z_{n,i})\Big] \\
=&\mathbb{E}\Big[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}} w_n\varepsilon_n^2(s)ds\Big] + \mathbb{E}\Big[\int_{Y_{n,i}+Z_{n,i}}^{Y_{n,i}+Z_{n,i}+Y_{n,i+1}} w_n\varepsilon_n^2(s)ds\Big] \\
&+ \lambda\mathbb{E}[Y_{n,i+1}] - \beta_n\mathbb{E}(Y_{n,i}+Z_{n,i}).
\end{aligned} \tag{4.99}
$$

In order to prove Lemma 4.6, we need to compute the second term in (4.99)*. First, let us introduce the following lemma which is more general than Lemma 5 in [34] and works for any OU process irrespective of the signal structure, i.e., the value of parameter θ. By using Dynkin's formula and optional stopping theorem, we get the following useful lemma.*

Lemma 4.7 *[34] Let $\tau \geq 0$ be a stopping time of the OU process $O_{n,t}$ with $\mathbb{E}[O_\tau^2] < \infty$, then*

$$
\mathbb{E}\left[\int_0^\tau O_{n,t}^2 dt\right] = \mathbb{E}\left[\frac{\sigma_n^2}{2\theta_n}\tau - \frac{1}{2\theta_n}O_{n,\tau}^2\right]. \tag{4.100}
$$

If, in addition, τ is the first exit time of a bounded set, then

$$\mathbb{E}\left[\tau\right] = \mathbb{E}[R_{n,1}(O_{n,\tau})], \tag{4.101}$$

$$\mathbb{E}\left[\int_0^\tau O_{n,t}^2 dt\right] = \mathbb{E}[R_{n,2}(O_{n,\tau})], \tag{4.102}$$

where $R_{n,1}(\cdot)$ and $R_{n,2}(\cdot)$ are defined in (4.37) *and* (4.38)*, respectively.*

Proof 4.8 *We first prove* (4.100)*. It is known that the OU process $O_{n,t}$ is a Feller process [96, Section 5.5]. By using a property of Feller process in [96, Theorem 3.32], we get that*

$$\begin{aligned} &O_{n,t}^2 - \int_0^t \mathcal{G}(O_{n,s}^2)ds \\ =&O_{n,t}^2 - \int_0^t (-\theta_n O_{n,s} 2O_{n,s} + \sigma_n^2)ds \\ =&O_{n,t}^2 - \sigma_n^2 t + 2\theta_n \int_0^t O_{n,s}^2 ds \end{aligned} \tag{4.103}$$

is a martingale. According to [76], the minimum of two stopping times is a stopping time and constant times are stopping times. Hence, $t\wedge\tau$ is a bounded stopping time for every $t \in [0, \infty)$, where $x\wedge y = \min\{x, y\}$. Then, by [76, Theorem 8.5.1], for all $t \in [0, \infty)$

$$\mathbb{E}\left[\int_0^{t\wedge\tau} O_{n,s}^2 ds\right] = \mathbb{E}\left[\frac{{\sigma_n}^2}{2\theta_n}(t\wedge\tau)\right] - \mathbb{E}\left[\frac{1}{2\theta_n}O_{n,t\wedge\tau}^2\right]. \tag{4.104}$$

Because $\mathbb{E}\left[\int_0^{t\wedge\tau} O_{n,s}^2 ds\right]$ and $\mathbb{E}[t\wedge\tau]$ are positive and increasing with respect to t, by using the monotone convergence theorem [76, Theorem 1.5.5], we get

$$\lim_{t\to\infty} \mathbb{E}\left[\int_0^{t\wedge\tau} O_{n,s}^2 ds\right] = \mathbb{E}\left[\int_0^{\tau} O_{n,s}^2 ds\right], \tag{4.105}$$

$$\lim_{t\to\infty} \mathbb{E}[(t\wedge\tau)] = \mathbb{E}[\tau]. \tag{4.106}$$

In addition, according to Doob's maximal inequality [76], we get that

$$\mathbb{E}\left[\sup_{0\le s\le\tau} O_{n,s}^2\right] \le 4\mathbb{E}[O_{n,\tau}^2] < \infty. \tag{4.107}$$

Because $0 \leq O^2_{n,t\wedge\tau} \leq \sup_{0\leq s\leq\tau} O^2_{n,s}$ *for all* t *and* (4.107) *implies that* $\sup_{0\leq s\leq\tau} O^2_{n,s}$ *is integratable, by invoking the dominated convergence theorem [76, Theorem 1.5.6], we have*

$$\lim_{t\to\infty} \mathbb{E}\left[O^2_{n,t\wedge\tau}\right] = \mathbb{E}\left[O^2_{n,\tau}\right]. \tag{4.108}$$

Combining (4.105)-(4.108), (4.100) *is proven.*

Next, we prove (4.101) *and* (4.102). *By using the solution of the ODE in Appendix 4.C, one can show that* $R_{n,1}(v)$ *in* (4.37) *is the solution to the following ODE*

$$\frac{\sigma_n^2}{2} R''_{n,1}(v) - \theta_n v R'_{n,1}(v) = 1, \tag{4.109}$$

and $R_{n,2}(v)$ *in* (4.38) *is the solution to the following ODE*

$$\frac{\sigma_n^2}{2} R''_{n,2}(v) - \theta_n v R'_{n,2}(v) = v^2. \tag{4.110}$$

In addition, $R_{n,1}(v)$ *and* $R_{n,2}(v)$ *are twice continuously differentiable. According to Dynkin's formula in [75, Theorem 7.4.1 and the remark afterwards], because the initial value of* $O_{n,t}$ *is* $O_{n,0} = 0$, *if* τ *is the first exit time of a bounded set, then*

$$\mathbb{E}_0[R_{n,1}(O_{n,\tau})] = R_{n,1}(0) + \mathbb{E}_0\left[\int_0^\tau 1 ds\right] = R_{n,1}(0) + \mathbb{E}_0[\tau], \tag{4.111}$$

$$\mathbb{E}_0[R_{n,2}(O_{n,\tau})] = R_{n,2}(0) + \mathbb{E}_0\left[\int_0^\tau O^2_{n,s} ds\right]. \tag{4.112}$$

Because $R_{n,1}(0) = R_{n,2}(0) = 0$, (4.101) *and* (4.102) *follow. This completes the proof.*

By using Lemma 4.7, we can write

$$\begin{aligned}&\mathbb{E}\bigg[\int_{Y_{n,i}+Z_{n,i}}^{Y_{n,i}+Z_{n,i}+Y_{n,i+1}} w_n \varepsilon_n^2(s) ds\bigg]\\ =&w_n \frac{\sigma_n^2}{2\theta_n}\mathbb{E}[Y_{n,i+1}] - w_n \frac{1}{2\theta_n}\mathbb{E}\big[O^2_{n,Y_{n,i}+Z_{n,i}+Y_{n,i+1}} - O^2_{n,Y_{n,i}+Z_{n,i}}\big],\end{aligned} \tag{4.113}$$

where

$$\begin{aligned}
&\mathbb{E}\left[O^2_{n,Y_{n,i}+Z_{n,i}+Y_{n,i+1}} - O^2_{n,Y_{n,i}+Z_{n,i}}\right] \\
=&\mathbb{E}\left[\left(O_{n,Y_{n,i}+Z_{n,i}} e^{-\theta_n Y_{n,i+1}} + \frac{\sigma_n}{\sqrt{-2\theta_n}} e^{-\theta_n Y_{n,i+1}} W_{1-e^{2\theta_n Y_{n,i+1}}}\right)^2 \right.\\
&\left. \qquad - O^2_{n,Y_{n,i}+Z_{n,i}}\right],
\end{aligned} \tag{4.114}$$

$$\begin{aligned}
=&\mathbb{E}\left[O^2_{n,Y_{n,i}+Z_{n,i}}(e^{-2\theta_n Y_{n,i+1}}-1) - \frac{\sigma_n^2}{2\theta_n} e^{-2\theta_n Y_{n,i+1}} W^2_{n,1-e^{2\theta_n Y_{n,i+1}}}\right] \\
&+ \mathbb{E}\left[2O_{n,Y_{n,i}+Z_{n,i}} e^{-\theta_n Y_{n,i+1}} \frac{\sigma_n}{\sqrt{-2\theta_n}} e^{-\theta_n Y_{n,i+1}} W_{n,1-e^{2\theta_n Y_{n,i+1}}}\right].
\end{aligned} \tag{4.115}$$

Because $Y_{n,i+1}$ is independent of $O_{n,Y_{n,i}+Z_{n,i}}$ and $W_{n,t}$, we have

$$\mathbb{E}\left[O^2_{n,Y_{n,i}+Z_{n,i}}(e^{-2\theta_n Y_{n,i+1}}-1)\right] = \mathbb{E}\left[O^2_{n,Y_{n,i}+Z_{n,i}}\right]\mathbb{E}\left[e^{-2\theta_n Y_{n,i+1}}-1\right], \tag{4.116}$$

and

$$\begin{aligned}
&\mathbb{E}\left[2O_{n,Y_{n,i}+Z_{n,i}} e^{-\theta_n Y_{n,i+1}} \frac{\sigma_n}{\sqrt{-2\theta_n}} e^{-\theta_n Y_{n,i+1}} W_{n,1-e^{2\theta_n Y_{n,i+1}}}\right] \\
=&\mathbb{E}\left[2O_{n,Y_{n,i}+Z_{n,i}}\right]\mathbb{E}\left[e^{-\theta_n Y_{n,i+1}} \frac{\sigma_n}{\sqrt{-2\theta_n}} e^{-\theta_n Y_{n,i+1}} W_{n,1-e^{2\theta_n Y_{n,i+1}}}\right] \\
\overset{(a)}{=}&\mathbb{E}\left[2O_{n,Y_{n,i}+Z_{n,i}}\right] \\
&\mathbb{E}\left[\mathbb{E}\left[e^{-\theta_n Y_{n,i+1}} \frac{\sigma_n}{\sqrt{-2\theta_n}} e^{-\theta Y_{n,i+1}} W_{n,1-e^{2\theta_n Y_{n,i+1}}} \middle| Y_{n,i+1}\right]\right],
\end{aligned} \tag{4.117}$$

where Step (a) is due to the law of iterated expectations. Because $\mathbb{E}[W_{n,t}] = 0$ for all constant $t \geq 0$, it holds for all realizations of $Y_{n,i+1}$ that

$$\mathbb{E}\left[e^{-\theta_n Y_{n,i+1}} \frac{\sigma_n}{\sqrt{-2\theta_n}} e^{-\theta_n Y_{n,i+1}} W_{n,1-e^{2\theta_n Y_{n,i+1}}} \middle| Y_{n,i+1}\right] = 0. \tag{4.118}$$

Hence,

$$\mathbb{E}\left[2O_{n,Y_{n,i}+Z_{n,i}}e^{-\theta_n Y_{n,i+1}}\frac{\sigma_n}{\sqrt{-2\theta_n}}e^{-\theta_n Y_{n,i+1}}W_{n,1-e^{2\theta_n Y_{n,i+1}}}\right]=0. \tag{4.119}$$

In addition,

$$\begin{aligned}
&\mathbb{E}\left[\frac{\sigma_n^2}{2\theta_n}e^{-2\theta_n Y_{n,i+1}}W^2_{n,1-e^{2\theta_n Y_{n,i+1}}}\right]\\
&\overset{(a)}{=}\frac{\sigma_n^2}{2\theta_n}\mathbb{E}\left[\mathbb{E}\left[e^{-2\theta_n Y_{n,i+1}}W^2_{n,1-e^{2\theta_n Y_{n,i+1}}}\middle|Y_{n,i+1}\right]\right]\\
&\overset{(b)}{=}\frac{\sigma_n^2}{2\theta_n}\mathbb{E}\left[e^{-2\theta_n Y_{n,i+1}}-1\right],
\end{aligned} \tag{4.120}$$

where Step (a) is due to the law of iterated expectations and Step (b) is due to $\mathbb{E}[W_{n,t}^2]=t$ *for all constant* $t\geq 0$*. Hence,*

$$\begin{aligned}
&\mathbb{E}\big[O^2_{n,Y_{n,i}+Z_{n,i}+Y_{n,i+1}}-O^2_{n,Y_{n,i}+Z_{n,i}}\big]\\
&=\mathbb{E}\big[O^2_{n,Y_{n,i}+Z_{n,i}}\big]\mathbb{E}[e^{-2\theta_n Y_{n,i+1}}-1]+\frac{\sigma_n^2}{2\theta_n}\mathbb{E}[1-e^{-2\theta_n Y_{n,i+1}}].
\end{aligned} \tag{4.121}$$

By using (4.121) *in* (4.113)*, we get that*

$$\begin{aligned}
&\mathbb{E}\bigg[\int_{Y_{n,i}+Z_{n,i}}^{Y_{n,i}+Z_{n,i}+Y_{n,i+1}}w_nO_s^2ds\bigg]\\
=&w_n\frac{\sigma_n^2}{2\theta_n}\mathbb{E}[Y_{n,i+1}]-w_n\frac{1}{2\theta_n}\mathbb{E}\big[O^2_{Y_{n,i}+Z_{n,i}}\big]\mathbb{E}[e^{-2\theta_n Y_{n,i+1}}-1]\\
&-w_n\frac{\sigma_n^2}{4\theta_n^2}\mathbb{E}[1-e^{-2\theta_n Y_{n,i+1}}],\\
=&w_n\frac{\sigma_n^2}{2\theta_n}\big\{\mathbb{E}[Y_{n,i+1}]-\gamma_n\big\}+w_n\gamma_n\mathbb{E}\left[O^2_{Y_{n,i}+Z_{n,i}}\right],
\end{aligned} \tag{4.122}$$

where γ_n is defined in (4.98)*. Substituting* (4.122) *into* (4.99) *yields*

$$\begin{aligned}
&\mathbb{E}\bigg[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}+Y_{n,i+1}} w_n O_s^2 ds + \lambda Y_{n,i+1} - \beta_n[Y_{n,i}+Z_{n,i}]\bigg]\\
=&\mathbb{E}\bigg[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}} w_n O_s^2 ds\bigg] + w_n\frac{\sigma_n^2}{2\theta_n}\big\{\mathbb{E}[Y_{n,i+1}] - \gamma_n\big\}\\
&+ w_n\gamma_n\mathbb{E}\left[O_{Y_{n,i}+Z_{n,i}}^2\right] + \lambda\mathbb{E}[Y_{n,i+1}] - \beta_n\mathbb{E}[Y_{n,i}+Z_{n,i}],
\end{aligned} \tag{4.123}$$

$$\begin{aligned}
=&\mathbb{E}\bigg[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}} (w_n O_s^2 - \beta_n) ds + w_n\gamma_n O_{Y_{n,i}+Z_{n,i}}^2\bigg]\\
&+ w_n\frac{\sigma_n^2}{2\theta_n}\big\{\mathbb{E}[Y_{n,i+1}] - \gamma_n\big\} - \beta_n\mathbb{E}[Y_{n,i}] + \lambda\mathbb{E}[Y_{n,i+1}],
\end{aligned} \tag{4.124}$$

from which (4.97) *follows.*

For any $s \geq 0$, define the σ-fields $\mathcal{F}_{n,t}^s = \sigma(O_{n,s+r} - O_{n,s} : r \in [0,t])$ and the right-continuous filtration $\mathcal{F}_{n,t}^{s+} = \cap_{r>t}\mathcal{F}_{n,r}^s$. Then, $\{\mathcal{F}_{n,t}^{s+}, t \geq 0\}$ is the filtration of the time-shifted OU process $\{O_{n,s+t} - O_{n,s}, t \in [0,\infty)\}$. Define $\mathfrak{M}_{n,s}$ as the set of integrable stopping times of $\{O_{n,s+t} - O_{n,s}, t \in [0,\infty)\}$, i.e.,

$$\mathfrak{M}_{n,s} = \{\tau \geq 0 : \{\tau \leq t\} \in \mathcal{F}_{n,t}^{s+}, \mathbb{E}[\tau] < \infty\}. \tag{4.125}$$

By using a sufficient statistic of (4.95), we can obtain

Lemma 4.8 *An optimal solution* $(Z_{n,0}, Z_{n,1}, \ldots)$ *to* (4.95) *satisfies*

$$\inf_{Z_{n,i}\in\mathfrak{M}_{Y_{n,i}}} \mathbb{E}\left[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}} (w_n\varepsilon_n^2(s) - \beta_n)ds + \gamma_n O_{Y_{n,i}+Z_{n,i}}^2 \middle| O_{Y_{n,i}}, Y_{n,i}\right], \tag{4.126}$$

where $\beta_n \geq 0$ and $\gamma_n \geq 0$ are defined in (4.96) *and* (4.98)*, respectively.*

Proof 4.9 *Because the $Y_{n,i}$'s are* i.i.d.*,* (4.123) *is determined by the control decision $Z_{n,i}$ and the information $(Y_{n,i}, O_{Y_{n,i}})$. Hence, $(Y_{n,i},$
$O_{Y_{n,i}})$ is a* sufficient statistic *for determining $Z_{n,i}$ in* (4.95)*. Therefore, there exists an optimal policy $(Z_{n,0}, Z_{n,1}, \ldots)$ to* (4.95)*, in which $Z_{n,i}$ is determined based on only $(Y_{n,i}, O_{Y_{n,i}})$. By*

this, (4.95) *is decomposed into a sequence of per-sample MDPs, given by* (4.126)*. This completes the proof.*

Next, we solve (4.126) by using free-boundary method for optimal stopping problems. Let consider an OU process $V_{n,t}$ with initial state $V_{n,0} = v$ and $\mu_n = 0$. Define the σ-fields $\mathcal{F}_{n,t}^{V} = \sigma(V_{n,s} : s \in [0,t])$, $\mathcal{F}_{n,t}^{V+} = \cap_{r>t}\mathcal{F}_{n,r}^{V}$, and the filtration $\{\mathcal{F}_{n,t}^{V+}, t \geq 0\}$ associated to $\{V_{n,t}, t \geq 0\}$. Define $\mathfrak{M}_V$ as the set of integrable stopping times of $\{V_{n,t}, t \in [0,\infty)\}$, i.e.,

$$\mathfrak{M}_V = \{\tau \geq 0 : \{\tau \leq t\} \in \mathcal{F}_{n,t}^{V+}, \mathbb{E}[\tau] < \infty\}. \tag{4.127}$$

Our goal is to solve the following optimal stopping problem for any given initial state $v \in \mathbb{R}$ and for any $\beta_n > 0$

$$\sup_{\tau \in \mathfrak{M}_V} \mathbb{E}_v \left[-w_n \gamma_n V_{n,\tau}^2 - \int_0^{\tau} (w_n V_{n,s}^2 - \beta_n) ds \right], \tag{4.128}$$

where $\mathbb{E}_v[\cdot]$ is the conditional expectation for given initial state $V_{n,0} = v$, where the supremum is taken over all stopping times τ of $V_{n,t}$, and γ_n is defined in (4.98). In this subsection, we focus on the case that β_n in (4.128) satisfies $\frac{\sigma_n^2}{2\theta_n}\mathbb{E}[1 - e^{-2\theta_n Y_{n,i}}] \leq \beta_n < \infty$.

In order to solve (4.128) for $\theta_n < 0$, we first find a candidate solution to (4.128) by solving a free boundary problem; then we prove that the free boundary solution is indeed the value function of (4.128):

The general optimal stopping theory in Chapter I of [74] tells us that the following guess of the stopping time should be optimal for Problem (4.128):

$$\tau_* = \inf\{t \geq 0 : |V_{n,t}| \geq v_*\}, \tag{4.129}$$

where $v_* \geq 0$ is the optimal stopping threshold to be found. Observe that in this guess, the continuation region $(-v_*, v_*)$ is assumed symmetric around zero. This is because the OU process is symmetric, i.e., the process $\{-V_{n,t}, t \geq 0\}$ is also an OU process started at $-V_{n,0} = -v$. Similarly, we can also argue that the value function of problem (4.128) should be even. According to [74, Chapter 8], and [75, Chapter 10], the value function and the optimal stopping threshold v_* should satisfy the following free boundary problem:

$$\frac{\sigma_n^2}{2}H''(v) - \theta_n v H'(v) = w_n v^2 - \beta_n, \quad v \in (-v_*, v_*), \tag{4.130}$$

$$H(\pm v_*) = -w_n \gamma_n v_*^2, \tag{4.131}$$

$$H'(\pm v_*) = \mp 2 w_n \gamma_n v_*. \tag{4.132}$$

In this sequel, we solve (4.130) to find $H(v)$.

We need to use the following indefinite integrals to solve (4.130) that can be obtained by [99, Sec. 15.3.1, (Eq. 36)], [85, Sec. 3.478 (Eq. 3), 8.250 (Eq. 1,4)]. Let $\theta_n = -\rho_n$.

$$\begin{aligned}\int \frac{2}{\sigma_n^2} w_n e^{-\frac{\theta_n}{\sigma_n^2}v^2} v^2 dv &= \frac{\sqrt{\pi} w_n \sigma_n}{2{\theta_n}^{\frac{3}{2}}} \operatorname{erf}\left(\frac{\sqrt{\theta_n}}{\sigma_n} v\right) - \frac{w_n v}{\theta_n} e^{-\frac{\theta_n}{{\sigma_n}^2}v^2} + C_1, \\ &= \frac{\sqrt{\pi} w_n \sigma_n}{-2j\rho_n\sqrt{\rho_n}} \operatorname{erf}\left(\frac{j\sqrt{\rho_n}}{\sigma_n} v\right) + \frac{w_n v}{\rho_n} e^{\frac{\rho_n}{{\sigma_n}^2}v^2} + C_1, \end{aligned} \tag{4.133}$$

$$\begin{aligned}\int \frac{2}{\sigma_n^2} \beta_n e^{-\frac{\theta_n}{\sigma_n^2}v^2} dv &= \frac{\sqrt{\pi}\beta_n}{\sigma_n\sqrt{\theta_n}} \operatorname{erf}\left(\frac{\sqrt{\theta_n}}{\sigma_n} v\right) + C_2, \\ &= \frac{\sqrt{\pi}\beta_n}{j\sigma_n\sqrt{\rho_n}} \operatorname{erf}\left(\frac{j\sqrt{\rho_n}}{\sigma_n} v\right) + C_2, \end{aligned} \tag{4.134}$$

$$\begin{aligned}&\int \operatorname{erf}\left(\frac{\sqrt{\theta_n}}{\sigma_n} v\right) e^{\frac{\theta_n}{{\sigma_n}^2}v^2} dv \\ &= \frac{\sigma_n}{\sqrt{\theta_n}\sqrt{\pi}} \frac{\theta_n}{{\sigma_n}^2} v^2 {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta_n}{\sigma_n^2} v^2\right) + C, \\ &= -\frac{\sigma_n}{j\sqrt{\rho_n}\sqrt{\pi}} \frac{\rho_n}{{\sigma_n}^2} v^2 {}_2F_2\left(1, 1; \frac{3}{2}, 2; -\frac{\rho_n}{\sigma_n^2} v^2\right) + C, \end{aligned} \tag{4.135}$$

$$\begin{aligned}&\int \left(\frac{\sqrt{\pi} w_n \sigma_n}{2{\theta_n}^{\frac{3}{2}}} - \frac{\sqrt{\pi}\beta_n}{\sigma_n\sqrt{\theta_n}}\right) \operatorname{erf}\left(\frac{\sqrt{\theta_n}}{\sigma_n} v\right) e^{\frac{\theta_n}{{\sigma_n}^2}v^2} dv \\ &= \left(\frac{w_n}{2\theta_n} - \frac{\beta_n}{{\sigma_n}^2}\right) v^2 \; {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta_n}{\sigma_n^2} v^2\right) + C_4, \\ &= \left(-\frac{w_n}{2\rho_n} - \frac{\beta_n}{{\sigma_n}^2}\right) v^2 \; {}_2F_2\left(1, 1; \frac{3}{2}, 2; -\frac{\rho_n}{\sigma_n^2} v^2\right) + C_4, \end{aligned} \tag{4.136}$$

$$\int C_3 e^{\frac{\theta_n}{\sigma_n{}^2}v^2} dv = C_5 \text{erfi}\left(\frac{\sqrt{\theta_n}}{\sigma_n} v\right) + C_6,$$

$$= C_5 \text{erfi}\left(-j\frac{\sqrt{\rho_n}}{\sigma_n} v\right) + C_6, \tag{4.137}$$

$$-\int \frac{w_n v}{\theta_n} dv = w_n \frac{v^2}{2\rho_n} + C_7, \tag{4.138}$$

where erf($\cdot$) and erfi($\cdot$) are the error function and imaginary error functions, respectively. Hence, $H(v)$ is given by

$$H(v) = \left(-\frac{w_n}{2\rho_n} - \frac{\beta_n}{\sigma_n^2}\right) v^2 {}_2F_2\left(1, 1; \frac{3}{2}, 2; -\frac{\rho_n}{\sigma_n^2} v^2\right) + C_1 \text{erfi}\left(j\frac{\sqrt{\rho_n}}{\sigma_n} v\right)$$
$$+ w_n \frac{v^2}{2\rho_n} + C_2,\ v \in (-v_*, v_*), \tag{4.139}$$

where C_1 and C_2 are constants to be found for satisfying (4.131)-(4.132), and erfi(x) is the imaginary error function, i.e.,

$$\text{erfi}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{t^2} dt. \tag{4.140}$$

Because $H(v)$ should be even but erfi(x) is odd, we should choose $C_1 = 0$. Further, in order to satisfy the boundary condition (4.131), C_2 is chosen as

$$C_2 = -\frac{1}{2\rho_n} \mathbb{E}\left(e^{2\rho_n Y_{n,i}}\right) v_*^2 + \left(\frac{1}{2\rho_n} + \frac{\beta_n}{\sigma_n^2}\right) {}_2F_2\left(1, 1; \frac{3}{2}, 2; -\frac{\rho_n}{\sigma_n^2} v_*^2\right) v_*^2, \tag{4.141}$$

where we have used (4.98). With this, the expression of $H(v)$ is obtained in the continuation region $(-v_*,\ v_*)$. In the stopping region $|v| \geq v_*$, the stopping time in (4.129) is simply $\tau_* = 0$, because $|V_{n,0}| = |v| \geq v_*$. Hence, if $|v| \geq v_*$, the objective value achieved by the sampling time (4.129) is

$$\mathbb{E}_v\left[-\gamma_n w_n v^2 - \int_0^0 (w_n V_{n,s}^2 - \beta_n) ds\right] = -\gamma_n w_n v^2. \tag{4.142}$$

Combining (4.139)-(4.142), we obtain a candidate of the value function for (4.128):

$$H(v) = \begin{cases} w_n \frac{v^2}{2\rho_n} - \left(\frac{w_n}{2\rho_n} + \frac{\beta_n}{\sigma_n{}^2}\right) {}_2F_2\big(1, 1; \frac{3}{2}, 2; -\frac{\rho_n}{\sigma_n^2} v^2\big) v^2 + C_2, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if} \quad |v| < v_*, \\ -\gamma_n w_n v^2, \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if} \quad |v| \geq v_*. \end{cases} \tag{4.143}$$

Next, we find a candidate value of the optimal stopping threshold v_*. By taking the gradient of $H(v)$, we get

$$H'(v) = w_n \frac{v}{\rho_n} - \left(\frac{w_n \sigma_n}{j\rho_n \sqrt{\rho_n}} + \frac{2\beta_n}{j\sigma_n \sqrt{\rho_n}}\right) F\left(j \frac{\sqrt{\rho}_n}{\sigma_n} v\right), \quad v \in (-v_*, v_*), \tag{4.144}$$

where

$$F(x) = e^{x^2} \int_0^x e^{-t^2} dt. \tag{4.145}$$

The boundary condition (4.132) implies that v_* is the root of

$$w_n \frac{v}{\rho_n} - \left(\frac{w_n \sigma_n}{j\rho_n \sqrt{\rho_n}} + \frac{2\beta_n}{j\sigma_n \sqrt{\rho_n}}\right) F\left(j \frac{\sqrt{\rho}_n}{\sigma_n} v\right) = -2\gamma_n w_n v. \tag{4.146}$$

Substituting (4.98) into (4.146), yields that v_* is the root of

$$-\left(w_n \frac{\sigma_n^2}{2\rho_n} + \beta_n\right) G\left(j \frac{\sqrt{\rho_n}}{\sigma_n} v\right) = -w_n \frac{\sigma_n^2}{2\rho_n} \mathbb{E}[e^{2\rho_n Y_{n,i}}], \tag{4.147}$$

where $G(\cdot)$ is defined in (4.18). By using (4.22) in (4.147), we get that

$$-\left(w_n \frac{\sigma_n^2}{2\rho_n} + \beta_n\right) K\left(\frac{\sqrt{\rho_n}}{\sigma_n} v\right) = -w_n \frac{\sigma_n^2}{2\rho_n} \mathbb{E}[e^{2\rho_n Y_{n,i}}], \tag{4.148}$$

where $K(\cdot)$ is defined in (4.19). Rearranging (4.148), we obtain the threshold as follows

$$v(\beta_n) = \frac{\sigma_n}{\sqrt{\rho_n}} K^{-1}\left(\frac{w_n \frac{\sigma_n^2}{2\rho_n} \mathbb{E}[e^{2\rho_n Y_{n,i}}]}{w_n \frac{\sigma_n^2}{2\rho_n} + \beta_n}\right) \tag{4.149}$$

Substituting $\rho_n = -\theta_n$ in (4.149), we get (4.24) for $\theta_n < 0$.

In addition, when $\theta_n \to 0$, (4.147) can be expressed as

$$\left(\sigma_n^2 - \frac{2\theta_n\beta_n}{w_n}\right) G\left(\frac{\sqrt{\theta_n}}{\sigma_n}v\right) = \sigma_n^2 \mathbb{E}\left[e^{-2\theta_n Y_{n,i}}\right], \tag{4.150}$$

The error function $\operatorname{erf}(x)$ has a Maclaurin series representation, given by

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\left[x - \frac{x^3}{3} + o(x^3)\right]. \tag{4.151}$$

Hence, the Maclaurin series representation of $G(x)$ in (4.18) is

$$G(x) = 1 + \frac{2x^2}{3} + o(x^2). \tag{4.152}$$

Let $x = \frac{\sqrt{\theta_n}}{\sigma_n}v$, we get

$$G\left(\frac{\sqrt{\theta_n}}{\sigma_n}v\right) = 1 + \frac{2}{3}\frac{\theta_n}{\sigma_n^2}v^2 + o(\theta). \tag{4.153}$$

In addition,

$$\mathbb{E}\left[e^{-2\theta_n Y_{n,i}}\right] = 1 - 2\theta_n \mathbb{E}[Y_{n,i}] + o(\theta_n). \tag{4.154}$$

Hence, (4.150) can be expressed as

$$\left(\sigma_n^2 - \frac{2\beta_n\theta_n}{\sigma_n^2 w_n}\right)\left[1 + \frac{2}{3}\frac{\theta_n}{\sigma_n^2}v^2 + o(\theta_n)\right] = \sigma_n^2(1 - 2\theta_n\mathbb{E}[Y_{n,i}] + o(\theta_n)). \tag{4.155}$$

Expanding (4.155), yields

$$\sigma_n^2 2\theta_n \mathbb{E}[Y_{n,i}] - \frac{2\beta_n\theta_n}{\sigma_n^2 w_n} + \frac{2}{3}\frac{\theta_n}{\sigma_n^2}v^2 + o(\theta_n) = 0. \tag{4.156}$$

Divided by θ and let $\theta \to 0$ on both sides of (4.156), yields

$$v^2 - \frac{1}{w_n}3(\beta_n - w_n\sigma_n^2\mathbb{E}[Y_{n,i}]) = 0. \tag{4.157}$$

Equation (4.157) has two roots $-(1/\sqrt{w_n})\sqrt{3(\beta_n - w_n\sigma_n^2\mathbb{E}[Y_{n,i}]}$), and $(1/\sqrt{w_n})$ $\sqrt{3(\beta_n - w_n\sigma_n^2\mathbb{E}[Y_{n,i}]}$).

If $v_* = -(1/\sqrt{w_n})\sqrt{3(\beta_n - w_n\sigma_n^2\mathbb{E}[Y_{n,i}]}$), the free boundary problem in (4.130)-(4.132) are invalid. Hence, the root of (4.157) is $v_* = (1/\sqrt{w_n})\sqrt{3(\beta_n - w_n\sigma_n^2\mathbb{E}[Y_{n,i}]}$), from which (4.24) follows for $\theta_n = 0$.

Verification of the Optimality of the Candidate Solution

Next, we use Itô's formula to verify the above candidate solution is indeed optimal, as stated in the following theorem:

Theorem 4.7 *If* $\frac{\sigma_n^2}{2\theta_n}\mathbb{E}[1 - e^{-2\theta_n Y_{n,i}}] \leq \beta_n < \infty$, *then for all* $v \in \mathbb{R}$, $H(v)$ *in* (4.143) *is the value function of the optimal stopping problem* (4.128). *In addition, the optimal stopping time for solving* (4.128) *is* τ_* *in* (4.129), *where* $v_* = v(\beta_n)$ *is given by* (4.24).

In order to prove Theorem 4.7, we need to establish the following properties of $H(v)$ in (4.143), for the case that $\frac{\sigma_n^2}{2\theta_n}\mathbb{E}[1 - e^{-2\theta_n Y_{n,i}}] \leq \beta_n < \infty$ is satisfied in (4.128):

Lemma 4.9 *[34]* $H(v) = \mathbb{E}_v\left[-\gamma_n V_{n,\tau_*}^2 - \int_0^{\tau_*}(V_{n,s}^2 - \beta_n)ds\right]$.

Lemma 4.10 *[34]* $H(v) \geq -\gamma_n v^2$ *for all* $v \in \mathbb{R}$.

A function $f(v)$ is said to be *excessive* for the process $V_{n,t}$ if

$$\mathbb{E}_v f(V_{n,t}) \leq f(v), \forall t \geq 0, v \in \mathbb{R}. \tag{4.158}$$

By using Itô's formula in stochastic calculus, we can obtain

Lemma 4.11 *[34] The function* $H(v)$ *is excessive for the process* $V_{n,t}$.

Now, we are ready to prove Theorem 4.7.

Proof 4.10 (Proof of Theorem 4.7) *In Lemmas 4.9-4.11, we have shown that* $H(v) = \mathbb{E}_v\left[-\gamma_n V_{n,\tau_*}^2 - \int_0^{\tau_*}(V_{n,s}^2 - \beta_n)ds\right]$, $H(v) \geq -\gamma_n v^2$, *and* $H(v)$ *is an excessive function. Moreover, from Lemma 4.9, we know that* $\mathbb{E}_v[\tau_*] < \infty$ *holds for all* $v \in \mathbb{R}$. *Hence,* $\mathbb{P}_v(\tau_* < \infty) = 1$ *for all* $v \in \mathbb{R}$. *These conditions and Theorem 1.11 in [74, Section 1.2] imply that* τ_* *is an optimal stopping time of* (4.128). *This completes the proof.*

Because (4.126) is a special case of (4.128), we can get from Theorem 4.7 that

Corollary 4.2 *If* $\frac{\sigma_n^2}{2\theta_n}\mathbb{E}[1-e^{-2\theta_n Y_{n,i}}] \leq \beta_n < \infty$, *then a solution to* (4.126) *is* $(Z_{n,1}(\beta_n), Z_{n,2}(\beta_n), \ldots)$, *where*

$$Z_{n,i}(\beta_n) = \inf\{t \geq 0 : |O_{n,Y_{n,i}+t}| \geq v(\beta_n)\}, \tag{4.159}$$

and $v(\beta_n)$ *is defined in* (4.24).

This concludes the proof.

4.D Proof of Theorem 4.1

If $\gamma = 0$, i.e., no sample from source is currently in service, from Proposition 4.1, we get that the optimal sampling policy of (5.11) is a threshold policy which is given by (4.55). Given Proposition 4.1, for an instantaneous estimation error $|\varepsilon_n(t) = \varepsilon|$, it is optimal not to schedule source n if

$$|\varepsilon| < v_n(\bar{m}_n(\lambda)), \tag{4.160}$$

where

$$\bar{m}_n(\lambda) = \frac{\mathbb{E}\left[\int_{D_{n,i}(\bar{m}_n(\lambda))}^{D_{n,i+1}(\bar{m}_n(\lambda))} w_n \varepsilon_n^2(s)ds\right] + \lambda\mathbb{E}[Y_{n,i+1}]}{\mathbb{E}[D_{n,i+1}(\bar{m}_n(\lambda)) - D_{n,i}(\bar{m}_n(\lambda))]}, \tag{4.161}$$

and $\bar{m}_n(\lambda)$ is the optimal objective value of (5.11). We use $\bar{m}_{n,\text{opt}}$ as the optimal objective value in (5.11). For convenience of the proof and to illustrate the dependency of the activation cost λ, we express it as a function of λ in this proof. The numerator in (4.161) represents the expected penalty of source n strating from i-th delivery time to $(i+1)$-th delivery time and the denominator represents the expected time from i-th delivery time to the end of $(i+1)$-th delivery time. In order to prove Theorem 4.1, we need to introduce the following Lemma.

Lemma 4.12 $\bar{m}_n(\lambda)$ *is a continuous and strictly increasing function of* λ.

Proof 4.11 *The* $(i+1)$*-th delivery time from source* n *is given by*

$$D_{n,i+1}(\bar{m}_n(\lambda)) = S_{n,i+1}(\bar{m}_n(\lambda)) + Y_{n,i+1}, \tag{4.162}$$

and for (4.55)*, the* $(i+1)$*-th sampling time is*

$$S_{n,i+1}(\bar{m}_n(\lambda)) = \inf\{t \geq D_{n,i}(\bar{m}_n(\lambda)) : |\varepsilon_n(t)| \geq v_n(\bar{m}_n(\lambda))\}. \tag{4.163}$$

Let the waiting time after the delivery of the i*-th sample is*

$$Z_{n,i}(\bar{m}_n(\lambda)) = \inf\{z \geq 0 : |\varepsilon_n(D_{n,i}(\bar{m}_n(\lambda)) + z)| \geq v_n(\bar{m}_n(\lambda))\}, \tag{4.164}$$

which represents the minimum time z *upto which it needs to wait after the delivery of the* i*-th sample before generating the* $(i+1)$*-th sample. Hence, by using* (4.162), (4.163)*, and* (4.164) *the sampling time* $S_{n,i}(\bar{m}_n(\lambda))$ *and the delivery time* $D_{n,i}(\bar{m}_n(\lambda))$ *can also be expressed as*

$$S_{n,i}(\bar{m}_n(\lambda)) = \sum_{j=0}^{i-1} Y_{n,j} + Z_{n,j}(\bar{m}_n(\lambda)), \tag{4.165}$$

$$D_{n,i}(\bar{m}_n(\lambda)) = \sum_{j=0}^{i-1} Y_{n,j} + Z_{n,j}(\bar{m}_n(\lambda)) + Y_{n,i}. \tag{4.166}$$

By substituting (4.165) *and* (4.166) *in* (4.161)*, we get that*

$$\bar{m}_n(\lambda) = \frac{\mathbb{E}\left[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}(\bar{m}_n(\lambda))+Y_{n,i+1}} w_n \varepsilon_n^2(s) ds\right] + \lambda \mathbb{E}[Y_{n,i+1}]}{\mathbb{E}[Y_{n,i+1} + Z_{n,i}(\bar{m}_n(\lambda))]}. \tag{4.167}$$

The optimal objective value $\bar{m}_n(\lambda)$ *in* (4.167) *is exactly equal to the root of the following equation:*

$$f(\beta_n) + \lambda \mathbb{E}[Y_{n,i+1}] = 0, \tag{4.168}$$

where

$$f(\beta_n) = \mathbb{E}\bigg[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}(\beta_n)+Y_{n,i+1}} w_n \varepsilon_n^2(s)ds\bigg] - \beta_n \mathbb{E}[Z_{n,i}(\beta_n) + Y_{n,i+1}]. \tag{4.169}$$

Because $f(\beta_n)$ is a concave, continuous, and strictly decreasing function of β_n [34, Lemma 2], from (4.168), it is evident that the root of (4.168) is unique and continuous in λ. Hence, $\bar{m}_n(\lambda)$ is unique and continuous in λ. From (4.168), we get that

$$f(\beta_n) = -\lambda \mathbb{E}[Y_{n,i+1}]. \tag{4.170}$$

For any $0 \leq \lambda_1 \leq \lambda_2$ and $\beta_n = \bar{m}_n(\lambda)$, from (4.170), we have

$$f(\bar{m}_n(\lambda_1)) = -\lambda_1 \mathbb{E}[Y_{n,i+1}], \tag{4.171}$$

$$f(\bar{m}_n(\lambda_2)) = -\lambda_2 \mathbb{E}[Y_{n,i+1}]. \tag{4.172}$$

As $f(\beta_n)$ is a continuous and strictly decreasing function of β, for any non-negative $\lambda_2 > \lambda_1$ implies $\bar{m}_n(\lambda_1) < \bar{m}_n(\lambda_2)$. Therefore, $\bar{m}_n(\lambda)$ is continuous and strictly increasing function of λ.

The next task is to show the properties of the threshold $v_n(\bar{m}_n(\lambda))$ in (4.160) by using the (4.24) for the three cases of Gauss-Markov processes. In that sequel, we need to use the following lemma.

Lemma 4.13 *The threshold $v_n(\bar{m}_n(\lambda))$ is continuous and strictly increasing in λ irrespective of the signal structure.*

Proof 4.12 *For $\theta_n > 0$, $v_n(\bar{m}_n(\lambda))$ is as follows*

$$v_n(\bar{m}_n(\lambda)) = \frac{\sigma_n}{\sqrt{\theta_n}} G^{-1}\bigg(\frac{w_n \frac{\sigma_n^2}{2\theta_n} \mathbb{E}[e^{-2\theta_n Y_{n,i}}]}{w_n \frac{\sigma_n^2}{2\theta_n} - \bar{m}_n(\lambda)}\bigg). \tag{4.173}$$

The derivative of $v_n(\bar{m}_n(\lambda))$ is given by

$$v_n'(\bar{m}_n(\lambda)) = \frac{\sigma_n}{\sqrt{\theta_n}}\left\{G^{-1}\left(\frac{w_n\frac{\sigma_n^2}{2\theta_n}\mathbb{E}\big[e^{-2\theta_n Y_{n,i}}\big]}{w_n\frac{\sigma_n^2}{2\theta_n} - \bar{m}_n(\lambda)}\right)\right\}'. \quad (4.174)$$

Let

$$G^{-1}\left(\frac{w_n\frac{\sigma_n^2}{2\theta_n}\mathbb{E}\big[e^{-2\theta_n Y_{n,i}}\big]}{w_n\frac{\sigma_n^2}{2\theta_n} - \bar{m}_n(\lambda)}\right) = y. \quad (4.175)$$

By using the property of derivative of an inverse function [116], $v_n'(\bar{m}_n(\lambda))$ in (4.179) *can be expressed as*

$$v_n'(\bar{m}_n(\lambda)) = \frac{\sigma_n}{\sqrt{\theta_n}}\frac{1}{G'(y)}\frac{w_n\frac{\sigma_n^2}{2\theta_n}\mathbb{E}\big[e^{-2\theta_n Y_{n,i}}\big]}{(w_n\frac{\sigma_n^2}{2\theta_n} - \bar{m}_n(\lambda))^2}, \quad (4.176)$$

where $G'(x)$ is as follows

$$G'(x) = -\frac{\sqrt{\pi}}{2}\frac{e^{x^2}}{x^2}\mathit{erf}(x) + \sqrt{\pi}e^{x^2}\mathit{erf}(x) + \frac{1}{x} > 0, \quad (4.177)$$

for all $x > 0$. Hence, by using Lemma 4.12 and the fact that $v_n'(\bar{m}_n(\lambda)) > 0$, it is proved that $v_n(\bar{m}_n(\lambda))$ is a strictly increasing function of λ.

In addition, for $\theta_n < 0$, $v_n(\bar{m}_n(\lambda))$ can be expressed as

$$v_n(\bar{m}_n(\lambda)) = \frac{\sigma_n}{\sqrt{-\theta_n}}K^{-1}\left(\frac{w_n\frac{\sigma_n^2}{2\theta_n}\mathbb{E}\big[e^{-2\theta_n Y_{n,i}}\big]}{w_n\frac{\sigma_n^2}{2\theta_n} - \bar{m}_n(\lambda)}\right). \quad (4.178)$$

The derivative of $v_n(\bar{m}_n(\lambda))$ is then given by

$$v_n'(\bar{m}_n(\lambda)) = \frac{\sigma_n}{\sqrt{-\theta_n}}\left\{K^{-1}\left(\frac{w_n\frac{\sigma_n^2}{2\theta_n}\mathbb{E}\big[e^{-2\theta_n Y_{n,i}}\big]}{w_n\frac{\sigma_n^2}{2\theta_n} - \bar{m}_n(\lambda)}\right)\right\}'. \quad (4.179)$$

Let

$$K^{-1}\left(\frac{w_n\frac{\sigma_n^2}{2\theta_n}\mathbb{E}\big[e^{-2\theta_n Y_{n,i}}\big]}{w_n\frac{\sigma_n^2}{2\theta_n} - \bar{m}_n(\lambda)}\right) = p. \quad (4.180)$$

Utilizing the property of the derivative of an inverse function, $v_n'(\bar{m}_n(\lambda))$ in (4.179) can be expressed as

$$v_n'(\bar{m}_n(\lambda)) = \frac{\sigma_n}{\sqrt{-\theta_n}} \frac{1}{K'(p)} \frac{w_n \frac{\sigma_n^2}{2\theta_n} \mathbb{E}[e^{-2\theta_n Y_{n,i}}]}{(w_n \frac{\sigma_n^2}{2\theta_n} - \bar{m}_n(\lambda))^2}, \tag{4.181}$$

where $K'(x)$ is as follows

$$K'(x) = -\frac{\sqrt{\pi}}{2} \frac{e^{-x^2}}{x^2} \mathit{erfi}(x) - \sqrt{\pi} e^{-x^2} \mathit{erfi}(x) + \frac{1}{x} < 0, \tag{4.182}$$

for all $x > 0$. Hence, from Lemma 4.12 and the fact that $v_n'(\bar{m}_n(\lambda)) > 0$, it is proved that $v_n(\bar{m}_n(\lambda))$ is a strictly increasing function of λ. By using similar proof arguments, the result can be proven for $\theta_n = 0$. Hence, combining the results for all of the three cases of θ_n, Lemma 4.13 is proven.

From (4.183), the set $\Psi_n(\lambda)$ is

$$\Psi_n(\lambda) = \{(\varepsilon, \gamma) : \gamma > 0 \,\text{or}\, |\varepsilon| < v_n(\bar{m}_n(\lambda))\}. \tag{4.183}$$

Moreover, for a given ε, if $\gamma = 0$ and $\varepsilon \in \Psi_n(\lambda_1)$, then

$$|\varepsilon| < v_n(\bar{m}_n(\lambda_1)). \tag{4.184}$$

Because Lemma 4.13 implies that $v_n(\bar{m}_n(\lambda))$ is continuous and strictly increasing in λ, we get that for $\gamma = 0$, $(\varepsilon, \gamma) \in \Psi_n(\lambda_2)$ for any $\lambda_1 < \lambda_2$. Hence, $\Psi_n(\lambda_1) \subseteq \Psi_n(\lambda_2)$. Thus from the definition of indexability, the arm n is indexable for all n. This completes the proof.

4.E Proof of Theorem 4.2

Substituting (4.183) into the definition of Whittle index in (4.28), we obtain that

$$W_n(\varepsilon, \gamma) = \inf_{\lambda} \{\lambda \in \mathbb{R} : \gamma > 0 \,\text{or}\, |\varepsilon| < v_n(\lambda)\}. \tag{4.185}$$

First, we consider the case when $\gamma = 0$. By using Lemma 4.13, (4.185) implies that the Whittle index $W_n(\varepsilon, \gamma)$ is unique and it satisfies the following at $\lambda = W_n(\varepsilon, \gamma)$:

$$|\varepsilon| = v_n(\bar{m}_n(W_n(\varepsilon, \gamma))), \tag{4.186}$$

where $v_n(\cdot)$ is defined in (4.24). First, consider the case of stable OU process (i.e., $\theta_n > 0$). Substituting (4.24) for $\theta_n > 0$ into (4.186), we get that

$$|\varepsilon| = \frac{\sigma_n}{\sqrt{\theta_n}} G^{-1} \frac{w_n \frac{\sigma_n^2}{2\theta_n} \mathbb{E}[e^{-2\theta_n Y_{n,i}}]}{w_n \frac{\sigma_n^2}{2\theta_n} - \bar{m}_n(W_n(\varepsilon, \gamma))}, \tag{4.187}$$

which implies

$$G\left(\frac{\sqrt{\theta_n}}{\sigma_n}|\varepsilon|\right) = \frac{w_n \frac{\sigma_n^2}{2\theta_n} \mathbb{E}[e^{-2\theta_n Y_{n,i}}]}{w_n \frac{\sigma_n^2}{2\theta_n} - \bar{m}_n(W_n(\varepsilon, \gamma))}. \tag{4.188}$$

After some rearrangements, (4.188) becomes

$$\bar{m}_n(W_n(\varepsilon, \gamma)) = \frac{w_n \frac{\sigma_n^2}{2\theta_n} \left(G\left(\frac{\sqrt{\theta_n}}{\sigma_n}|\varepsilon|\right) - \mathbb{E}[e^{-2\theta_n Y_{n,i}}]\right)}{G\left(\frac{\sqrt{\theta_n}}{\sigma_n}|\varepsilon|\right)}. \tag{4.189}$$

The optimal objective value $\bar{m}_n(W_n(\varepsilon))$ to problem (5.11) is defined by

$$\begin{aligned} &\bar{m}_n(W_n(\varepsilon, \gamma)) = \\ &\frac{\mathbb{E}\left[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}(\varepsilon, \bar{m}_n(W_n(\varepsilon)))+Y_{n,i+1}} w_n O_{n,s}^2 ds\right] + W_n(\varepsilon, \gamma)\mathbb{E}[Y_{n,i+1}]}{\mathbb{E}[Z_{n,i}(\varepsilon, \bar{m}_n(W_n(\varepsilon, \gamma))) + Y_{n,i+1}]}. \end{aligned} \tag{4.190}$$

Substituting (4.190) into (4.189) implies

$$\begin{aligned} &\frac{\mathbb{E}\left[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}(\varepsilon, \bar{m}_n(W_n(\varepsilon)))+Y_{n,i+1}} w_n O_{n,s}^2 ds\right] + W_n(\varepsilon, \gamma)\mathbb{E}[Y_{n,i+1}]}{\mathbb{E}[Z_{n,i}(\varepsilon, \bar{m}_n(W_n(\varepsilon, \gamma))) + Y_{n,i+1}]} \\ &= \frac{w_n \frac{\sigma_n^2}{2\theta_n} \left(G\left(\frac{\sqrt{\theta_n}}{\sigma_n}|\varepsilon|\right) - \mathbb{E}[e^{-2\theta_n Y_{n,i}}]\right)}{G\left(\frac{\sqrt{\theta_n}}{\sigma_n}|\varepsilon|\right)}, \end{aligned} \tag{4.191}$$

which yields

$$\mathbb{E}\left[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}(\varepsilon,\bar{m}_n(W_n(\varepsilon,\gamma)))+Y_{n,i+1}} w_n O_{n,s}^2 ds\right] + W_n(\varepsilon,\gamma)\mathbb{E}[Y_{n,i+1}] =$$
$$\mathbb{E}[Z_{n,i}(\varepsilon,\bar{m}_n(W_n(\varepsilon,\gamma)))+Y_{n,i+1}] \frac{w_n \frac{\sigma_n^2}{2\theta_n}\Big(G\Big(\frac{\sqrt{\theta_n}}{\sigma_n}|\varepsilon|\Big) - \mathbb{E}[e^{-2\theta_n Y_{n,i}}]\Big)}{G\left(\frac{\sqrt{\theta_n}}{\sigma_n}|\varepsilon|\right)}. \tag{4.192}$$

After rearranging (4.192), we get that

$$\begin{aligned} W_n(\varepsilon,\gamma) = & \frac{1}{\mathbb{E}[Y_{n,i}]}\Bigg\{\mathbb{E}[Z_{n,i}(\varepsilon,\bar{m}_n(W_n(\varepsilon,\gamma))) + Y_{n,i+1}] \\ & \frac{w_n \frac{\sigma_n^2}{2\theta_n}\left(G\left(\frac{\sqrt{\theta_n}}{\sigma_n}|\varepsilon|\right) - \mathbb{E}[e^{-2\theta_n Y_{n,i}}]\right)}{G\left(\frac{\sqrt{\theta_n}}{\sigma_n}|\varepsilon|\right)} \\ & - \mathbb{E}\left[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}(\varepsilon,\bar{m}_n(W_n(\varepsilon,\gamma)))+Y_{n,i+1}} w_n O_{n,s}^2 ds\right]\Bigg\}, \end{aligned} \tag{4.193}$$

where (4.193) holds because $Y_{n,i}$'s are *i.i.d.*. In Theorem 4.2, we have used $Z_{n,i}(\varepsilon)$ to represent $Z_{n,i}(\varepsilon,\bar{m}_n(W_n(\varepsilon,\gamma)))$ because $Z_{n,i}$ is dependent on $m_n(W_n(\varepsilon,\gamma))$ through the state ε.

In addition, for unstable OU process, when $\theta_n < 0$, by using the similar proof arguments, we can prove (4.31) for $\theta_n < 0$.

Subsequently, when $\theta_n = 0$, by using (4.24), (4.28), and Lemma 4.13, at $\lambda = W_n(\varepsilon,\gamma)$, we get that

$$|\varepsilon| = \frac{1}{\sqrt{w_n}}\sqrt{3(\bar{m}_n(W_n(\varepsilon,\gamma)) - w_n\sigma_n^2\mathbb{E}[Y_{n,i}])}, \tag{4.194}$$

which implies

$$w_n\varepsilon^2 = 3(\bar{m}_n(W_n(\varepsilon,\gamma)) - w_n\sigma_n^2\mathbb{E}[Y_{n,i}]). \tag{4.195}$$

After some rearrangements, (4.195) becomes

$$w_n\left(\frac{\varepsilon^2}{3} + \sigma_n^2\mathbb{E}[Y_{n,i}]\right) = \bar{m}_n(W_n(\varepsilon,\gamma)). \tag{4.196}$$

Substituting (4.190) into (4.196), we obtain that

$$w_n\left(\frac{\varepsilon^2}{3}+\sigma_n^2\mathbb{E}[Y_{n,i}]\right)=\frac{\mathbb{E}\left[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}(\varepsilon,\bar{m}_n(W_n(\varepsilon,\gamma)))+Y_{n,i+1}} w_n O_{n,s}^2 ds\right]+W_n(\varepsilon,\gamma)\mathbb{E}[Y_{n,i}]}{\mathbb{E}[Z_{n,i}(\varepsilon,\bar{m}_n(W_n(\varepsilon,\gamma)))+Y_{n,i+1}]}, \tag{4.197}$$

which yields

$$\begin{aligned} &w_n\mathbb{E}\left[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}(\varepsilon,m_n(W_n(\varepsilon,\gamma)))+Y_{n,i+1}} O_{n,s}^2 ds\right]+W_n(\varepsilon,\gamma)\mathbb{E}[Y_{n,i}] \\ &= w_n\mathbb{E}[Y_{n,i}+Z_{n,i}(\varepsilon,m_n(W_n(\varepsilon,\gamma)))]\left(\frac{\varepsilon^2}{3}+\sigma_n^2\mathbb{E}[Y_{n,i}]\right), \end{aligned} \tag{4.198}$$

from which (4.30) follows for $\theta_n = 0$.

Next, we consider $\gamma > 0$. From Definition 4.1 and Definition 4.3, the possible infimum cost λ for which to activate and not to activate are equally desirable is $-\infty$, from which (4.32) follows. This concludes the proof.

4.F Proof of Lemma 5.2

In order to prove Lemma 5.2, we need to consider the following two cases:

Case 1: If $|\varepsilon_n(D_{n,i})| = |O_{n,D_{n,i}-S_{n,i}}| = |O_{n,Y_{n,i}}| \geq |\varepsilon|$, then $S_{n,i+1} = D_{n,i}$. Hence,

$$\begin{aligned} \mathbb{E}[D_{n,i+1}(\varepsilon)-D_{n,i}(\varepsilon)] &= \mathbb{E}[S_{n,i+1}(\varepsilon)+Y_{n,i+1}-S_{n,i+1}(\varepsilon)], \\ &= \mathbb{E}[Y_{n,i+1}]. \end{aligned} \tag{4.199}$$

Using the fact that the $Y_{n,i}$'s are independent of the OU process, we can obtain

$$\begin{aligned}
&\mathbb{E}\bigg[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds\bigg|O_{n,Y_{n,i}}, |O_{n,Y_{n,i}}| \geq |\varepsilon|\bigg] \\
=&\mathbb{E}\bigg[\int_{Y_{n,i}}^{Y_{n,i}+Y_{n,i+1}} O_{n,s}^2 ds\bigg|O_{n,Y_{n,i}}, |O_{n,Y_{n,i}}| \geq |\varepsilon|\bigg] \\
=&\mathbb{E}\bigg[\int_{0}^{Y_{n,i}+Y_{n,i+1}} O_{n,s}^2 ds\bigg|O_{n,Y_{n,i}}, |O_{n,Y_{n,i}}| \geq |\varepsilon|\bigg] \\
&-\mathbb{E}\bigg[\int_{0}^{Y_{n,i}} O_{n,s}^2 ds\bigg|O_{n,Y_{n,i}}, |O_{n,Y_{n,i}}| \geq |\varepsilon|\bigg]. \qquad (4.200)
\end{aligned}$$

By invoking Lemma 4.7, we get that

$$\begin{aligned}
&\mathbb{E}\bigg[\int_{0}^{Y_{n,i}+Y_{n,i+1}} O_{n,s}^2 ds\bigg|O_{n,Y_{n,i}}, |O_{n,Y_{n,i}}| \geq |\varepsilon|\bigg] \\
=&R_{n,2}(O_{Y_{n,i}+Y_{n,i+1}}), \qquad (4.201)
\end{aligned}$$

$$\mathbb{E}\bigg[\int_{0}^{Y_{n,i}} O_{n,s}^2 ds\bigg|O_{n,Y_{n,i}}, |O_{n,Y_{n,i}}| \geq |\varepsilon|\bigg] = R_{n,2}(O_{Y_{n,i}}). \qquad (4.202)$$

Substituting (4.201) and (4.202) into (4.200), it becomes

$$\begin{aligned}
&\mathbb{E}\bigg[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds\bigg|O_{n,Y_{n,i}}, |O_{n,Y_{n,i}}| \geq |\varepsilon|\bigg] \\
=&R_{n,2}(O_{Y_{n,i}+Y_{n,i+1}}) - R_{n,2}(O_{Y_{n,i}}), \\
=&R_{n,2}(|\varepsilon| + O_{n,Y_{n,i+1}}) - R_{n,2}(O_{Y_{n,i}}), \qquad (4.203)
\end{aligned}$$

where (4.203) holds because at $t = D_{n,i}(\varepsilon)$, the estimation error $O_{n,Y_{n,i}}$ reaches the threshold $|\varepsilon|$.

Case 2: If $|\varepsilon_n(D_{n,i})| = |O_{n,Y_{n,i}}| < |\varepsilon|$, then, almost surely,

$$|\varepsilon_n(S_{n,i+1})| = |\varepsilon|. \qquad (4.204)$$

Then,

$$\begin{aligned}&\mathbb{E}[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)]\\=&\mathbb{E}[D_{n,i+1}(\varepsilon) - S_{n,i+1}(\varepsilon) + S_{n,i+1}(\varepsilon) - S_{n,i}(\varepsilon) + S_{n,i}(\varepsilon) - D_{n,i}(\varepsilon)].\end{aligned} \tag{4.205}$$

Because $D_{n,i+1}(\varepsilon) = S_{n,i+1}(\varepsilon) + Y_{n,i+1}$, by invoking Lemma 4.7, we can obtain the remaining expectations in (4.205) which are given by

$$\mathbb{E}\left[S_{n,i+1}(\varepsilon) - S_{n,i}(\varepsilon)\middle|O_{n,Y_{n,i}}, |O_{n,Y_{n,i}}| < |\varepsilon|\right] = R_{n,1}(\varepsilon), \tag{4.206}$$

$$\mathbb{E}\left[D_{n,i}(\varepsilon) - S_{n,i}(\varepsilon)\middle|O_{n,Y_{n,i}}, |O_{n,Y_{n,i}}| < |\varepsilon|\right] = R_{n,1}(O_{n,Y_{n,i}}). \tag{4.207}$$

Using (4.206) and (4.207), we get that

$$\begin{aligned}&\mathbb{E}\left[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)\middle|O_{n,Y_{n,i}}, |O_{n,Y_{n,i}}| < |\varepsilon|\right]\\=&\mathbb{E}[Y_{n,i+1}] + R_{n,1}(\varepsilon) - R_{n,1}(O_{n,Y_{n,i}}).\end{aligned} \tag{4.208}$$

In addition,

$$\begin{aligned}&\mathbb{E}\left[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds\middle|O_{n,Y_{n,i}}, |O_{n,Y_{n,i}}| < |\varepsilon|\right]\\=&\mathbb{E}\left[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}(\varepsilon)+Y_{n,i+1}} O_{n,s}^2 ds\right],\\=&\mathbb{E}\left[\int_0^{Y_{n,i}+Z_{n,i}(\varepsilon)+Y_{n,i+1}} O_{n,s}^2 ds\right] - \mathbb{E}\left[\int_0^{Y_{n,i}} O_{n,s}^2 ds\right].\end{aligned} \tag{4.209}$$

By invoking Lemma 4.7 again, we can obtain

$$\begin{aligned}&\mathbb{E}\left[\int_0^{Y_{n,i}+Z_{n,i}(\varepsilon)+Y_{n,i+1}} O_{n,s}^2 ds\middle|O_{n,Y_{n,i}}, |O_{n,Y_{n,i}}| < |\varepsilon|\right]\\=&R_{n,2}(O_{n,Y_{n,i}+Z_{n,i}(\varepsilon)+Y_{n,i+1}}),\end{aligned} \tag{4.210}$$

$$\mathbb{E}\left[\int_0^{Y_{n,i}} O_{n,s}^2 ds\middle|O_{n,Y_{n,i}}, |O_{n,Y_{n,i}}| < |\varepsilon|\right] = R_{n,2}(O_{n,Y_{n,i}}). \tag{4.211}$$

By using (4.210) and (4.211) in (4.209), we have

$$\begin{aligned}
&\mathbb{E}\bigg[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds\bigg|O_{n,Y_{n,i}}, |O_{n,Y_{n,i}}| < |\varepsilon|\bigg] \\
=&R_{n,2}(O_{n,Y_{n,i}+Z_{n,i}(\varepsilon)+Y_{n,i+1}}) - R_{n,2}(O_{n,Y_{n,i}}), && (4.212)\\
=&R_{n,2}(|O_{n,Y_{n,i}}| + O_{n,Y_{n,i+1}}) - R_{n,2}(O_{n,Y_{n,i}}), && (4.213)
\end{aligned}$$

where (4.213) holds because at $t = D_{n,i}(\varepsilon)$, the estimation error $O_{n,Y_{n,i}}$ is below the threshold $|\varepsilon|$.

By combining (4.199) and (4.208) of the two cases, yields

$$\begin{aligned}
\mathbb{E}\left[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)\middle|O_{n,Y_{n,i}}\right] =&\max\{R_{n,1}(|\varepsilon|) - R_{n,1}(O_{n,Y_{n,i}}), 0\} \\
&+ \mathbb{E}[Y_{n,i+1}]. && (4.214)
\end{aligned}$$

By taking the expectation over $O_{n,Y_{n,i}}$ in (4.214) gives

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)\middle|O_{n,Y_{n,i}}\right]\right] \\
=&\mathbb{E}[\max\{R_{n,1}(|\varepsilon|) - R_{n,1}(O_{n,Y_{n,i}}), 0\} + Y_{n,i+1}] \\
=&\mathbb{E}[\max\{R_{n,1}(|\varepsilon|) - R_{n,1}(O_{n,Y_{n,i}}), 0\} + R_{n,1}(O_{n,Y_{n,i}})], && (4.215)
\end{aligned}$$

where (4.215) follows from the fact that $Y_{n,i}$'s are *i.i.d.* and (4.101) in Lemma 4.7. Because $R_{n,1}(\cdot)$ is an even function, form (4.215) we get that

$$\mathbb{E}\left[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)\middle|O_{n,Y_{n,i}}\right] = \max\{R_{n,1}(\max\{|\varepsilon|, |O_{n,Y_{n,i}}|\}). \qquad (4.216)$$

Similarly, by combining (4.203) and (4.213) of the two cases, yields

$$\begin{aligned}
&\mathbb{E}\bigg[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds\bigg|O_{n,Y_{n,i}}\bigg] \\
=&R_{n,2}(\max\{|\varepsilon|, |O_{n,Y_{n,i}}|\} + O_{n,Y_{n,i+1}}) - R_{n,2}(O_{n,Y_{n,i}}). && (4.217)
\end{aligned}$$

Finally, by taking the expectation over $O_{n,Y_{n,i}}$ in (4.214) and (4.217) and using the fact that $R_{n,1}(\cdot)$ and $R_{n,2}(\cdot)$ are even functions, Lemma 5.2 is proven.

4.G Proof of Lemma 4.2

Because λ represents the cost to activate an arm, it is optimal in (5.5) to activate a *dummy bandit* only when $\lambda < 0$. Conversely, when $\lambda \geq 0$, it is optimal not to activate the dummy bandit. Hence, from Definition 4.1, the *dummy bandits* are always indexable. In addition, from Definition 4.3 and the fact that the *dummy bandits* are activated only when $\lambda < 0$, we get $W_0(\varepsilon, \gamma) = 0$.

4.H Proof of Theorem 4.5

In order to prove Theorem 4.5, we first show that (4.66) and (4.70) are equivalent to each other. For single source, the source weight $w_1 = 1$ and the transmission cost $\lambda = 0$.

We first show the proof for stable OU process, i.e., for $\theta_1 > 0$. When $\varepsilon = v_1(\beta_1)$ in (4.67), we have

$$\begin{aligned} G\left(\frac{\sqrt{\theta_1}}{\sigma_1} v_1(\beta_1)\right) =& G\left(\frac{\sqrt{\theta_1}}{\sigma_1}\frac{\sigma_1}{\sqrt{\theta_1}} G^{-1}\left(\frac{\frac{\sigma_1^2}{2\theta_1}\mathbb{E}[e^{-2\theta_1 Y_{1,i}}]}{\frac{\sigma_1^2}{2\theta_1} - \beta_1}\right)\right), \\ =& \frac{\sigma_1^2 \mathbb{E}[e^{-2\theta_1 Y_{1,i}}]}{\sigma_1^2 - 2\theta_1\beta_1}. \end{aligned} \tag{4.218}$$

Substituting (4.218) into (4.29) for single source results

$$\begin{aligned} W_1(v_1(\beta_1), \gamma) =& \frac{1}{\mathbb{E}[Y_{1,i}]}\Bigg\{\mathbb{E}[D_{1,i+1}(\varepsilon) - D_{1,i}(\varepsilon)]\frac{\sigma_1^2}{2\theta_1}\left(1 - \frac{\mathbb{E}[e^{-2\theta_1 Y_{1,i}}]}{\frac{\sigma_1^2 \mathbb{E}[e^{-2\theta_1 Y_{1,i}}]}{\sigma_1^2 - 2\theta_1\beta_1}}\right) \\ &- \mathbb{E}\left[\int_{D_{1,i}(\varepsilon)}^{D_{1,i+1}(\varepsilon)} \varepsilon_1^2(s) ds\right]\Bigg\}, \end{aligned} \tag{4.219}$$

which becomes

$$W_1(v_1(\beta_1),\gamma) = \frac{1}{\mathbb{E}[Y_{1,i}]}\left\{\mathbb{E}[D_{1,i+1}(\varepsilon)-D_{1,i}(\varepsilon)]\beta_1-\mathbb{E}\left[\int_{D_{1,i}(\varepsilon)}^{D_{1,i+1}(\varepsilon)}\varepsilon_1^2(s)ds\right]\right\}, \tag{4.220}$$

The parameter β_1 in (4.220) can be found from (4.68) and (4.69), which is exactly equal to the optimal objective value $\bar{m}_{1,\text{opt}}$. Hence, substituting β_1 in (4.220) yields

$$W_1(v_1(\beta_1),\gamma) = 0. \tag{4.221}$$

If $\varepsilon > v_1(\beta_1)$, as $G(x)$ is a strictly increasing function in $[x,\infty)$, we have

$$G\left(\frac{\sqrt{\theta_1}}{\sigma_1}\varepsilon\right) > G\left(\frac{\sqrt{\theta_1}}{\sigma_1}v_1(\beta_1)\right), \tag{4.222}$$

which yields

$$\left(1-\frac{\mathbb{E}[e^{-2\theta_1 Y_{1,i}}]}{G\left(\frac{\sqrt{\theta_1}}{\sigma_1}\varepsilon\right)}\right) > \left(1-\frac{\mathbb{E}[e^{-2\theta_1 Y_{1,i}}]}{G\left(\frac{\sqrt{\theta_1}}{\sigma_1}v_1(\beta_1)\right)}\right). \tag{4.223}$$

From the above arguments, it is proved that $W_1(\varepsilon,\gamma) > W_1(v_1(\beta_1),\gamma)$ $= 0$ for $\varepsilon > v_1(\beta_1)$. Similarly, as $G(x)$ is an even function, for $\varepsilon < v_1(\beta_1)$, it holds that $W_1(\varepsilon,\gamma) < W_1(v_1(\beta_1),\gamma) = 0$.

By using the similar proof arguments we can show that for all θ_1, $W_1(v_1(\beta_1),\gamma) = 0$, $W_1(\varepsilon,\gamma) > 0$ for $\varepsilon > v_1(\beta_1)$, and $W_1(\varepsilon,\gamma) < 0$ for $\varepsilon < v_1(\beta_1)$. Hence, the two statements in (4.66) and (4.70) are equivalent to each other. This result also illustrate in Fig. 4.2 from which it is evident that $W_1(\varepsilon,\gamma)$ is an even function. We prove the optimality of Proposition 4.1 for any number of sources in Appendix 4.C. Hence, Theorem 4.5 is also optimal. This completes the proof.

4.I Proof of Theorem 4.3

If $\gamma = 0$, i.e., no sample from source is currently in service, from Proposition 4.2, we get that for an AoI $\Delta_n(t) = \delta$, it is optimal not to schedule source n if

$$\mathbb{E}[p(\delta + Y_{n,i+1})] < \bar{m}_{n,\text{age}}(\lambda), \tag{4.224}$$

where

$$\bar{m}_{n,\text{age}}(\lambda) = \frac{\mathbb{E}\left[\int_{D_{n,i}(\bar{m}_{n,\text{age}}(\lambda))}^{D_{n,i+1}(\bar{m}_{n,\text{age}}(\lambda))} w_n p_n(\Delta_n(s))ds\right] + \lambda \mathbb{E}[Y_{n,i+1}]}{\mathbb{E}[D_{n,i+1}(\bar{m}_{n,\text{age}}(\lambda)) - D_{n,i}(\bar{m}_{n,\text{age}}(\lambda))]}, \tag{4.225}$$

and $\bar{m}_{n,\text{age}}(\lambda)$ is the optimal objective value to (4.54). We use $\bar{m}_{n,\text{age-opt}}$ as the optimal objective value in (4.54). For convenience of the proof and to illustrate the dependency of the activation cost λ, we express it as a function of λ in the rest of the proofs.

According to (4.45), (4.46), and Lemma 4.12, $\bar{m}_{n,\text{age}}(\lambda)$ is a continuous and strictly increasing function of λ.

By utilizing (4.58), if $\gamma = 0$, for a given δ, if $(\delta, \gamma) \in \Psi_{n,\text{age}}(\lambda_1)$, then

$$\mathbb{E}[p_n(\delta + Y_{n,i+1})] < \bar{m}_{n,\text{age}}(\lambda_1). \tag{4.226}$$

By using the fact that $\bar{m}_{n,\text{age}}(\lambda)$ is continuous and strictly increasing in λ, we get that $(\delta, \gamma) \in \Psi_{n,\text{age}}(\lambda_2)$ for any $\lambda_1 < \lambda_2$. Hence, $\Psi_{n,\text{age}}(\lambda_1) \subseteq \Psi_{n,\text{age}}(\lambda_2)$. Thus from the definition of indexability, the arm n is indexable for all n. This concludes the proof.

4.J Proof of Theorem 4.4

When $\gamma = 0$, from Definition 4.3, we get that

$$W_n(\delta, \gamma) = \inf_{\lambda}\{\lambda \in \mathbb{R} : (\delta, \gamma) \in \Psi_{n,\text{age}}(\lambda)\}. \tag{4.227}$$

By utilizing (4.58) into (4.227), we obtain that

$$W_n(\delta,\gamma) = \inf_{\lambda}\{\lambda \in \mathbb{R} : \gamma > 0 \,\text{or}\, \mathbb{E}[p_n(\delta + Y_{n,i+1})] < \bar{m}_{n,\text{age}}(\lambda)\}. \tag{4.228}$$

At $\lambda = W_n(\delta,\gamma)$, we have

$$w_n\mathbb{E}[p_n(\delta + Y_{n,i+1})] = \bar{m}_{n,\text{age}}(W_n(\delta,\gamma)). \tag{4.229}$$

Substituting (4.225) into (4.229), we get that

$$\begin{aligned} & w_n\mathbb{E}[p_n(\delta + Y_{n,i+1})] = \\ & \frac{\mathbb{E}\left[\int_{D_{n,i}(\bar{m}_{n,\text{age}}(\lambda))}^{D_{n,i+1}(\bar{m}_{n,\text{age}}(\lambda))} w_n O_{n,s}^2 ds\right] + W_n(\delta,\gamma)\mathbb{E}[Y_{n,i+1}]}{\mathbb{E}[D_{n,i+1}(\bar{m}_{n,\text{age}}(\lambda)) - D_{n,i}(\bar{m}_{n,\text{age}}(\lambda))]}, \end{aligned} \tag{4.230}$$

which yields

$$\begin{aligned} & w_n\mathbb{E}[p_n(\delta + Y_{n,i+1})]\mathbb{E}[D_{n,i+1}(\bar{m}_{n,\text{age}}(\lambda)) - D_{n,i}(\bar{m}_{n,\text{age}})] \\ =& \mathbb{E}\left[\int_{D_{n,i}(\bar{m}_{n,\text{age}}(\lambda))}^{D_{n,i+1}(\bar{m}_{n,\text{age}}(\lambda))} w_n p_n(\Delta_n(s))ds\right] + W_n(\delta,\gamma)\mathbb{E}[Y_{n,i+1}], \end{aligned} \tag{4.231}$$

from which (4.59) follows because the $Y_{n,i}$'s are *i.i.d.*.

When $\gamma > 0$, from (4.58) and (4.228), (4.61) yields. This completes the proof.

4.K Proof of Lemma 4.3

In order to prove Lemma 4.3, we need to consider the following two cases:

Case 1: If $\mathbb{E}\left[p_n(\Delta_n(D_{n,i} + Y_{n,i+1})\right] = \mathbb{E}\left[p_n(\Delta_n(D_{n,i}) + Y_{n,i+1})\right] = \mathbb{E}\left[p_n(Y_{n,i} + Y_{n,i+1})\right] \geq \mathbb{E}\left[p_n(\delta + Y_{n,i+1})\right]$, then, $S_{n,i+1} = D_{n,i}$. Hence,

$$\begin{aligned} & \mathbb{E}[D_{n,i+1}(\delta) - D_{n,i}(\delta)\big|\delta, Y_{n,i}] \\ =& \mathbb{E}[S_{n,i+1}(\delta) + Y_{n,i+1} - S_{n,i+1}(\delta)\big|\delta, Y_{n,i}], && (4.232) \\ =& \mathbb{E}[Y_{n,i+1}] = \mathbb{E}[Y_{n,i}], && (4.233) \end{aligned}$$

where (4.233) holds because the $Y_{n,i}$'s are *i.i.d.*. In addition,

$$\begin{aligned}
&\mathbb{E}\left[\int_{D_{n,i}(\delta)}^{D_{n,i+1}(\delta)} p_n(s)ds\middle|\delta, Y_{n,i}\right] \\
=&\mathbb{E}\left[\int_{Y_{n,i}}^{Y_{n,i}+Y_{n,i+1}} p_n(s)ds\middle|\delta, Y_{n,i}\right] \\
=&\mathbb{E}\left[\int_{0}^{Y_{n,i}+Y_{n,i+1}} p_n(s)ds\middle|\delta, Y_{n,i}\right] - \mathbb{E}\left[\int_{0}^{Y_{n,i}} p_n(s)ds\middle|\delta, Y_{n,i}\right], \\
=&R_{n,3}(Y_{n,i}+Y_{n,i+1}) - R_{n,3}(Y_{n,i}). \qquad (4.234)
\end{aligned}$$

Case 2: If $\mathbb{E}\left[p_n(\Delta_n(D_{n,i}+Y_{n,i+1})\right] = \mathbb{E}\left[p_n(Y_{n,i}+Y_{n,i+1})\right] < \mathbb{E}\left[p_n(\delta+Y_{n,i+1})\right]$, then, almost surely,

$$\mathbb{E}\left[p_n(\Delta_n(S_{n,i+1}+Y_{n,i+1})\right] = \mathbb{E}\left[p_n(\delta+Y_{n,i+1})\right]. \qquad (4.235)$$

Then,

$$\mathbb{E}[D_{n,i+1}(\delta) - D_{n,i}(\delta)\big|\delta, Y_{n,i}] = \mathbb{E}[Z_{n,i}(\delta) + Y_{n,i+1}\big|\delta, Y_{n,i}] = \delta. \qquad (4.236)$$

In addition,

$$\begin{aligned}
&\mathbb{E}\left[\int_{D_{n,i}(\delta)}^{D_{n,i+1}(\delta)} p_n(s)ds\middle|\delta, Y_{n,i}\right] \\
=&\mathbb{E}\left[\int_{Y_{n,i}}^{Y_{n,i}+Z_{n,i}(\delta)+Y_{n,i+1}} p_n(s)ds\right], \\
=&\mathbb{E}\left[\int_{0}^{Y_{n,i}+Z_{n,i}(\delta)+Y_{n,i+1}} p_n(s)ds\right] - \mathbb{E}\left[\int_{0}^{Y_{n,i}} p_n(s)ds\right] \\
=&R_{n,3}(\delta+Y_{n,i+1}) - R_{n,3}(Y_{n,i}). \qquad (4.237)
\end{aligned}$$

By combining (4.233) and (4.236) of the two cases, yields

$$\mathbb{E}\left[D_{n,i+1}(\delta) - D_{n,i}(\delta)\middle|\delta, Y_{n,i}\right] = \max\{\delta, Y_{n,i}\}. \qquad (4.238)$$

Similarly, by combining (4.234) and (4.237) of the two cases, yields

$$\begin{aligned}&\mathbb{E}\left[\int_{D_{n,i}(\delta)}^{D_{n,i+1}(\delta)} p_n(s)ds \middle| Y_{n,i}\right]\\ =&R_{n,3}(\max\{\delta, Y_{n,i}\} + Y_{n,i+1}) - R_{n,3}(Y_{n,i}) \end{aligned} \tag{4.239}$$

Finally, by taking the expectation over $Y_{n,i}$ in (4.238) and (4.239), Lemma 4.3 is proven.

Chapter 5

Remote Estimation for Situational Awareness Maximization on Safety-critical Systems

5.1 Introduction

A broad range of safety-critical systems is ubiquitous across the world. For instance, in industrial automation, it is essential to continuously monitor the safety of various machines [117]. In patient health monitoring, precise tracking of the glucose level or heart rate is imperative to swiftly implement precautionary measures when they are required [118]. In disaster monitoring, it is important to promptly track any consistent changes in temperature or humidity, as they could indicate a possible disaster [119]. In these safety-critical situations, the monitoring system needs timely access and accurately interpret the states of remote systems. Any misunderstanding of the system state can lead to severe consequences.

In practice, multiple sensors are required to track various safety-critical situations. One challenge to efficiently utilize these sensor information in real-time is the limited capacity of the communication medium. Moreover, continuous monitoring of all sensors at all times is unnecessary if sensor measurements change slowly. Conversely, some sensor information may have more crucial content than others and hence need more attention. Therefore, it is efficient to adopt an on-demand approach. In this context, we consider a pull-based system where a receiver selects sensors and requests information when required. This selective retrieval of information ensures that the system receives essential information in a timely manner while minimizing unnecessary resource consumption.

In this chapter, we consider a pull-based status updating system consisting of multiple agents monitoring the status of different safety-critical situations. A receiver selects agents to transmit their updates through multiple unreliable channels. Due to transmission errors, the received packet may not be fresh. One performance metric that characterizes data freshness is the age of information (AoI) [1]. Let $U(t)$ be the generation time of the freshest observation that has been delivered to the receiver by time t. The AoI, as a function of t, is defined as $\Delta(t) = t - U(t)$ which exhibits a linear growth with time t and drops down to a smaller value whenever a fresher observation is delivered. However, the time difference represented by AoI can only capture the timeliness of the information but it cannot capture its significance. Hence, relying solely on AoI-based decision-making is not perfect. This is particularly relevant in safety-critical situations where misunderstanding about the situation can lead to significant performance loss. By exploiting the knowledge of the signal observation along with AoI in decision making, the incurred performance loss can be significantly improved. A key observation in this study is that any misinterpretation of a dangerous situation yields higher loss compared to the misinterpretation of a safe situation. Based on the above mentioned insights, we introduce a framework for quantifying the *cost of a dangerous situation* that characterizes the performance loss caused by situational unawareness.

In this chapter, we utilize both AoI and the most recently received observation to estimate the current situation in the environment. For any given signal observation, the system performance degradation due to situational unawareness can be expressed as a function of the AoI which can be non-monotonic [67, 120, 121]. The goal of this chapter is to find the optimal scheduling policy to select agents and to request observations while improving the system performance. The contributions of this work are as follows:

- We introduce a novel framework where the current status of a safety-critical system is estimated based on the AoI and the latest received signal observation. The loss for wrong estimation is modeled by general loss functions L. Specifically, the loss of wrongly estimating a dangerous situation is significantly high. By considering general loss functions, we mitigate the limitations of 0-1 loss, quadratic loss, and logarithmic loss as L can quantify the costs associated with the unawareness of a potentially dangerous situation. By adopting appropriate loss functions L, our results can be applied to health, safety, and security monitoring, which are paramount.

- We formulate a penalty function that represents the expected loss L given the AoI and the received observation. This formulation works for both optimal and sub-optimal estimators. We obtain an information-theoretic bound of the penalty function by utilizing the optimal estimator and the L-conditional entropy (or generalized conditional entropy) given the AoI and the received observation where L represents general loss functions [120, 122, 123]. In literature, there exists numerous metrics that address the state of the information content along with AoI such as Age of Incorrect Information (AoII) [58], Age of Synchronization (AoS) [57], Version AoI [26], AoI at Query (QAoI) [60]. These studies could not explain this information-theoretic bound. Moreover, the penalty considered in prior studies was monotonically growing functions of age whereas our study allows non-monotonic age penalty functions.

- In literature, numerous freshness metrics based on entropy and mutual information were considered [4,53,66]. In this chapter, we demonstrate that for optimal estimators the penalty function can be expressed as L-conditional entropy. In contrast to earlier studies, the entropy-based freshness metric in our study holds a significant contribution to remote estimation and prediction.

- We consider a multi-agent, multi-channel pull-based status updating problem. Our findings demonstrate that when utilizing one-time slot transmission time and optimal estimators, it is beneficial to always keep the servers busy (see Theorem 5.2). However, channel resource limitations prevent all agents from transmitting information continuously. To address this issue, we formulate the multi-agent, multi-channel transmission scheduling problem as a Restless Multi-armed Bandit (RMAB) with expanding state space. Because the state space contains the information on all historically received observations and their AoI values. We simplify the state space by showing that the latest received observation and its age is a sufficient statistic of all historically received observations and their AoI values.

- We utilize relaxation and the Lagrangian method to decompose the original problem into multiple separated Markov Decision Processes (MDPs). our analysis is different from existing literature [48,53,121,124,125] that utilizes belief MDP and represents the state space using belief states, i.e., the probability distribution over all possible states. In our analysis, instead of formulating any belief MDP, we formulate an MDP where the state of each MDP is the latest received observation and its age. Our state space is linearly growing with age whereas in earlier studies the state space is geometrically increasing with the belief states.

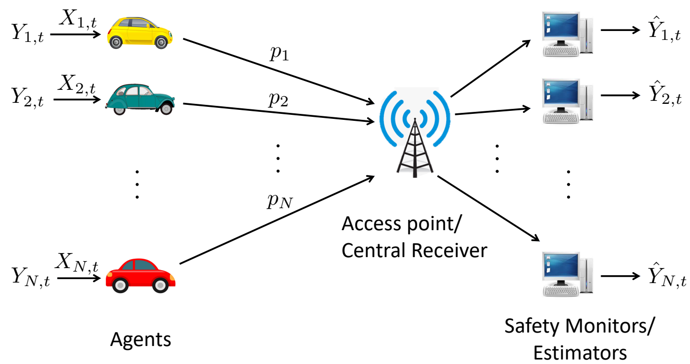


Figure 5.1: A multi-agent, multi-channel safety monitoring system.

- We solve each MDP by dynamic programming [126]. By utilizing the solution to the MDPs, we provide a low-complexity "Maximum Gain First" scheduling policy which is asymptotically optimal. Our results work for both reliable and unreliable channels.

- Numerical results illustrate that our multi-agent, multi-channel scheduling policy achieves up to 100 times performance gain over periodic updating policy and up to 10 times over randomized policy which randomly selects agents depending on the number of available channels.

5.2 System Model

This section describes the multi-agent, multi-channel system depicted in Figure 5.1.

5.2.1 Communication Model

Let us consider the pull-based status updating system depicted in Figure 5.1, where a central receiver pulls the status updates (e.g., car location) of N agents to monitor their safety (e.g., safe, cautious, dangerous). Let $X_{n,t}$ be the status of agent n at time t. The signal $X_{n,t}$ is a finite-state Markov chain with B states, where $2 \leq B < \infty$. We assume that the status updates of N agents are independent of each other. Let $Y_{n,t}$ quantify the safety level associated with agent n, where $X_{n,t}$ and $Y_{n,t}$ evolve according to the Markovian relationship illustrated in Figure 5.2, which can be described by the following three Markov chains for all t:, (i) $X_{n,t-1} \leftrightarrow X_{n,t} \leftrightarrow X_{n,t+1}$, (ii) $Y_{n,t-1} \leftrightarrow X_{n,t} \leftrightarrow X_{n,t-1}$, and (iii) $Y_{n,t} \leftrightarrow X_{n,t} \leftrightarrow X_{n,t+1}$.

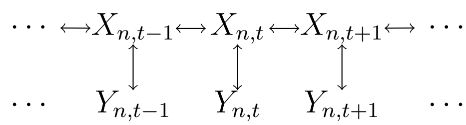


Figure 5.2: Relationship between $X_{n,t}$ and $Y_{n,t}$.

The status updates of N agents are transmitted through M unreliable wireless channels. In response to the pull request from the central receiver, each agent n generates and submits a time-stamped updating message $(X_{n,t}, t)$ to one wireless channel. We assume that it takes one-time slot to transmit a message update to the receiver. Due to wireless channel fading, the transmission of the status updates becomes unreliable. Let p_n be the probability of a successful transmission from agent n, irrespective of the selected wireless channel.

Our systems consist of N estimators. The goal of each estimator n is to estimate the safety level $Y_{n,t}$ by utilizing causally received updates from agent n available at the receiver. Due to transmission errors, the delivered information at the receiver may not be fresh and is represented by $X_{n,t-\Delta_n(t)}$ that is generated $\Delta_n(t)$ times ago. The time difference $\Delta_n(t)$ is usually called *age of information (AoI)* [1], which represents the staleness of the status of the n-th agent.

Based on the latest available information, the n-th estimator estimates the safety level $Y_{n,t}$ and outputs $\hat{y} = \phi_n(\mathbf{\Xi}_{n,t}, \mathbf{H}_{n,t}) \in \mathcal{Y}$, where $\phi_n(\cdot,\cdot)$ is a function of all historically received packets $\mathbf{H}_{n,t}$ and their AoI values $\mathbf{\Xi}_{n,t}$. The estimator function $\phi_n(\cdot,\cdot)$ is quite general, it could be an optimal or sub-optimal estimator.

5.2.2 Loss Model for Situational Awareness

The loss due to the unawareness of potential danger is characterized by a loss function $L : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$, where $L(y, \hat{y})$ is the incurred loss if $Y_{n,t} = y$ is the actual safety level and $\hat{y}$ is estimated value of the safety level. To better understand the behavior of the loss L, consider the example in Figure 5.1, where the fire and smoke indicates a *dangerous* situation, only smoke indicates a *cautious* situation, and otherwise, it is a *safe* situation. Let us consider the following parameters of the loss function L for the event depicted in Figure 5.1:

$$L(\textit{dangerous, safe}) = 1000,$$
$$L(\textit{safe, dangerous}) = 5,$$
$$L(\textit{dangerous, cautious}) = 100,$$
$$L(\textit{cautious, dangerous}) = 5,$$
$$L(\textit{cautious, safe}) = 10,$$
$$L(\textit{safe, cautious}) = 1.$$

If the *dangerous* situation is wrongly estimated as *safe*, the loss will be significantly high due to the huge risk of damage. Conversely, if the *safe* situation is wrongly estimated as *dangerous*, the loss will be small. This is because even if the estimation is incorrect, the risk of damage is small. The loss for wrongly estimating a *dangerous* situation as *cautious* also has some impact because a *dangerous* situation always involves high risk of damage. However, the parameters used above may change according to the application context. Moreover, the loss for perfect estimation is zero, i.e., $L(\text{safe, safe}) = L(\text{danger, danger}) = 0$. It might look counter-intuitive that the loss associated with a dangerous situation is 0. However, this is because even when the actual safety state is *dangerous*, the situational awareness is good due to perfect estimation[1].

The loss function L effectively captures the impact of situational awareness. Our results apply to general loss functions including the well-known loss functions in the literature such as 0-1 loss, quadratic loss, logarithmic loss, etc. The significance of the loss function L lies in the fact that it can address safety issues based on situational awareness, which existing loss functions cannot address.

[1]In this study, we are interested in maximizing situational awareness and not optimizing the control policy. Hence, we consider $L(\textit{dangerous, dangerous}) = L(\textit{cautious, cautious}) = 0$.The joint design of the scheduler and controller could be an interesting future direction. However, this is out of the scope of this study.

5.2.3 Scheduling Policy

Let $\pi = (\mu_n(0), \mu_n(1), \ldots)_{n=1}^{N}$ denote a scheduling policy, where $\mu_n(t) \in \{0, 1\}$ is the decision variable to schedule agent n at every time slot t. If agent n is scheduled for transmission at time slot t, then $\mu_n(t) = 1$; otherwise $\mu_n(t) = 0$. Let Π denote the set of all causal scheduling policies in which every decision is made by using the current and historical information available at the receiver. If agent n is scheduled for transmission at time slot t, then $\mu_n(t) = 1$; otherwise $\mu_n(t) = 0$. Let Π denote the set of all causal scheduling policies in which every decision is made by using the current and historical information available at the receiver.

If agent n is not scheduled for transmission, i.e., $\mu_n(t) = 0$, then AoI will increase by 1, i.e., $\Delta_n(t) = \Delta_n(t-1) + 1$. If agent n is scheduled for transmission, i.e., $\mu_n(t) = 1$ and the transmission succeeds (with probability p_n), then AoI will drops to 1, i.e., $\Delta_n(t) = 1$; otherwise, if transmission fails (with probability $1 - p_n$), then AoI will increase by 1, i.e., $\Delta_n(t) = \Delta_n(t-1) + 1$.

5.3 Problem Formulation

Our goal is to find an optimal scheduling policy that minimizes the time-average sum of the expected loss of the N agents due to the unawareness of potential danger. The scheduling problem is formulated as

$$\mathsf{L}_{\text{opt}} = \inf_{\pi \in \Pi} \limsup_{T \to \infty} \sum_{n=1}^{N} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}_{\pi}\big[L(Y_{n,t}, \phi_n(\mathbf{\Xi}_{n,t}, \mathbf{H}_{n,t}))\big] \tag{5.1}$$

$$\text{s.t.} \sum_{n=1}^{N} \mu_n(t) \leq M, \mu_n(t) \in \{0, 1\}, t = 0, 1, \ldots, \tag{5.2}$$

where L_{opt} is the optimum value of (5.1). Because our system consists of M channels, $\sum_{n=1}^{N} \mu_n(t) \leq M$ is required to hold for all time slot t.

Problem (5.1)-(5.2) is a Restless Multi-armed Bandit (RMAB) because the age process of associated with each agent n continues to evolve regardless of whether agent n is selected for transmission [127]. Solving RMAB problems and finding optimal solutions are generally challenging and PSPACE-hard [128]. To find a solution for (5.1)-(5.2) is much harder because it is a complicated RMAB with expanding state space. This complexity arises because it involves the history of the signal observations $\mathbf{H}_{n,t}$ and their corresponding AoI values $\Xi_{n,t}$ at every time slot t. Consequently, a new piece information will be added to the history at every time slot which makes the problem significantly more challenging to solve. However, we are able to reduce the state space of problem (5.1)-(5.2) by utilizing the sufficient statistic of the history information. The details are provided in Section 5.4.

5.4 Problem Simplification

In order to simplify problem (5.1)-(5.2), we leverage the sufficient statistic of the history [126]. By using the sufficient statistic, we can significantly reduce the complexity of problem (5.1)-(5.2). In this sequel, we have the following theorem.

Theorem 5.1 *If $X_{n,t}$ is a Markov chain and $Y_{n,t}$ follows the relationship illustrated in Figure 5.2, then $(\Delta_n(t), X_{n,t-\Delta_n(t)})$ is a sufficient statistic of $(\Xi_{n,t}, \mathbf{H}_{n,t})$ for estimating $Y_{n,t}$.*

Then, RMAB (5.1)-(5.2) can be equivalently expressed as the following problem:

$$\mathsf{L}_{\text{opt}} = \inf_{\pi \in \Pi} \limsup_{T \to \infty} \sum_{n=1}^{N} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}_{\pi}\left[L(Y_{n,t}, \phi_n(\Delta_n(t), X_{n,t-\Delta_n(t)}))\right] \tag{5.3}$$

$$\text{s.t.} \sum_{n=1}^{N} \mu_n(t) \leq M, \mu_n(t) \in \{0,1\}, t = 0, 1, \ldots. \tag{5.4}$$

By this, we obtain an equivalent RMAB (5.1)-(5.2) with a reduced state space, where the age $\Delta_n(t)$ and the latest received observation $X_{n,t-\Delta_n(t)}$ of agent n at time slot t are the state of the n-th bandit.

In literature, there exists numerous studies that utilized belief states (i.e., the probability distribution over all possible states) to reduce the state space of each bandit n in RMAB problems [53, 121, 124, 125]. Such formulations render the state space uncountable and leads to the curse of dimensionality. Although [53, 121, 124, 125] made the state space countable under a positive recurrent assumption and used a sufficiently large truncated age value, the state space still exhibits a quadratic increase with the age. For a truncated set $\{1, 2, \ldots, \tau\}$

of AoI values, the state space increases as $\tau \times |\mathcal{X}_n|^2$, where $X_{n,t} \in \mathcal{X}_n$ represents the n-th bandit process. The difference between the formulation in [53, 121, 124, 125] and problem (5.3)-(5.4) is that we do not need to utilize belief states. Instead, we use the latest received observation and the corresponding age value. As a result, our state space remains much smaller that demonstrate linear growth with age, such as $\tau \times |\mathcal{X}_n|$. Consequently, the state space remains much simpler than [53, 121, 124, 125].

5.4.1 Restless Multi-armed Bandit: Relaxation and Lagrangian Decomposition

To find an optimal solution to RMAB (5.3)-(5.4) is quite challenging and is still PSPACE-hard. Because constraint (5.4) need to be satisfied at every time slot t. A Whittle index Policy is known to be an efficient approach to solve RMAB problems which requires to satisfy a condition called indexability [127], [113].

A key challenge in solving problem (5.3)-(5.4) is that indexability is very difficult to establish. This difficulty arises due to the following reasons: (i) The state of each bandit of RMAB (5.3)-(5.4) exhibits a complicated transition, (ii) the transmission channels are unreliable, and (iii) the expected penalty associated with each bandit can be either monotonic or non-monotonic function of the AoI while most of the previous studies considered monotonic penalty function of AoI [4, 41, 64, 112] is a non-monotonic function of the age. In addition, we allow the estimator function $\phi_n(\cdot, \cdot)$ to be both optimal and sub-optimal where most of the prior studies considered optimal estimators [53, 121]. Hence, (5.3)-(5.4) is a more challenging problem than the problems studied in [4, 41, 53, 64, 112, 121] and the requirements to solve RMAB (5.3)-(5.4) become quite complicated. However, we are able to develop a Maximum Gain First policy that does not need to satisfy indexability.

5.4.2 Relaxation and Lagrangian Decomposition

In standard RMAB problems, the constraint (5.4) needs to be satisfied with equality, i.e., exactly M bandits are activated at any time slot t. However, in our problem, constraint (5.4) activates less than M bandits at any time t. Following [115, Section 5.1.1] , [129, Section IV-A], we introduce M additional *dummy bandits* that never change state and therefore, they incur 0 cost. If a *dummy bandit* is activated, it occupies one channel but does not incur any cost. Let $\mu_0(t) \in \{1, 2, \ldots, M\}$ be the number of *dummy bandits* that are activated at time slot t. After incorporating these *dummy bandits*, the RMAB (5.3)-(5.4) can be expressed as

$$\mathsf{L}_{\text{opt}} = \inf_{\pi\in\Pi} \limsup_{T\to\infty} \sum_{n=1}^{N} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}_\pi\big[L(Y_{n,t}, \phi_n(\Delta_n(t), X_{n,t-\Delta_n(t)}))\big] \tag{5.5}$$

$$\text{s.t.} \sum_{n=0}^{N} \mu_n(t) = M, \mu_0(t) \in \{1, 2, \ldots, M\}, t = 0, 1, \ldots,$$

$$\mu_n(t) \in \{0, 1\}, n = 1, 2, \ldots, t = 0, 1, \ldots, \tag{5.6}$$

which is an RMAB with an equality constraint. Because the *dummy bandits* never change state, problem (5.3)-(5.4) and (5.5)-(5.6) are equivalent. Therefore, a policy that optimizes (5.3)-(5.4) will also optimize (5.5)-(5.6).

Next, we follow the standard relaxation and Lagrangian decomposition procedure for RMAB [127] and relax the constraint (5.6) and obtain the following relaxed problem:

$$\mathsf{L}_{\text{opt}} = \inf_{\pi\in\Pi} \limsup_{T\to\infty} \sum_{n=1}^{N} \mathbb{E}_\pi\left[\frac{1}{T} \sum_{t=0}^{T-1} L(Y_{n,t}, \phi_n(\Delta_n(t), X_{n,t-\Delta_n(t)}))\right], \tag{5.7}$$

$$\text{s.t.} \limsup_{T\to\infty} \sum_{n=1}^{N} \mathbb{E}_\pi\left[\frac{1}{T} \sum_{t=0}^{T-1} \mu_n(t)\right] = M,$$

$$\mu_0(t) \in \{1, 2, \ldots, M\}, t = 0, 1, \ldots,$$

$$\mu_n(t) \in \{0, 1\}, n = 1, 2, \ldots, t = 0, 1, \ldots. \tag{5.8}$$

The relaxed constraint (5.8) needs to be satisfied on average, instead of satisfying at every time slot t. To solve the relaxed problem (5.7)-(5.8), we take a dual cost $\lambda \geq 0$ (also known as Lagrange multiplier) for the relaxed constraint. The dual problem is given by

$$\sup_{\lambda\geq 0} \bar{L}(\lambda), \tag{5.9}$$

where

$$\bar{L}(\lambda) = \inf_{\pi\in\Pi} \limsup_{T\to\infty} \mathbb{E}_{\pi}\left[\sum_{n=1}^{N}\frac{1}{T}\sum_{t=0}^{T-1} L(Y_{n,t}, \phi_n(\Delta_n(t), X_{n,t-\Delta_n(t)}))+\lambda\left(\sum_{n=0}^{N}\mu_n(t) - M\right)\right]. \tag{5.10}$$

The term $\frac{1}{T}\sum_{t=0}^{T-1}\sum_{n=0}^{N}\lambda M$ in (5.10) does not depend on policy π and hence can be removed. For a given λ, problem (5.10) can be decomposed into $(N+1)$ separated sub-problems and each sub-problem associated with agent n is formulated as

$$\bar{L}_n(\lambda) = \inf_{\pi_n\in\Pi_n} \limsup_{T\to\infty} \mathbb{E}_{\pi_n}\left[\frac{1}{T}\sum_{t=0}^{T-1} L(Y_{n,t}, \phi_n(\Delta_n(t), X_{n,t-\Delta_n(t)})) + \lambda\mu_n(t)\right], \tag{5.11}$$

where $\bar{L}_n(\lambda)$ is the optimum value of (5.11), $\pi_n = (\mu_n(0), \mu_n(1), \ldots)$ is the sub-scheduling policy for agent n, and Π_n is the set of all causal sub-scheduling policies of agent n. Problem (5.11) is a per-bandit problem associated with bandit n. On the other hand, the sub-problem associated with the *dummy bandits* is given by

$$\bar{L}_0(\lambda) = \inf_{\pi_0\in\Pi_0} \limsup_{T\to\infty} \mathbb{E}_{\pi_0}\left[\frac{1}{T}\sum_{t=0}^{T-1}\lambda\mu_0(t)\right], \tag{5.12}$$

where $\bar{L}_0(\lambda)$ is the optimum value of (5.12), $\pi_0 = \{\mu_0(t), t = 0, 1, \ldots\}$, and Π_0 is the set of all causal activation policies π_0.

5.4.3 MDP Framework of the Decomposed Problem

Given transmission cost λ, Problem (5.11) can be cast as an average-cost infinite horizon MDP. The components of MDP (5.11) are described below:

- State: At any time slot t, the state of the MDP is the age $\Delta_n(t)$ and the latest received observation $X_{n,t-\Delta_n(t)}$ of agent n.
- Action: At any time slot t, the action is defined by $\mu_n(t) \in \{0, 1\}$ which denotes the scheduling decision for agent n. Each bandit n associated with problem (5.3)-(5.4) is a Markov Decision Process (MDP) with two actions: active and passive. If a packet from

agent n is requested and submitted to a channel at time slot t, then restless bandit n takes an active action at time slot t; otherwise, bandit n is made passive at time slot t.

- State Transitions: The AoI process $\Delta_n(t)$ for agent n evolves as follows

$$\Delta_n(t) = \begin{cases} 1, & \text{w. prob.}\, p_n, \quad \text{if}\, \mu_n(t) = 1, \\ \Delta_n(t-1)+1, & \text{w. prob.}\, 1-p_n, \,\text{if}\, \mu_n(t) = 1, \\ \Delta_n(t-1)+1, & \text{w. prob.}\, 1, \quad \text{if}\, \mu_n(t) = 0. \end{cases}$$

  The latest received observation $X_{n,t-\Delta_n(t)}$ evolves as

$$X_{n,t-\Delta_n(t)} = \begin{cases} X_{n,t}, & \text{w. prob.}\, p_n, \quad \text{if}\, \mu_n(t) = 1, \\ X_{n,t-\Delta_n(t)}, & \text{w. prob.}\, 1-p_n, \,\text{if}\, \mu_n(t) = 1, \\ X_{n,t-\Delta_n(t)}, & \text{w. prob.}\, 1, \quad \text{if}\, \mu_n(t) = 0. \end{cases}$$

- Penalty: Given $\Delta_n(t) = \delta$ and $X_{n,t-\Delta_n(t)} = x$, the expected penalty for agent n between two consecutive transmissions is defined as

$$L_n(\delta, x) = \mathbb{E}[L(Y_{n,t}, \phi_n(\Delta_n(t), X_{n,t-\Delta_n(t)}))|\Delta_n(t) = \delta, X_{n,t-\Delta_n(t)} = x]. \qquad (5.13)$$

The function $L_n(\delta, x)$ has some interesting properties: (i) $L_n(\delta, x)$ exhibits information-theoretic interpretation depending on the estimator $\phi_n(\cdot,\cdot)$, (ii) $L_n(\delta, x)$ is non-monotonic with age δ given observation x. The details are provided in Section 5.5.

5.5 Properties of $L_n(\delta, x)$ in (5.13)

In this section, we analyze the behavior of the penalty function $L_n(\cdot,\cdot)$ in (5.13). Our analysis reveals two interesting observations: (i) $L_n(\cdot,\cdot)$ exhibits an information-theoretic behavior based on different properties of the estimator $\phi_n(\cdot,\cdot)$: (a) If $\phi_n(\cdot,\cdot)$ be the optimal estimator that minimizes (5.13), then $L_n(\cdot,\cdot)$ can be represented as L-conditional entropy,

(b) if $\phi_n(\cdot,\cdot)$ be the optimal estimator of another problem, i.e., $\phi_n(\cdot,\cdot)$ minimizes any problem other than (5.13), then $L_n(\cdot,\cdot)$ can be represented as L-conditional cross-entropy, and (c) if $\phi_n(\cdot,\cdot)$ is not an estimator, then $L_n(\cdot,\cdot)$ remains the same as (5.13), (ii) For a given received observation x, $L_n(\cdot,\cdot)$ is not necessarily a monotonic function of the age. Both of these observations provide useful insights and understanding to solve (5.3)-(5.4). The details are provided below:

5.5.1 Information-theoretic Interpretation

Assume that $Y_{n,t}$ is conditionally independent of $\Delta_n(t)$ given $X_{n,t-\Delta_n(t)}$. Under this assumption, we present the information-theoretic analysis of $L_n(\delta, x)$ in the following three cases:

Case 1: If $\phi_n(\cdot,\cdot)$ is the optimal estimator of the underlying data distribution given by

$$\phi_n(\delta, x) = \underset{\hat{y}}{\operatorname{argmin}}\, \mathbb{E}_{Y\sim P_{Y_{n,t}|X_{n,t-\delta}=x}}[L(Y,\hat{y})], \tag{5.14}$$

then (5.13) can be expressed as

$$L_n(\delta, x) = H_L(Y_{n,t}|X_{n,t-\delta} = x), \tag{5.15}$$

where $H_L(Y_{n,t}|X_{n,t-\delta} = x)$ is the generalized conditional entropy of $Y_{n,t}$ given the latest received observation generated δ times ago at time slot t [122, 123], [120]. The generalized conditional entropy is given by

$$H_L(Y_{n,t}|X_{n,t-\delta} = x) = \min_{\hat{y}\in\mathcal{Y}} \mathbb{E}_{Y\sim P_{Y_{n,t}|X_{n,t-\delta}=x}}[L(Y,\hat{y})]. \tag{5.16}$$

For optimal estimator $\phi_n(\delta, x)$ in (5.14), $L_n(\delta, x)$ becomes

$$L_n(\delta, x) = \min_{\hat{y}\in\mathcal{Y}} \mathbb{E}_{Y\sim P_{Y_{n,t}|X_{n,t-\delta}=x}}[L(Y,\hat{y})], \tag{5.17}$$

which provides a lower bound of (5.17). This lower bound is closely related to the concept of generalized entropy [122, 123] or specifically, the L-entropy [120]. For a random variable

Y, the L-entropy is given by

$$H_L(Y) = \min_{a \in \mathcal{A}} \mathbb{E}_{Y \sim P_Y}[L(Y, a)]. \tag{5.18}$$

Let a_Y be the optimal solution to (5.18), or specifically the optimal estimator associated with the random variable Y, which is also called a Bayes estimator [122]. The L-conditional entropy of Y given $X = x$ can be defined as [120, 122, 123]

$$H_L(Y|X = x) = \min_{a \in \mathcal{A}} \mathbb{E}_{Y \sim P_{Y|X=x}}[L(Y, a)]. \tag{5.19}$$

Comparing (5.14), (5.17), and (5.16), it is evident that for the optimal estimator, $L_n(\delta, x)$ is indeed L-conditional entropy. By this, we obtain an information-theoretic lower bound for $L_n(\delta, x)$. It represents the minimum achievable penalty that characterizes performance degradation due to the lack of knowledge of the situation.

Case 2: If $\phi_n(\cdot, \cdot)$ is the optimal estimator of another data distribution given by,

$$\phi_n(\delta, x) = \underset{\hat{y}}{\operatorname{argmin}}\, \mathbb{E}_{Y \sim P_{\tilde{Y}_{n,t}|\tilde{X}_{n,t-\delta}=x}}[L(Y, \hat{y})], \tag{5.20}$$

then (5.13) can be expressed as

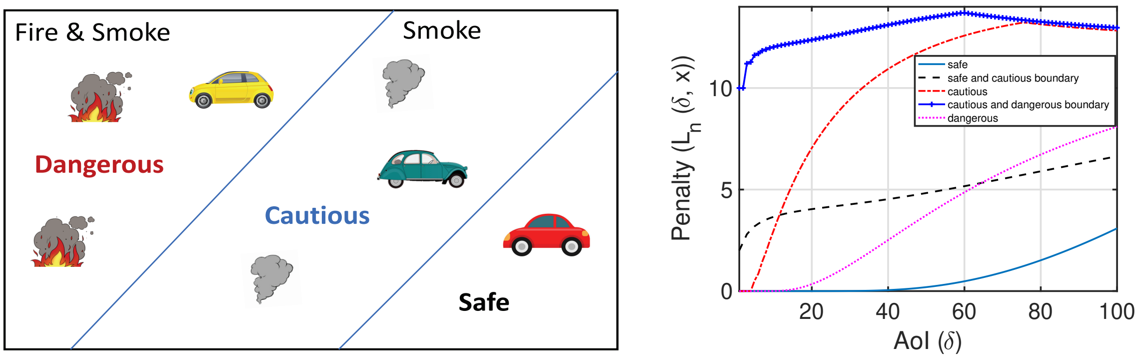
$$q_n(\delta, x) = H_L(P_{Y_{n,t}|X_{n,t-\delta}=x}; P_{\tilde{Y}_{n,t}|\tilde{X}_{n,t-\delta}=x}), \tag{5.21}$$

which is the L-conditional cross entropy [120] between $Y_{n,t}$ and $\tilde{Y}_{n,t}$ given the age and the latest received observation at time slot t. The L-conditional cross-entropy is given by

$$H_L(P_{Y_{n,t}|X_{n,t-\delta}=x}; P_{\tilde{Y}_{n,t}|\tilde{X}_{n,t-\delta}=x}) = \mathbb{E}_{P_{Y_{n,t}|X_{n,t-\delta}=x}}[L(Y, \hat{y}_{P_{\tilde{Y}_{n,t}|\tilde{X}_{n,t-\delta}=x}}]. \tag{5.22}$$

For the optimal estimator of (5.20) $L_n(\delta, x)$ can be expressed as

$$L_n(\delta, x) = \mathbb{E}_{P_{Y_{n,t}|X_{n,t-\delta}=x}}[L(Y, \hat{y}_{P_{\tilde{Y}_{n,t}|\tilde{X}_{n,t-\delta}=x}}]. \tag{5.23}$$



(a) Safety regions

(b) Penalty vs AoI given observation

Figure 5.3: (a) Safety regions and (b) Penalty ($L_n(\delta, x)$) vs AoI (δ) for given observations.

This result in (5.23) is closely related to L-conditional cross entropy [120]. The L-cross entropy between random variables Y and $\tilde{Y}$ is given by

$$H_L(Y; \tilde{Y}) = \mathbb{E}_{Y\sim P_Y}[L(Y, a_{\tilde{Y}})], \tag{5.24}$$

where $a_{\tilde{Y}}$ is the optimal estimator associated with random variable $\tilde{Y}$. In addition, the L-conditional cross-entropy between Y and $\tilde{Y}$ given $X = x$ is

$$H_L(Y; \tilde{Y}|X = x) = \mathbb{E}_{Y\sim P_{Y|X=x}}[L(Y, a_{\tilde{Y}|\tilde{X}=x})], \tag{5.25}$$

where $a_{\tilde{Y}|\tilde{X}=x}$ is the optimal estimator associated with random variables $\tilde{Y}|\tilde{X} = x$. Comparing (5.20), (5.22), and (5.23), we can conclude that if $\phi_n(\cdot,\cdot)$ is the optimal estimator of another data distribution, then $L_n(\delta, x)$ becomes L-conditional cross entropy.

Case 3: If $\phi_n(\cdot,\cdot)$ is not any optimal estimator, then $L_n(\delta, x)$ remains the same as (5.13).

5.5.2 Non-monotonic Information Aging

Our analysis reveals that $L_n(\delta, x)$ can be a non-monotonic function of age, particularly when the knowledge of the surrounding situation is considered. Also, we analyze $L_n(\delta, x)$ with received observation x given δ which demonstrates that the region for frequent updating increases with δ.

To do this experiment, we consider a safety-critical system where N robots are moving in a region illustrated in Figure 5.3. This region is equally divided into 400 positions and the received observation $X_{n,t}$ of robot n is represented by the position $x = (x_1, x_2)$ of robot n at time t. The safety level $Y_{n,t}$ is divided into three regions: $\{$*safe, cautious, dangerous*$\}$. A robot n can randomly move in any four directions: *up, down, left*, and *right* with equal probability 0.2. If robot n is in the leftmost position, then moving left means it will stay in the same position, similar criteria are applied for the rightmost, upmost, and downmost positions. The losses considered in this experiment are: L*(cautious, safe)* $= 50,$ L*(safe, cautious)* $= 10,$ L*(dangerous, safe)* $= 200,$ L*(safe, dangerous)* $= 10,$ L*(dangerous, cautious)* $= 50$, L*(cautious, dangerous)*$= 20$, and L*(dangerous, dangerous)* $=$ L*(cautious, cautious)* $=$ L*(safe, safe)* $= 0$. We consider the optimal estimator in this experiment.

$L_n(\delta, x)$ for fixed x and varying δ

The penalty vs AoI curve for the given observation is illustrated in Figure 5.3(b). When a robot is in a *safe*, *cautious*, or *dangerous* region which is far from the *safe* and *cautious* or *cautious* and *dangerous* boundary, the penalty is initially close to zero for small AoI values and increases gradually with increasing age. This phenomenon tells us that we do not need to update frequently when a robot is far from the boundary region. However, if the robot moves closer to the boundary between *safe* and *cautious* or *cautious* and *dangerous*, the penalty increases very quickly because of the uncertainty of its position in the subsequent time slots. Hence, we need to update very frequently if any robot is close to the boundary. Because situational awareness is not good at the boundary, frequent updating is required to keep the loss small. With the increase in age, the curves at the boundary start decreasing, specifically, when AoI becomes large. This is because, for larger AoI values, the best decision for the estimator is to estimate a "*dangerous*" and as the loss for wrongly estimating "*safe*" or "*cautious*" as "*dangerous*" has less impact, the penalty reduces. Similarly, the other curves can be explained.

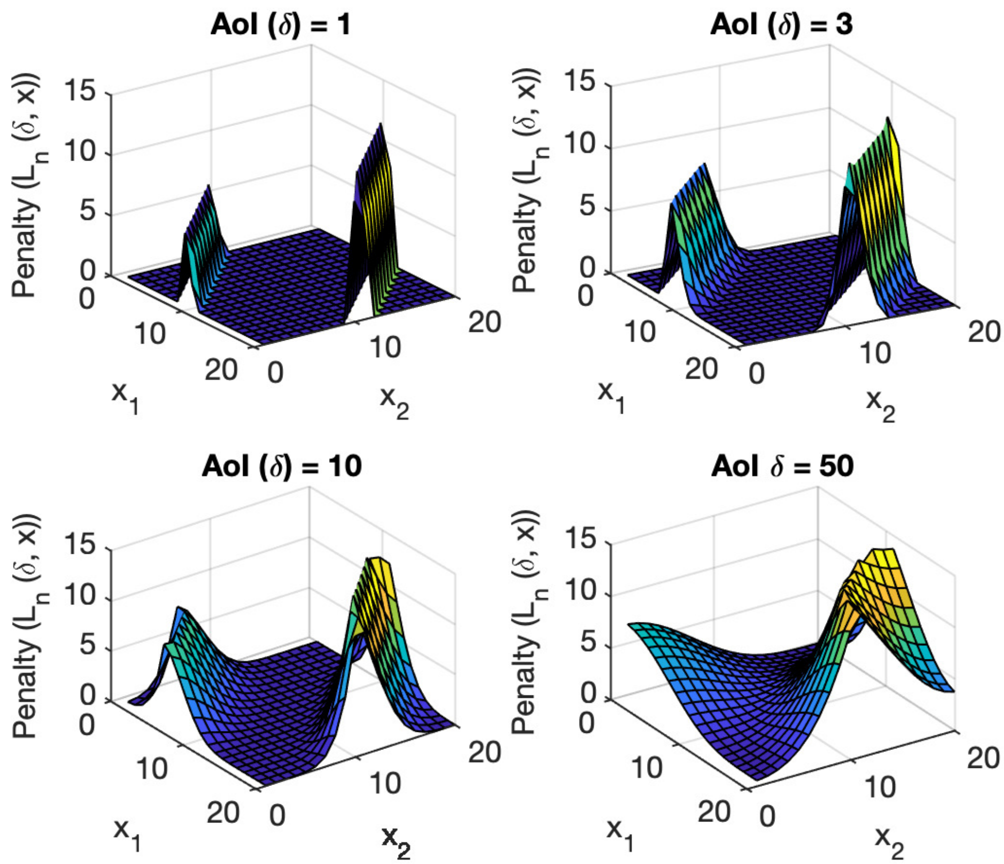


Figure 5.4: Penalty $L_n(\delta, x)$ vs received observation x for fixed AoI δ.

$L_n(\delta, x)$ **for fixed** δ **and varying** x

Figure 5.4 demonstrates the penalty vs received observation x curve for different AoI δ values. In this figure, when AoI is small, i.e., $\delta = 1$, the loss is high only at the two boundary regions which illustrates that we need to update frequently if the received observation is at the boundary. With increasing δ, the loss curve spreads to the adjacent regions of the boundaries. Hence, the region for frequent updating is also increasing with increasing δ.

Because the penalty curves are not necessarily monotonic with age, only considering the non-decreasing functions of the age is not sufficient for performance analysis of safety-critical systems. The proposed metrics in prior works, i.e., AoII, VoI, AoS, QAoI cannot explain this non-monotonicity with age.

5.6 Penalty Minimization: An Information-theoretic Perspective

In Section 5.5.1, we demonstrate that for optimal estimators, the penalty function $L_n(\delta, x)$ can be represented as L-conditional entropy. By leveraging this insight, we obtain that for optimal estimators, always sending updates, or specifically taking active actions at every time slot benefits the system by reducing the average penalty of the system.

Assume the estimator $\phi_n(\cdot)$ in (5.3) is the optimal estimator. In addition, we have available channel resources at every time slot. By this, constraint (5.4) will always be satisfied. Under these assumptions and by utilizing (5.15) in (5.3), we can write problem (5.1) as

$$\mathsf{L}_{\text{opt}} = \inf_{\pi \in \Pi} \limsup_{T \to \infty} \sum_{n=1}^{N} \mathbb{E}_{\pi}\left[\frac{1}{T}\sum_{t=0}^{T-1} H_L(Y_{n,t}|\Delta_n(t), X_{n,t-\Delta_n(t)})\right]. \tag{5.26}$$

Problem (5.26) can be decomposed into N separated sub-problems and each sub-problem is a MDP that can be solved by Dynamic programming [126]. The optimal policy of the subproblem associated with agent n satisfies the following Bellman optimality equation:

$$\begin{aligned} J_n(\delta, x) = & H_L(Y_{n,\delta}|X_{n,0} = x) - \mathsf{L}_{n,\text{opt}} + \min\{J_n(\delta+1, x), (1-p_n)J_n(\delta+1, x) \\ & + p_n\mathbb{E}[J_n(1, X_{n,0})|X_{n,-\delta} = x]\}, \end{aligned} \tag{5.27}$$

where $J_n(\delta, x)$ is the value function associated with state (δ, x) and $\mathsf{L}_{n,\text{opt}}$ is the optimal value of the subproblem associated with agent n. Because $(Y_{n,t}|X_{n,t-\delta} = x)$ and $X_{n,t}$ are time-homogeneous, (5.27) holds.

As explained in [126], the optimal value function can be derived by using value iteration and the sequence of value functions $J_{n,k}(\delta, x)$ can be written as

$$\begin{aligned} J_{n,k+1}(\delta, x) = & H_L(Y_{n,\delta}|X_{n,0} = x) - \mathsf{L}_{n,\text{opt}} + \min\{J_{n,k}(\delta+1, x), (1-p_n)J_{n,k}(\delta+1, x) \\ & + p_n\mathbb{E}[J_{n,k}(1, X_{n,0})|X_{n,-\delta} = x]\}, \end{aligned} \tag{5.28}$$

which converges to $\lim_{k\to\infty} J_{n,k} = J_n$ for any $J_{n,0}$. After some rearrangements, we can write (5.28) as

$$J_{n,k+1}(\delta, x) = H_L(Y_{n,\delta}|X_{n,0} = x) - \mathsf{L}_{n,\text{opt}} + J_{n,k}(\delta+1, x) + p_n \min\{0, -J_{n,k}(\delta+1, x) + \mathbb{E}[J_{n,k}(1, X_{n,0})|X_{n,-\delta} = x]\}. \tag{5.29}$$

Then, we have the following lemma which illustrates sending is beneficial at every k.

Lemma 5.1 *For any k, it holds that $J_{n,k}(\delta+1, x) \geq \mathbb{E}[J_{n,k}(1, X_{n,0})|X_{n,-\delta} = x]$.*

Lemma 5.1 states that the penalty for not sending at iteration step k is higher than sending and therefore, taking the active action is beneficial to reduce the penalty. One interesting observation from (5.29) is that each time a packet is successfully delivered with probability p_n, a new piece of information about the agent's signal value is added with the existing information ($X_{n,t-\delta} = x$) (see the term $\mathbb{E}[J_{n,k}(1, X_{n,0})|X_{n,t-\delta} = x]$ in (5.28)). This new information plays a crucial role in reducing the system penalty and hence benefits the system through sending. In this sequel, we introduce the following useful lemma which illustrates that more information reduces the L-conditional entropy.

Lemma 5.2 *For random variables X, Y, and Z, it holds that $H_L(Y|Z = z) \geq H_L(Y|X, Z = z)$, where*

$$H_L(Y|Z = z) = \min_{a\in\mathcal{A}} \mathbb{E}[L(Y, a)|Z = z], \tag{5.30}$$

$$H_L(Y|X, Z=z) = \sum_{x\in\mathcal{X}} P(X=x|Z=z) H_L(Y|X=x, Z=z). \tag{5.31}$$

Proof 5.1 *See Appendix 5.A.*

Given Lemma 5.2, we are ready to prove Lemma 5.1. The proof of Lemma 5.1 is provided in Appendix 5.B.

Theorem 5.2 *For optimal estimators, the optimal policy π in (5.26) chooses the active action at every time slot t.*

Algorithm 10 Maximum Gain First Policy

1: At time $t = 0$:
2: Input λ^* which is the optimal solution to (5.9).
3: Input $\alpha_{n,\lambda^*}(\delta, x)$ in (5.36) for every agent n.
4: For all time $t = 0, 1, \ldots,$
5: Update $(\Delta_n(t), X_{n,t-\Delta_n(t)})$ for all agent n.
6: Update current "gain" $\alpha_{n,\lambda^*}(\Delta_n(t), X_{n,t-\Delta_n(t)})$ for all agent n.
7: Choose at most M agents with highest "gain".

Proof 5.2 *See Appendix 5.C.*

Theorem 5.2 states that if $\phi_n(\cdot, \cdot)$ is the optimal estimator, then it is always better to send. Though the penalty $L_n(\delta, x)$ is not necessarily a monotonic function of the age, the insights obtained from Lemma 5.2 tell us that having additional information helps reduce the average penalty. Theorem 5.2 holds for the situation when there are no channel resource constraints. The original problem stated in (5.3)-(5.4) has a channel resource constraint, therefore, all of the agents cannot submit their updates at every time slot when $N > M$. On the other hand, the estimator function $\phi_n(\cdot, \cdot)$ in (5.3)-(5.4) is not necessarily the optimal estimator. Theorem 5.2 holds for optimal estimators. To solve problem (5.3)-(5.4) for arbitrary estimators with a channel resource constraint, we have to design an efficient scheduling policy that minimizes the time-average sum of the expected penalty of the N sources ensuring that constraint (5.4) is satisfied. We provide the details in the next section.

5.7 Maximum Gain First Policy

For a given transmission cost λ, the per-bandit problem (5.11) can be cast as an average-cost infinite horizon MDP with state $(\Delta_n(t), X_{n,t-\Delta_n(t)})$. We solve (5.11) by using dynamic programming [126]. The Bellman optimality equation for the MDP in (5.11) is

$$h_{n,\lambda}(\delta, x) = \min_{\mu \in \{0,1\}} Q_{n,\lambda}(\delta, x, \mu), \tag{5.32}$$

where $h_{n,\lambda}(\delta, x)$ is the relative-value function of the average-cost MDP and $Q_{n,\lambda}(\delta, x, \mu)$ is the relative action-value function defined as

$$Q_{n,\lambda}(\delta, x, \mu) = \begin{cases} L_n(\delta, x) - \bar{L}_n(\lambda) + h_{n,\lambda}(\delta + 1, x), & \text{if } \mu = 0, \\ L_n(\delta, x) - \bar{L}_n(\lambda) + (1 - p_n)h_{n,\lambda}(\delta + 1, x) & \\ +p_n\mathbb{E}[h_{n,\lambda}(1, X_{n,0})|X_{n,\delta} = x] + \lambda, & \text{otherwise.} \end{cases} \quad (5.33)$$

The relative-value function $h_{n,\lambda}(\delta, x)$ can be computed by using the relative value iteration algorithm for average-cost MDP [126].

Let $\pi^*_{n,\lambda} = (\mu^*_{n,\lambda}(1), \mu^*_{n,\lambda}(2), \ldots)$ be an optimal solution to (5.11). The optimal decision at time slot t for agent n is given by

$$\mu^*_{n,\lambda}(t) = \underset{\mu \in \{0,1\}}{\text{argmin}}\, Q_{n,\lambda}(\Delta_n(t), X_{n,t-\Delta_n(t)}, \mu), \quad (5.34)$$

where the dual cost is iteratively updated using the dual sub-gradient ascent method with step size $\beta > 0$ [130]:

$$\lambda(j + 1) = \lambda(j) + \frac{\beta}{j}\left(\sum_{n=1}^{N} \mu_{n,\lambda(j)}(j) - M\right), \quad (5.35)$$

for j-th iteration. Let λ^* be the optimal dual cost to problem (5.9) to which $\lambda(t)$ converges. Then, we can apply $(\pi_{n,\lambda^*})_{n=1}^N$ for the relaxed problem (5.7)-(5.8). But applying this policy to the original problem (5.3)-(5.4) may violate the constraint (5.4). Following [53, 54], we define the "gain" $\alpha_{n,\lambda}(\delta, x)$ for choosing the action $\mu_{n,\lambda}(t)$ as

$$\alpha_{n,\lambda}(\delta, x) = Q_{n,\lambda}(\delta, x, 0) - Q_{n,\lambda}(\delta, x, 1). \quad (5.36)$$

If $Q_{n,\lambda}(\delta, x, 0) > Q_{n,\lambda}(\delta, x, 1)$, i.e., $\alpha_{n,\lambda}(\delta, x) > 0$, it is optimal to schedule agent n. If $Q_{n,\lambda}(\delta, x, 0) < Q_{n,\lambda}(\delta, x, 1)$, i.e., $\alpha_{n,\lambda}(\delta, x) < 0$, it is optimal to not to schedule agent n.

Substituting (5.33) into (5.36), we get

$$\alpha_{n,\lambda}(\delta, x) = p_n\bigg(h_{n,\lambda}(\delta+1, x) - \mathbb{E}[h_{n,\lambda}(1, X_{n,0})|X_{n,-\delta} = x]\bigg) - \lambda. \tag{5.37}$$

By utilizing the "gain" in (5.37) as the priority measurement for choosing action μ, we provide a low-complexity algorithm for solving problem (5.3)-(5.4) in Algorithm 10 which takes the optimal dual cost λ^* and the precomputed gain $\alpha_{n,\lambda^*}(\delta, x)$ associated with λ^* as input. Then, for all $t \geq 0$, the state $(\Delta_n(t), X_{n,t-\Delta_n(t)})$ and the associated "gain" $\alpha_{n,\lambda^*}(\Delta_n(t), X_{n,t-\Delta_n(t)})$ are updated. Finally, Algorithm 10 maximizes the "Net-gain" (total gain of all agents) of the system at time t. This is done by selecting at most M agents having the maximum "gain" at time slot t. The benefit of "Maximum Gain First Policy" in Algorithm 10 is that it does not need to satisfy the indexability condition.

5.8 Asymptotic Optimality

In this section, we demonstrate that the "Maximum Gain First Policy" in Algorithm 10 is asymptotically optimal in the same asymptotic regime as the Whittle index policy [127]. In this scenario, all N bandits are generalized to N classes, and the number of bandits in each class and the number of channels M are scaled by a parameter γ, while maintaining a constant ratio between them. Two bandits are said to be in the same class if they have identical penalty functions and transition probabilities. The dummy bandits belong to the same class. None of the N agents have the same penalty functions and transition probabilities. Therefore, we have $N + 1$ distinct class of bandits.

Let $V^{\gamma}_{\pi_{\text{gain}}}$ be the expected long-term average cost under policy π_{gain}. The policy π_{gain} will be asymptotically optimal if $V^{\gamma}_{\pi_{\text{gain}}} \leq V^{\gamma}_{\pi}$ for all $\pi \in \Pi$ as γ approaches ∞, while maintaining a constant ratio $\gamma N/\gamma M$. To prove the asymptotic optimality, (i) we first introduce a linear program in Section 5.8.1 for solving the relaxed optimization problem (5.7)-(5.8), (ii) next, by using the solution to the LP, we define a uniform global attractor in Section 5.8.2. In the relaxed problem (5.7)-(5.8), we have $N + M$ bandits with N agents and M dummy bandits, and M channels. We assume that in the state (δ, x), large AoI values, i.e., $\delta > \delta_{\text{high}}$ are rarely visited if δ_{high} is sufficiently large.

Let $W^n_{\delta,x}(t)$ be the fraction of class-n bandits in state (δ, x) at time t and $U^{n,\mu}_{\delta,x}(t)$ be the fraction of class-n bandits in state (δ, x) at time t for which decision $\mu \in \{0, 1\}$ is taken. If $\mu = 0$, no agent is scheduled for transmission; otherwise, if $\mu = 1$, an agent is scheduled for transmission. Given state (δ, x) and action μ of a class-n bandit, let $P^{(n,\mu)}_{(\delta',x'),(\delta,x)}$ be the transition probability from state (δ, x) to a state (δ', x') for a class-n bandit under action μ. Define

$$w^n_{\delta,x} = \limsup_{T\to\infty} \sum_{t=0}^{T-1} \frac{1}{T}\mathbb{E}[W^n_{\delta,x}(t)], \tag{5.38}$$

$$u^{n,\mu}_{\delta,x} = \limsup_{T\to\infty} \sum_{t=0}^{T-1} \frac{1}{T}\mathbb{E}[U^{n,\mu}_{\delta,x}(t)]. \tag{5.39}$$

If $\mu = 1$, a channel is occupied by a bandit. In this scenario, the time-average expected fraction of bandits from class-n occupying a channel is determined by

$$\sum_{(\delta,x)} u^{n,1}_{\delta,x}. \tag{5.40}$$

Let $\mathbf{V}^m(t)$, $\mathbf{U}^m(t)$, $\mathbf{v}^m$, and $\mathbf{u}^m$ be the vectors that contain $V^n_{\delta,x}(t)$, $U^{n,\mu}_{\delta,x}(t)$, $v^n_{\delta,x}$, and $u^{n,\mu}_{\delta,x}$, respectively, for all δ, x, and μ.

5.8.1 Linear Program for Solving Problem (5.7)-(5.8)

By utilizing $u^{n,\mu}_{\delta,x}$ in (5.39), the on average constraint in (5.8) can be written as

$$\sum_{n=0}^{N} \sum_{\mu=1} u^{n,\mu}_{\delta,x} = N. \tag{5.41}$$

Let $\bar{L}_{\text{rel}}$ be the optimal objective value of the relaxed problem (5.7)-(5.8). By solving the following LP, we can obtain $\bar{L}_{\text{rel}}$:

$$\min_{(\mathbf{u}^n)_{n=0}^N} \sum_{n=1}^{N} \sum_{\delta,x,\mu} \bar{L}_n(\delta,x) u_{\delta,x}^{n,\mu} \tag{5.42}$$

$$\text{s.t.} \sum_{n=0}^{N} \sum_{\mu=1} u_{\delta,x}^{n,\mu} = N, \tag{5.43}$$

$$\sum_{\mu} u_{\delta,x}^{n,\mu} = \sum_{\delta',x',\mu} u_{\delta',x'}^{n,\mu} P_{(\delta,x),(\delta',x')}^{(n,\mu)}, \forall n, \delta, x, \tag{5.44}$$

$$\sum_{\delta,x,\mu} u_{\delta,x}^{n,\mu} = 1, \forall n, \tag{5.45}$$

$$0 \leq \mathbf{u}^n \leq 1, \forall n. \tag{5.46}$$

5.8.2 Uniform Global Attractor Condition

For a policy π, we can have the following mapping

$$\Psi_\pi((\mathbf{w}^n)_{n=1}^N) = \mathbb{E}_\pi[(\mathbf{W}^n(t+1))_{n=1}^N | (\mathbf{W}^n(t))_{n=1}^N = (\mathbf{w}^n)_{n=1}^N]. \tag{5.47}$$

We define the t-th iteration of the maps $\Psi_{\pi,t\geq 0}(\cdot)$ as follows

$$\Psi_{\pi,0}((\mathbf{w}^n)_{n=1}^N) = (\mathbf{w}^n)_{n=1}^N, \tag{5.48}$$

$$\Psi_{\pi,t+1}((\mathbf{w}^n)_{n=1}^N) = \Psi_\pi(\Psi_{\pi,t}((\mathbf{w}^n)_{n=1}^N)). \tag{5.49}$$

Definition 5.1 ***Uniform Global attractor.*** *An equilibrium point* $(w^{n^*})_{n=1}^N$ *given by the optimal solution of* (5.42)-(5.46) *is a uniform global attractor of* $\Psi_{\pi,t\geq 0}(\cdot)$*, i.e., for all* $\epsilon > 0$*, there exists* $T(\epsilon)$ *such that for all* $t \geq T(\epsilon)$ *and for all* $(w^{n^*})_{n=1}^N)$*, one has* $||\Psi_{\pi,t}(((w^n)_{n=1}^N)) - (w^{n^*})_{n=1}^N))||_1 \leq \epsilon$.

Theorem 5.3 *Under Definition 5.1, the policy* π_{gain} *is asymptotically optimal.*

5.9 Numerical Results

In this section, we evaluate the performance of the following policies:

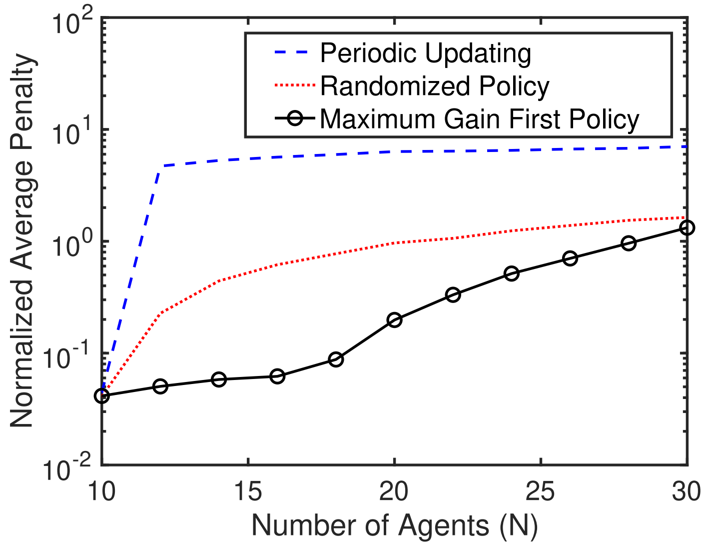


Figure 5.5: Normalized average penalty vs Number of agents (N) where Number of channels are $M = 10$ with success probability 0.95.

- Periodic Updating: The agents generate updates at every time slot and store them in a FIFO queue. Whenever a channel is available, an update from the queue is sent.
- Randomized Policy: If M channel resources are available, this policy randomly selects at most M agents.
- Maximum Gain First Policy: The policy provided in Algorithm 10.

We consider the same experimental setup of Figure 5.3 where 5 robots follow a deterministic policy (they follow a fixed path). The cost associated with these 5 robots is zero at every time slot because given an initial state, the position of these robots can be uniquely determined by following the deterministic policy. The goal of the other $N - 5$ robots is to move and scan the environment (e.g., Mars Rovers [131]) and send updates when requested. We do not consider any termination state for these robots, the goal is to keep scanning for an infinite time horizon. Our system consists of $M = 10$ erasure channels and the success probability is 0.95.

The performance comparison of the three policies mentioned above is provided in Figure 5.5. The normalized average penalty in Figure 5.5 is obtained by dividing time-average cost by the number of robots. From the figure, until $N \leq M$, all of the three policies show the same performance. Whenever $N > M$, periodic updating starts getting worse because the queue length is getting higher. In our simulation, we have used a buffer size of 20 for periodic updating. Moreover, the randomized policy randomly selects at most 5 agents for sending updates, whereas the maximum gain first policy decides a smarter way by considering the AoI and the state of the surrounding situation. The performance gain of the maximum gain first policy is up to 100 times compared to periodic updating and up to 10 times compared to the randomized policy.

5.10 Conclusion

We address the importance of situational awareness in safety-critical systems. The general loss function L has practical importance and an appropriate design of L can address many safety-critical issues. In the future, we will study systems where multiple agents can arrive and leave the system at any time. Another interesting direction is to consider a finite time horizon problem where there is a termination state while encountering a danger.

Appendix

5.A Proof of Lemma 5.2

From the definition of L-conditional entropy in (5.19), we get that

$$\begin{aligned}
& H_L(Y|Z=z) \\
=& \min_{a\in\mathcal{A}} \mathbb{E}[L(Y,a)|Z=z], \\
=& \min_{a\in\mathcal{A}} \sum_{y\in\mathcal{Y}} P(Y=y|Z=z)L(y,a), \\
=& \min_{a\in\mathcal{A}} \sum_{y\in\mathcal{Y}}\sum_{x\in\mathcal{X}} P(Y=y|X=x,Z=z)P(X=x|Z=z)L(y,a), \\
=& \min_{a\in\mathcal{A}} \sum_{x\in\mathcal{X}} P(X=x|Z=z)\sum_{y\in\mathcal{Y}} P(Y=y|X=x,Z=z)L(y,a), \\
\geq& \sum_{x\in\mathcal{X}} P(X=x|Z=z)\min_{a\in\mathcal{A}}\sum_{y\in\mathcal{Y}} P(Y=y|X=x,Z=z)L(y,a),
\end{aligned} \tag{5.50}$$

where (5.50) holds because $\min(f(w)+g(w)) \geq \min f(w) + \min g(w)$ for all w. Continuing from (5.50), we get that

$$\begin{aligned}
& H_L(Y|Z=z) \\
\geq& \sum_{x\in\mathcal{X}} P(X=x|Z=z)\min_{a\in\mathcal{A}} \mathbb{E}[L(Y,a)|X=x,Z=z].
\end{aligned} \tag{5.51}$$

Utilizing (5.30), we obtain that [120, 122, 123]

$$H_L(Y|X=x,Z=z) = \min_{a\in\mathcal{A}} \mathbb{E}[L(Y,a)|X=x,Z=z]. \tag{5.52}$$

Substituting (5.52) into (5.51) yields

$$\begin{aligned}
& H_L(Y|Z=z) \\
\geq& \sum_{x\in\mathcal{X}} P(X=x|Z=z)H_L(Y|X=x,Z=z), \\
=& H_L(Y|X,Z=z),
\end{aligned} \tag{5.53}$$

where (5.53) follows from (5.31). This completes the proof.

5.B Proof of Lemma 5.1

Without loss of generality, we can assume that for all (δ, x), $J_{n,0}(\delta, x) = 0$. At $k = 0$, (5.28) becomes

$$J_{n,1}(\delta, x) = H_L(Y_{n,\delta}|X_{n,0} = x) - \mathsf{L}_{n,\text{opt}}. \tag{5.54}$$

In this case, a minimum value can be achieved by both sending and not sending. Hence, we can conclude that sending is beneficial at iteration step k=0. Next, at $k = 1$, (5.28) becomes

$$\begin{aligned} J_{n,2}(\delta, x) = &H_L(Y_{n,\delta}|X_{n,0} = x) - \mathsf{L}_{n,\text{opt}} + \min\{J_{n,1}(\delta+1, x), (1-p_n)J_{n,1}(\delta+1, x) \\ &+ p_n\mathbb{E}[J_{n,1}(1, X_{n,0})|X_{n,-\delta} = x]\}. \end{aligned} \tag{5.55}$$

Sending will be beneficial at $k = 1$ if

$$J_{n,1}(\delta+1, x) \geq \mathbb{E}[J_{n,1}(1, X_{n,0})|X_{n,-\delta} = x]. \tag{5.56}$$

From the right-side term in (5.56), we get

$$\begin{aligned} &\mathbb{E}[J_{n,1}(1, X_{n,0})|X_{n,-\delta}] \\ =&\sum_{z\in\mathcal{X}} J_{n,1}(1, z)P(X_{n,0} = z|X_{n,-\delta} = x) \\ =&\sum_{z\in\mathcal{X}} H_L(Y_{n,t}|X_{n,0} = z)P(X_{n,0} = z|X_{n,-\delta} = x) - \mathsf{L}_{n,\text{opt}}, \\ =&H_L(Y_{n,t}|X_{n,0}, X_{n,-\delta} = x) - \mathsf{L}_{n,\text{opt}}, \end{aligned} \tag{5.57}$$

where (5.57) holds from Lemma 5.2.

In (5.54), no new information is obtained because $J_{n,1}(\delta+1, x) = H_L(Y_{n,\delta+1}|X_{n,0} = x) - \mathsf{L}_{n,\text{opt}}$. However, in (5.57), a new piece of information is added. By comparing (5.54) and (5.57) and by utilizing Lemma 5.2, we can conclude that

$$H_L(Y_{n,\delta+1}|X_{n,0} = x) \geq H_L(Y_{n,t}|X_{n,0}, X_{n,-\delta} = x) \tag{5.58}$$

from which (5.56) follows. Therefore, sending is beneficial for $k = 1$. Next, at $k = 2$, sending will be beneficial if

$$J_{n,2}(\delta + 1, x) \geq \mathbb{E}[J_{n,2}(1, X_{n,0})|X_{n,-\delta} = x]. \tag{5.59}$$

Utilizing (5.54), (5.56), and (5.57) in (5.55) yields

$$\begin{aligned} J_{n,2}(\delta, x) =& H_L(Y_{n,\delta}|X_{n,0} = x) - 2\mathsf{L}_{n,\text{opt}} + (1 - p_n)H_L(Y_{n,\delta+1}|X_{n,0} = x) + \\ & p_n H_L(Y_{n,1}|X_{n,0}, X_{n,-\delta} = x). \end{aligned} \tag{5.60}$$

Similarly, the right-side term of (5.59) is given by

$$\begin{aligned} \mathbb{E}[J_{n,2}(1, X_{n,0})|X_{n,-\delta}=x] =& H_L(Y_{n,1}|X_{n,0}, X_{n,-\delta}=x) - 2\mathsf{L}_{n,\text{opt}} + (1-p_n)H_L(Y_{n,2}|X_{n,0}, X_{n,-\delta}=x) \\ & + p_n H_L(Y_{n,2}|X_{n,0}, X_{n,1}, X_{n,-\delta}=x). \end{aligned} \tag{5.61}$$

From (5.60) and (5.61), we observe that every term in (5.61) has a piece of additional information. Then, by utilizing Lemma 5.2, we can conclude that (5.59) holds. Therefore, sending is beneficial at $k = 2$.

Next, assume that this result holds up to interation step k so that the following is true:

$$J_{n,k}(\delta + 1, x) \geq \mathbb{E}[J_{n,k}(1, X_{n,0})|X_{n,-\delta} = x]\}. \tag{5.62}$$

By using the result in (5.63), we have to prove that for iteration step $k + 1$, the following holds.

$$J_{n,k+1}(\delta + 1, x) \geq \mathbb{E}[J_{n,k+1}(1, X_{n,0})|X_{n,-\delta} = x]\}. \tag{5.63}$$

The right side of (5.63) contains $H_L(Y_{n,k}|X_{n,0}, X_{n,1}, \ldots, X_{n,k}, X_{n,-\delta} = x)$ which has more information than the left-side term. Similarly, it can be shown that all the terms on the right side of (5.63) have more information than the left side. By utilizing Lemma 5.2, we can conclude that sending is beneficial at iteration step $k + 1$.

Hence, it is beneficial to send for all k if a channel is available. This concludes the proof.

5.C Proof of Theorem 5.2

From Lemma 5.1, we get that sending is beneficial at every iteration step k. Hence, sending is beneficial at every time slot. Therefore, Theorem 5.2 follows from Lemma 5.1.

Chapter 6

Concluding Remarks and Future Work

In this dissertation, we study how to improve the performance of real-time monitoring applications by considering both the data signal value and the data freshness. To that end, we have designed low-complexity sampling and scheduling algorithms. Numerical evidences illustrate that our proposed algorithms achieve high performance gain over existing age-based scheduling policies. The following research problems could be an interesting directions based on our findings.

Joint Optimal Scheduling and Control. With the emerging number of interconnected devices, an efficient design of jointly optimal scheduling and control policies has significant importance where communication delays and uncertainties significantly impact control performance. This is relevant to systems that have safety requirements. My goal is to develop joint optimal scheduling and control policies that seamlessly optimize communication and control actions to improve situational awareness. In contrast with our existing studies explained in Chapter 4, we can consider the state of the scheduler and controller to make the system more practical. Leveraging insights from optimization theory, we strive to design algorithms that dynamically adapt scheduling decisions based on the real-time requirements of the control system. By tightly coupling the scheduling and control aspects, the system performance can be improved. This research not only contributes to the theoretical foundations of networked control systems but also holds practical implications for diverse applications, such as industrial automation to autonomous vehicles, remote surgery to health monitoring, where the interconnection of wireless communications and control policies is crucial for achieving optimal system performance.

Differential Privacy for Enhancing Security in Remote Estimation. Differential privacy offers a rigorous framework for quantifying and mitigating the risks associated with the disclosure of sensitive information. We aim to incorporate this privacy analysis with our earlier sampling and scheduling based studies in Chapter 2 and 3. By integrating differential privacy mechanisms into the data transmission strategies, we aim to establish a policy that jointly optimize the privacy protection and the estimation error. This approach ensures not only the optimization of mean square error and data security but also addresses concerns related to privacy-preserving data sharing. Practical examples of this framework can be found in scenarios where remote monitoring of sensitive information, such as financial transactions or personal health records, requires a balance between accuracy, security, and individual privacy. Through this approach, our research focuses on contributing comprehensive solutions that navigate the complex landscape of wireless communication by simultaneously improving system performance and data privacy.

Demonstration of Designed Policy. Our goal is to implement and validate the research projects by using the designed control and transmission scheduling policy utilizing robotics as a practical testbed. With a primary focus on safety-critical systems as illustrated in Chapter 4, the incorporation of advanced control policies and transmission scheduling strategies is essential for ensuring the reliability of these systems. Practical examples include autonomous vehicles, where the coordination of control actions and transmission schedules is crucial for timely and secure communication between vehicles and infrastructure. Additionally, in manufacturing environments with collaborative robots, the synchronization of control and communication plays a pivotal role in maintaining safety protocols. Through this research, I aim to contribute to the evidence of their effectiveness in enhancing the safety and performance of critical systems.

Bibliography

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

[2] X. Song and J. W. S. Liu, "Performance of multiversion concurrency control algorithms in maintaining temporal consistency," in *Proceedings., Fourteenth Annual International Computer Software and Applications Conference*, Oct 1990, pp. 132–139.

[3] E. Altman, R. El-Azouzi, D. S. Menasche, and Y. Xu, "Forever young: Aging control for hybrid networks," in *ACM MobiHoc*, 2019.

[4] Y. Sun and B. Cyr, "Sampling for data freshness optimization: Non-linear age functions," *J. Commun. Netw.*, vol. 21, no. 3, pp. 204–219, 2019.

[5] ——, "Information aging through queues: A mutual information perspective," in *IEEE SPAWC Workshop*, 2018.

[6] A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, "Age-optimal sampling and transmission scheduling in multi-source systems," in *ACM MobiHoc*, 2019.

[7] C. Kam, S. Kompella, G. D. Nguyen, and A. Ephremides, "Effect of message transmission path diversity on status age," *IEEE Trans. Inf. Theory*, vol. 62, no. 3, pp. 1360–1374, Mar. 2016.

[8] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7492–7508, Nov. 2017.

[9] C. Kam, S. Kompella, G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, "On the age of information with packet deadlines," *IEEE Trans. Inf. Theory*, vol. 64, no. 9, pp. 6419–6428, Sept. 2018.

[10] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Trans. Inf. Theory*, vol. 65, no. 3, pp. 1807–1827, Mar. 2019.

[11] Q. He, D. Yuan, and A. Ephremides, "Optimal link scheduling for age minimization in wireless systems," *IEEE Trans. Inf. Theory*, vol. 64, no. 7, pp. 5381–5394, July 2018.

[12] C. Joo and A. Eryilmaz, "Wireless scheduling for information freshness and synchrony: Drift-based design and heavy-traffic analysis," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2556–2568, Dec 2018.

[13] A. M. Bedewy, Y. Sun, and N. B. Shroff, "Minimizing the age of the information through queues," *IEEE Trans. Inf. Theory*, vol. 65, no. 8, pp. 5215–5232, Aug 2019.

[14] ——, "The age of information in multihop networks," *IEEE/ACM Trans. Netw.*, vol. 27, no. 3, pp. 1248–1257, June 2019.

[15] Y. Sun, E. Uysal-Biyikoglu, and S. Kompella, "Age-optimal updates of multiple information flows," in *IEEE INFOCOM AoI Workshop*, 2018.

[16] I. Kadota, A. Sinha, and E. Modiano, "Scheduling algorithms for optimizing age of information in wireless networks with throughput constraints," *IEEE/ACM Trans. Netw.*, vol. 27, no. 4, pp. 1359–1372, 2019.

[17] R. Talak, S. Karaman, and E. Modiano, "Optimizing information freshness in wireless networks under general interference constraints," in *ACM MobiHoc*, 2018.

[18] N. Lu, B. Ji, and B. Li, "Age-based scheduling: Improving data freshness for wireless real-time traffic," in *ACM MobiHoc*, 2018.

[19] A. Maatouk, Y. Sun, A. Ephremides, and M. Assaad, "Status updates with priorities: Lexicographic optimality," in *IEEE/IFIP WiOpt*, 2020.

[20] B. Zhou and W. Saad, "Joint status sampling and updating for minimizing age of information in the Internet of things," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7468–7482, Nov 2019.

[21] ——, "Minimum age of information in the Internet of things with non-uniform status packet sizes," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1933–1947, 2020.

[22] A. Kosta, N. Pappas, and V. Angelakis, *Age of Information: A New Concept, Metric, and Tool.* Now Publishers Inc, 2018.

[23] Y. Sun, I. Kadota, R. Talak, and E. Modiano, *Age of Information: A New Metric for Information Freshness.* Morgan & Claypool, 2019.

[24] A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, "Optimal sampling and scheduling for timely status updates in multi-source networks," *IEEE Trans. Inf. Theory*, vol. 67, no. 6, pp. 4019–4034, 2021.

[25] Y. Xiao and Y. Sun, "A dynamic jamming game for real-time status updates," in *IEEE INFOCOM AoI Workshop*, April 2018, pp. 354–360.

[26] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, 2021.

[27] G. M. Lipsa and N. C. Martins, "Remote state estimation with communication costs for first-order LTI systems," *IEEE Trans. Auto. Control*, vol. 56, no. 9, pp. 2013–2025, Sept. 2011.

[28] B. Hajek, K. Mitzel, and S. Yang, "Paging and registration in cellular networks: Jointly optimal policies and an iterative algorithm," *IEEE Trans. Inf. Theory*, vol. 54, no. 2, pp. 608–622, Feb 2008.

[29] A. Nayyar, T. Başar, D. Teneketzis, and V. V. Veeravalli, "Optimal strategies for communication and remote estimation with an energy harvesting sensor," *IEEE Trans. Auto. Control*, vol. 58, no. 9, pp. 2246–2260, Sept. 2013.

[30] M. Rabi, G. V. Moustakides, and J. S. Baras, "Adaptive sampling for linear state estimation," *SIAM Journal on Control and Optimization*, vol. 50, no. 2, pp. 672–702, 2012.

[31] K. Nar and T. Başar, "Sampling multidimensional Wiener processes," in *IEEE CDC*, Dec. 2014, pp. 3426–3431.

[32] X. Gao, E. Akyol, and T. Başar, "Optimal communication scheduling and remote estimation over an additive noise channel," *Automatica*, vol. 88, pp. 57 – 69, 2018.

[33] J. Chakravorty and A. Mahajan, "Remote estimation over a packet-drop channel with Markovian state," *IEEE Trans. Auto. Control*, vol. 65, no. 5, pp. 2016–2031, 2020.

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

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

[36] C.-H. Tsai and C.-C. Wang, "Unifying AoI minimization and remote estimation: Optimal sensor/controller coordination with random two-way delay," in *IEEE INFOCOM*, 2020.

[37] N. Guo and V. Kostina, "Optimal causal rate-constrained sampling for a class of continuous Markov processes," in *IEEE ISIT*, 2020.

[38] A. Arafa, K. Banawan, K. G. Seddik, and H. Vincent Poor, "Timely estimation using coded quantized samples," in *IEEE ISIT*, 2020, pp. 1812–1817.

[39] V. Jog, R. J. La, and N. C. Martins, "Channels, learning, queueing and remote estimation systems with a utilization-dependent component," 2019, coRR, abs/1905.04362.

[40] A. M. Bedewy, Y. Sun, R. Singh, and N. B. Shroff, "Optimizing information freshness using low-power status updates via sleep-wake scheduling," in *ACM MobiHoc*, 2020, pp. 51–60.

[41] V. Tripathi and E. Modiano, "A Whittle index approach to minimizing functions of age of information," in *IEEE Allerton*, 2019, pp. 1160–1167.

[42] Y.-P. Hsu, "Age of information: Whittle index for scheduling stochastic arrivals," in *IEEE ISIT*, 2018, pp. 2634–2638.

[43] Z. Tang, Z. Sun, N. Yang, and X. Zhou, "Whittle index-based scheduling policy for minimizing the cost of age of information," *IEEE Communications Letters*, vol. 26, no. 1, pp. 54–58, 2021.

[44] B. Yin, S. Zhang, Y. Cheng, L. X. Cai, Z. Jiang, S. Zhou, and Z. Niu, "Only those requested count: Proactive scheduling policies for minimizing effective age-of-information," in *IEEE INFOCOM*, 2019, pp. 109–117.

[45] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling policies for minimizing age of information in broadcast wireless networks," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2637–2650, 2018.

[46] M. K. C. Shisher and Y. Sun, "How does data freshness affect real-time supervised learning?" in *ACM MobiHoc*, 2022, pp. 31–40.

[47] G. Chen, S. C. Liew, and Y. Shao, "Uncertainty-of-information scheduling: A restless multiarmed bandit framework," *IEEE Trans. Inf. Theory*, vol. 68, no. 9, pp. 6151–6173, 2022.

[48] S. H. A. Ahmad, M. Liu, T. Javidi, Q. Zhao, and B. Krishnamachari, "Optimality of myopic sensing in multichannel opportunistic access," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4040–4050, 2009.

[49] J. Wang, X. Ren, Y. Mo, and L. Shi, "Whittle index policy for dynamic multichannel allocation in remote state estimation," *IEEE Trans. Auto. Control*, vol. 65, no. 2, pp. 591–603, 2020.

[50] G. Xiong, X. Qin, B. Li, R. Singh, and J. Li, "Index-aware reinforcement learning for adaptive video streaming at the wireless edge," in *ACM MobiHoc*, 2022, pp. 81–90.

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

[52] Y. Chen and A. Ephremides, "Scheduling to minimize age of incorrect information with imperfect channel state information," *Entropy*, vol. 23, no. 12, p. 1572, 2021.

[53] G. Chen and S. C. Liew, "An index policy for minimizing the uncertainty-of-information of markov sources," *IEEE Trans. Inf. Theory*, vol. 70, no. 1, pp. 698–721, 2024.

[54] M. K. C. Shisher, B. Ji, I.-H. Hou, and Y. Sun, "Learning and communications co-design for remote inference systems: Feature length selection and transmission scheduling," *IEEE J. Sel. Areas Inf. Theory*, vol. 4, pp. 524–538, 2023.

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

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

[57] J. Zhong, R. D. Yates, and E. Soljanin, "Two freshness metrics for local cache refresh," in *IEEE ISIT*, 2018, pp. 1924–1928.

[58] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides, "The age of incorrect information: A new performance metric for status updates," *IEEE/ACM Trans. Netw.*, vol. 28, p. 2215–2228, oct 2020.

[59] X. Zheng, S. Zhou, and Z. Niu, "Urgency of information for context-aware timely status updates in remote control systems," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 11, pp. 7237–7250, 2020.

[60] J. Holm, A. E. Kalør, F. Chiariotti, B. Soret, S. K. Jensen, T. B. Pedersen, and P. Popovski, "Freshness on demand: Optimizing age of information for the query process," in *IEEE ICC*, 2021, pp. 1–6.

[61] N. Pappas and M. Kountouris, "Goal-oriented communication for real-time tracking in autonomous systems," in *IEEE ICAS*, 2021, pp. 1–5.

[62] M. E. Ildiz, O. T. Yavascan, E. Uysal, and O. T. Kartal, "Query age of information: Optimizing AoI at the right time," in *IEEE ISIT*, 2022, pp. 144–149.

[63] ——, "Pull or wait: How to optimize query age of information," *IEEE J. Sel. Areas Inf. Theory*, pp. 1–1, 2023.

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

[65] Z. Wang, M.-A. Badiu, and J. P. Coon, "A framework for characterizing the value of information in hidden Markov models," *IEEE Trans. Inf. Theory*, vol. 68, no. 8, pp. 5203–5216, 2022.

[66] T. Soleymani, S. Hirche, and J. S. Baras, "Optimal self-driven sampling for estimation based on value of information," in *IEEE WODES*, 2016, pp. 183–188.

[67] M. K. C. Shisher, H. Qin, L. Yang, F. Yan, and Y. Sun, "The age of correlated features in supervised learning based forecasting," in *IEEE INFOCOM Workshops*, 2021, pp. 1–8.

[68] G. E. Uhlenbeck and L. S. Ornstein, "On the theory of the Brownian motion," *Phys. Rev.*, vol. 36, pp. 823–841, Sept. 1930.

[69] J. L. Doob, "The Brownian movement and stochastic equations," *Annals of Mathematics*, vol. 43, no. 2, pp. 351–369, 1942.

[70] L. Evans, S. Keef, and J. Okunev, "Modelling real interest rates," *Journal of Banking and Finance*, vol. 18, no. 1, pp. 153 – 165, 1994.

[71] A. Cika, M. Badiu, and J. Coon, "Quantifying link stability in ad hoc wireless networks subject to Ornstein-Uhlenbeck mobility," in *IEEE ICC*, 2019.

[72] H. Kim, J. Park, M. Bennis, and S. Kim, "Massive UAV-to-ground communication and its stable movement control: A mean-field approach," in *IEEE SPAWC*, June 2018, pp. 1–5.

[73] E. Vinogradov, H. Sallouha, S. D. Bast, M. M. Azari, and S. Pollin, "Tutorial on UAVs: A blue sky view onwireless communication," *Journal of Mobile Multimedia*, vol. 14, no. 4, p. 395–468, 2018.

[74] G. Peskir and A. N. Shiryaev, *Optimal Stopping and Free-Boundary Problems.* Basel, Switzerland: Birkhäuswer Verlag, 2006.

[75] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, 5th ed. Springer-Verlag Berlin Heidelberg, 2000.

[76] R. Durrett, *Probability: Theory and Examples*, 4th ed. Cambridge Univerisity Press, 2010.

[77] P. J. Haas, *Stochastic Petri Nets: Modelling, Stability, Simulation.* New York, NY: Springer New York, 2002.

[78] S. M. Ross, *Applied Probability Models with Optimization Applications.* San Francisco, CA: Holden-Day, 1970.

[79] H. Mine and S. Osaki, *Markovian Decision Processes.* New York: Elsevier, 1970.

[80] D. Hayman and M. Sobel, *Stochastic models in Operations Research, Volume II: Stochastic Optimizations.* New York: McGraw-Hill, 1984.

[81] E. A. Feinberg, "Constrained semi-Markov decision processes with average rewards," *Zeitschrift für Operations Research*, vol. 39, no. 3, pp. 257–288, 1994.

[82] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Belmont, MA: Athena Scientific, 2005, vol. 1.

[83] R. A. Maller, G. Müller, and A. Szimayer, "Ornstein-Uhlenbeck processes and extensions," in *Handbook of Financial Time Series*, T. Mikosch, J.-P. Kreiß, R. A. Davis, and T. G. Andersen, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 421–437.

[84] T. Soleymani, S. Hirche, and J. S. Baras, "Optimal information control in cyber-physical systems," *IFAC-PapersOnLine*, vol. 49, no. 22, pp. 1 – 6, 2016.

[85] I. Gradshteyn and I. Ryzhik, "Table of integrals, series, and products (7th ed.)," *Academic Press*, 2007.

[86] F. W. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, "Nist handbook of mathematical functions hardback and cd-rom," *Cambridge University Press*, 2010.

[87] J. H. Mathews and K. K. Fink, *Numerical Methods Using Matlab.* Simon & Schuster, Inc., 1998.

[88] C.-H. Tsai and C.-C. Wang, "Age-of-information revisited: Two-way delay and distribution-oblivious online algorithm," in *IEEE ISIT*, 2020.

[89] M. Spivak, *Calculus*, 4th ed. Publish or Perish, 2008.

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

[91] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex Analysis and Optimization.* Belmont, MA: Athena Scientific, 2003.

[92] A. N. Borodin and P. Salminen, *Handbook of Brownian Motion – Facts and Formulae.* Basel, Switzerland: Birkhäuswer Verlag, 1996.

[93] S. M. Ross, *Stochastic Processes*, 2nd ed. John Wiley& Sons, Inc., 1996.

[94] H. Amann, *Ordinary Differential Equations An Introduction to Nonlinear Analysis.* Berlin: Walter De Gruyter, 1990.

[95] G. Strang, *Calculus.* Wellesley-Cambridge Press, 1991.

[96] T. M. Liggett, *Continuous Time Markov Processes: An Introduction.* Providence, Rhode Island: American Mathematical Society, 2010.

[97] C. Jia and G. Zhao, "Moderate maximal inequalities for the Ornstein-Uhlenbeck process," *Proc. Amer. Math. Soc.*, vol. 148, pp. 3607–3615, 2020.

[98] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York, NY, USA: Springer-Verlag New York, Inc., 1994.

[99] A. Jeffrey and H.-H. Dai, *Handbook of Mathematical Formulas and Integrals.* Academic Press, 1995.

[100] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. Academic Press, 2007.

[101] P. Morters and Y. Peres, *Brownian Motion.* Cambridge Univerisity Press, 2010.

[102] E. Altman, R. El-Azouzi, D. S. Menasche, and Y. Xu, "Forever young: Aging control for hybrid networks," in *ACM MobiHoc*, 2019, p. 91–100.

[103] O. E. Barndorff-Nielsen and N. Shephard, *Modelling by Levy processes for financial econometrics.* Birkhauser, 2001, pp. 283–318.

[104] K. R. Ghusinga, V. Srivastava, and A. Singh, "Driving an ornstein-uhlenbeck process to desired first-passage time statistics," in *18th European Control Conference (ECC)*, 2019, pp. 869–874.

[105] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, “Remote estimation of the Wiener process over a channel with random delay,” in *IEEE ISIT*, 2017, pp. 321–325.

[106] J. Nauta, Y. Khaluf, and P. Simoens, “Using the Ornstein-Uhlenbeck process for random exploration,” in *COMPLEXIS2019*. SciTePress, 2019, pp. 59–66.

[107] Y. Nie and V. Linetsky, “Sticky reflecting Ornstein-Uhlenbeck diffusions and the Vasicek interest rate model with the sticky zero lower bound,” *Stochastic Models*, vol. 36, no. 1, pp. 1–19, 2020.

[108] K. Bartoszek, S. Glémin, I. Kaj, and M. Lascoux, “Using the Ornstein-Uhlenbeck process to model the evolution of interacting populations,” *Journal of theoretical biology*, vol. 429, pp. 35–45, 2017.

[109] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, “Continuous control with deep reinforcement learning,” *arXiv*, 2015.

[110] T. Z. Ornee and Y. Sun, “Performance bounds for sampling and remote estimation of Gauss-Markov processes over a noisy channel with random delay,” in *IEEE SPAWC*, 2021, pp. 1–5.

[111] R. D. Yates, “Lazy is timely: Status updates by an energy harvesting source,” in *IEEE ISIT*. IEEE, 2015, pp. 3008–3012.

[112] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, “Update or wait: How to keep your data fresh,” *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.

[113] R. R. Weber and G. Weiss, “On an index policy for restless bandits,” *Journal of applied probability*, vol. 27, no. 3, pp. 637–648, 1990.

[114] P. Whittle, “Restless bandits: activity allocation in a changing world,” *Journal of Applied Probability*, vol. 25A, pp. 287–298, 1988.

[115] I. M. Verloop, “Asymptotically optimal priority policies for indexable and nonindexable restless bandits,” *The Annals of Applied Probability*, vol. 26, no. 4, 2016.

[116] G. Strang, “Calculus,” *SIAM*, 1991.

[117] A. Grau, M. Indri, L. L. Bello, and T. Sauter, “Industrial robotics in factory automation: From the early stage to the internet of things,” in *IEEE IECON*, 2017, pp. 6159–6164.

[118] S. Abdulmalek, A. Nasir, W. A. Jabbar, M. A. Almuhaya, A. K. Bairagi, M. A.-M. Khan, and S.-H. Kee, “Iot-based healthcare-monitoring system towards improving quality of life: A review,” in *Healthcare*, vol. 10, no. 10, 2022, p. 1993.

[119] M. Seenivasan, M. Arularasu, K. Senthilkumar, and R. Thirumalai, “Disaster prevention and control management in automation: a key role in safety engineering,” *Procedia Earth and Planetary Science*, vol. 11, pp. 557–565, 2015.

[120] M. K. C. Shisher and Y. Sun, "How does data freshness affect real-time supervised learning?" in *ACM MobiHoc*, 2022, pp. 31–40.

[121] G. Chen, S. C. Liew, and Y. Shao, "Uncertainty-of-information scheduling: A restless multi-armed bandit framework," *IEEE Trans. Inf. Theory*, 2022.

[122] A. P. Dawid, "Coherent measures of discrepancy, uncertainty and dependence, with applications to bayesian predictive experimental design," *Department of Statistical Science, University College London*, vol. 139, 1998.

[123] F. Farnia and D. Tse, "A minimax approach to supervised learning," *Advances in Neural Information Processing Systems*, vol. 29, 2016.

[124] W. Ouyang, A. Eryilmaz, and N. B. Shroff, "Asymptotically optimal downlink scheduling over markovian fading channels," in *IEEE INFOCOM*, 2012, pp. 1224–1232.

[125] K. Liu and Q. Zhao, "Indexability of restless bandit problems and optimality of whittle index for dynamic multichannel access," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5547–5567, 2010.

[126] D. P. Bertsekas *et al.*, "Dynamic programming and optimal control 3rd edition, volume ii," *Belmont, MA: Athena Scientific*, vol. 1, 2011.

[127] P. Whittle, "Restless bandits: activity allocation in a changing world," *Journal of Applied Probability*, vol. 25A, pp. 287–298, 1988.

[128] C. Papadimitriou and J. Tsitsiklis, "The complexity of optimal queueing network control," in *IEEE CCC*, 1994, pp. 318–322.

[129] T. Z. Ornee and Y. Sun, "A Whittle index policy for the remote estimation of multiple continuous Gauss-Markov processes over parallel channels," *arXiv preprint arXiv:2305.04809*, accepted by ACM MobiHoc 2023.

[130] A. Nedic and A. Ozdaglar, "Subgradient methods in network resource allocation: Rate analysis," in *IEEE Annual Conference on Information Sciences and Systems*, 2008, pp. 1189–1194.

[131] B. K. Muirhead, "Mars rovers, past and future," in *IEEE aerospace conference*, vol. 1, 2004.

Vita

B.Sc. in Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology 2017

M.S. in Electrical Engineering, Auburn University 2022

Graduate Research and Teaching Assistant in Electrical Engineering, Auburn University. 2018-2024

Publications

1. **Tasmeen Zaman Ornee** and Yin Sun, "Sampling for Remote Estimation through Queues: Age of information and Beyond," in *Proceedings of IEEE International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, 2019, pp. 1-8. (Best Paper Award)

2. **Tasmeen Zaman Ornee** and Yin Sun, "Sampling and Remote Estimation for the Ornstein-Uhlenbeck Process through Queues: Age of information and Beyond," in *IEEE/ACM Transactions on Networking*, 2021, vol. 29, no. 5, pp. 1962-1975. (Recommended for fast-tracked review)

3. **Tasmeen Zaman Ornee** and Yin Sun, "Performance Bounds for Sampling and Remote Estimation of Gauss-Markov processes over A Noisy Channel with Random Delay," in *Proceedings of IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2021, pp. 1-5. (Invited paper)

4. **Tasmeen Zaman Ornee** and Yin Sun, "Age of Information and Remote Estimation," in *Age of Information: Foundations and Applications*, B. Zhou, W. Saad, H. S. Dhillon, N. Pappas, and M. Abd-Elmagid eds. *Cambridge University Press (CUP)*, 2023.

5. **Tasmeen Zaman Ornee** and Yin Sun, "A Whittle Index Policy for the Remote Estimation of Multiple Continuous Gauss-Markov Processes over Parallel Channels," in *Proceedings of the Twenty-fourth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing (MobiHoc '23)*, 2023, pp. 91–100.

6. **Tasmeen Zaman Ornee** and Yin Sun, "Remote Estimation of Gauss-Markov Processes over Multiple Channels: A Whittle Index Policy," submitted to *IEEE/ACM Transactions on Networking*, 2023.

7. **Tasmeen Zaman Ornee**, Md Kamran Chowdhury Shisher, Clement Kam, and Yin Sun, "Context-aware Status Updating: Wireless Scheduling for Maximizing Situational

Awareness in Safety-critical Systems," in *Proceedings of IEEE Military Communications Conference (MILCOM), Workshop on QuAVoI*, 2023, pp. 194-200.