A Whittle Index Policy for the Remote Estimation of Multiple Continuous Gauss-Markov Processes over Parallel Channels

Tasmeen Zaman Ornee
Auburn University
Auburn, AL, USA
tzo0017@auburn.edu

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

ABSTRACT

In this paper, we study a sampling and transmission scheduling problem for multi-source remote estimation, where a scheduler determines when to take samples from multiple continuous-time Gauss-Markov processes and send the samples over multiple channels to remote estimators. The sample transmission times are *i.i.d.* across samples and 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) problem with a continuous state space. We prove that the 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 Age of Information (AoI) minimization over parallel channels with *i.i.d.* random transmission times. Our results unite two theoretical frameworks 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. Moreover, the methodology employed to derive threshold-based sampling strategies in the single-source, single-channel scenario plays a crucial role in establishing indexability and evaluating the Whittle index in the more intricate multi-source, multi-channel scenario. Our numerical results show that the proposed policy achieves high performance gain over the existing policies when some of the Gauss-Markov processes are highly unstable.

CCS CONCEPTS

• **Networks → Network performance evaluation**; • **Theory of computation → Design and analysis of algorithms**;

MobiHoc '23, October 23–26, 2023, Washington, DC, USA

ACM ISBN 978-1-4503-9926-5/23/10…$15.00
https://doi.org/10.1145/3565287.3610263

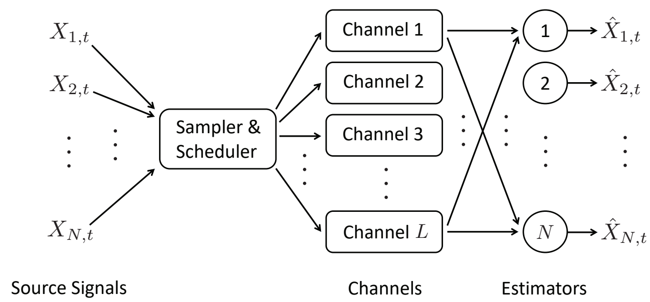

Figure 1: A multi-source, multi-channel remote estimation system.

KEYWORDS

Ornstein-Uhlenbeck process, Wiener process, remote estimation, Whittle index, restless multi-armed bandit

ACM Reference Format:
Tasmeen Zaman Ornee and Yin Sun. 2023. A Whittle Index Policy for the Remote Estimation of Multiple Continuous Gauss-Markov Processes over Parallel Channels. In *The Twenty-fourth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing (MobiHoc '23), October 23–26, 2023, Washington, DC, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3565287.3610263

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, one metric named *Age of Information* (AoI) has drawn significant attention in recent years, e.g., [14], [34]. 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 [38], [39], [16], [35] and illustrated to be useful as a metric of information freshness in sampling, estimation, and control [49], [35].

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 in a problem of sampling a Wiener process [37]. Subsequently, it was generalized to the case of (stable) Ornstein-Uhlenbeck (OU) process in [28].

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 [14]. In this paper, we consider a remote estimation system with N source-estimator pairs and L channels, as illustrated in Figure 1. Each source n is a continuous-time Gauss-Markov process $X_{n,t}$, defined as the solution of a Stochastic Differential Equation (SDE)

$$dX_{n,t} = \theta_n(\mu_n - X_{n,t})dt + \sigma_n dW_{n,t}, \tag{1}$$

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 [6]. If $\theta_n = 0$, then $X_{n,t} = \sigma_n W_{n,t}$ is a scaled Wiener process [20]. 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 [22], interest rates [23], commodity prices [7], robotic swarms [15], biological processes [2], control systems (e.g., the transfer of liquids or gases in and out of a tank) [18], state exploration in deep reinforcement learning [17], 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 1) and derive an exact expression for the Whittle index (Theorem 2 and Lemma 1). These results generalize prior studies on the remote estimation of a single Gauss-Markov process [27, 28, 36] 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 [30] for solving optimal stopping problems and Dynkin's formula [25] 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 4-5 and Lemma 3). Technically, these results carry forth and expand upon prior findings on Whittle index based AoI minimization [10, 12, 41] in the following manner: In [10, 12, 41], the transmission time remains constant, resulting in the optimality of the zero-wait sampling policy defined in [39, 48]. 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 [27, 28, 35, 36] and Whittle index-based scheduling [10, 12, 41]. 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 ([27, 28, 36]) and a Whittle index-based scheduling policy (see Theorems 3, 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.

2 RELATED WORK

Remote state estimation has received considerable attention in numerous studies, e.g., see [1, 21, 27, 28, 31, 36, 42, 45] and two recent surveys [11, 43]. Optimal sampling of one-dimensional and multi-dimensional Wiener processes with zero-delay, perfect channel was studied in [21, 31]. A dynamic programming method was used in [31] to find the optimal sampling policy of the stable OU processes numerically for the case of zero-delay, perfect channel. A connection between remote estimation and AoI minimization was first reported in [36], where optimal sampling strategies were obtained for the remote estimation of the Wiener process over a channel with *i.i.d.* random transmission times. This study was further generalized to the case of the stable OU process in [28], where the optimal sampling strategy was derived analytically. In

[42], the authors considered remote estimation of the Wiener process with random two-way delay. When the system state follows a binary ON-OFF Markov process, Whittle index scheduling policies for remote estimation were developed in [1]. Our study makes a contribution on the remote estimation of multiple Gauss-Markov processes (possibly with different distributions), by showing indexability and providing an analytical expression of the Whittle index.

Moreover, AoI-based scheduling for timely status updating has been studied extensively in, e.g., [3–5, 9, 10, 12, 13, 32, 40, 41, 50]. A detailed survey on AoI was presented in [49]. In [9], the authors showed that under inference constraints, the scheduling problem for minimizing the age in wireless networks is NP-hard. In [4], 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 [3]. AoI minimization in single-hop networks was considered in [13]. AoI-based scheduling with timely throughput constraints was considered in [12]. A Whittle index-based scheduling algorithm for minimizing AoI for stochastic arrivals was considered in [10]. In [40], [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 [5]. A Whittle index policy for signal-agnostic remote estimation was studied in [45] for minimizing increasing AoI functions. In [32], the authors proposed a Whittle index policy for minimizing non-monotonic AoI functions. In the present paper, we propose a Whittle index policy for AoI-based, signal-agnostic remote estimation for *i.i.d.* random transmission times.

3 MODEL AND FORMULATION

3.1 System Model

Consider a remote estimation system with N source-estimator pairs and L channels, which is shown in Figure 1. Each source n is a continuous-time Gauss-Markov process $X_{n,t}$, as defined in (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 $S_{n,i} \le S_{n,i+1}$. 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 $S_{n,i} \le G_{n,i}$, 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, $D_{n,i} \le G_{n,i+1}$. 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 $U_n(t) = \max_i\{S_{n,i} : D_{n,i} \le t, i = 1, 2, \ldots\}$ 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 [14, 34]

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

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

$$\Delta_n(t) = t - S_{n,i}, \text{if } t \in [D_{n,i}, D_{n,i+1}),\ i = 0, 1, \ldots. \quad (3)$$

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.

3.2 MMSE Estimator

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

$$X_{n,t} = \begin{cases} X_{n,S_{n,i}} e^{-\theta_n(t-S_{n,i})} + \mu_n\left[1 - e^{-\theta_n(t-S_{n,i})}\right] \\ + \frac{\sigma_n}{\sqrt{2\theta_n}} W_{n,1-e^{-2\theta_n(t-S_{n,i})}}, & \text{if } \theta_n > 0, \\ \sigma_n W_{n,t}, & \text{if } \theta_n = 0, \\ X_{n,S_{n,i}} e^{-\theta_n(t-S_{n,i})} + \mu_n\left[1 - e^{-\theta_n(t-S_{n,i})}\right] \\ + \frac{\sigma_n}{\sqrt{-2\theta_n}} W_{n,e^{-2\theta_n(t-S_{n,i})}-1}, & \text{if } \theta_n < 0, \end{cases} \quad (4)$$

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) for the stable OU process and the scaled Wiener process were provided in [19]. The third expression in (4) for the unstable OU process is proven in the technical report [29] of the present paper.

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) $M_{n,t} = \{(S_{n,i}, X_{n,S_{n,i}}, G_{n,i}, D_{n,i}) : D_{n,i} \le t, i = 1, 2, \ldots\}$, 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 $\max_i\{D_{n,i} : D_{n,i} \le t, i = 1, 2, \ldots\}$. Similar to [28, 31, 33, 36], 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 [27, 28]

$$\begin{aligned} \hat{X}_{n,t} &= \mathbb{E}[X_{n,t} | M_{n,t}] \\ &= \begin{cases} X_{n,S_{n,i}} e^{-\theta_n(t-S_{n,i})} + \mu_n\left[1 - e^{-\theta_n(t-S_{n,i})}\right], & \text{if } \theta_n \ne 0, \\ \sigma_n W_{n,S_{n,i}}, & \text{if } \theta_n = 0. \end{cases} \end{aligned} \quad (5)$$

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

$$\varepsilon_n(t) = X_{n,t} - \hat{X}_{n,t}. \quad (6)$$

By substituting (4) and (5) into (6), if $t \in [D_{n,i}, D_{n,i+1})$, then

$$\varepsilon_n(t) = \begin{cases} \frac{\sigma_n}{\sqrt{2\theta_n}} W_{n,1-e^{-2\theta_n(t-S_{n,i})}}, & \text{if } \theta_n > 0, \\ \sigma_n(W_{n,t} - W_{n,S_{n,i}}), & \text{if } \theta_n = 0, \\ \frac{\sigma_n}{\sqrt{-2\theta_n}} W_{n,e^{-2\theta_n(t-S_{n,i})}-1}, & \text{if } \theta_n < 0. \end{cases} \quad (7)$$

3.3 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

$$\inf_{\pi \in \Pi} \limsup_{T\to\infty} \sum_{n=1}^N w_n \mathbb{E}_\pi \left[\frac{1}{T} \int_0^T \varepsilon_n^2(t) dt \right] \tag{8}$$

$$\text{s.t.} \sum_{n=1}^N c_n(t) \le L,\ c_n(t) \in \{0, 1\}, n = 1, 2, \ldots, N, t \in [0, \infty), \tag{9}$$

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 our technical report [29], we prove that in the optimal policies to (8)-(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 MAIN RESULTS

4.1 Signal-aware Scheduling

Problem (8)-(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 [46, 47]. 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 [46]. In this section, we develop a Whittle index policy for solving problem (8)-(9) in three steps: (i) first, we relax the constraint (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 (8)-(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.

4.1.1 Relaxation and Lagrangian Dual 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 (9). Following [44, 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 (8)-(9) is equivalent to

$$\inf_{\pi \in \Pi} \limsup_{T\to\infty} \sum_{n=1}^N w_n \mathbb{E}_\pi \left[\frac{1}{T} \int_0^T \varepsilon_n^2(t) dt \right] \tag{10}$$

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

$$c_n(t) \in \{0, 1\}, n = 1, 2, \ldots, N, t \in [0, \infty), \tag{11}$$

which is an RMAB with an equality constraint.

Following the standard relaxation and Lagrangian dual decomposition procedure in the Whittle index theory [47], we relax the constraint (11) as

$$\limsup_{T\to\infty} \sum_{n=0}^N \mathbb{E}_\pi \left[\frac{1}{T} \int_0^T c_n(t) dt \right] = L. \tag{12}$$

The relaxed constraint (12) only needs to be satisfied on average, whereas (11) is required to hold at any time t. Then, the RMAB (10)-(11) is reformulated as

$$\inf_{\pi \in \Pi} \limsup_{T\to\infty} \sum_{n=1}^N w_n \mathbb{E}_\pi \left[\frac{1}{T} \int_0^T \varepsilon_n^2(t) dt \right] \tag{13}$$

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

$$c_0(t) \in \{0, 1, \ldots, L\}, c_n(t) \in \{0, 1\}, n=1, 2, \ldots, N, t \in [0, \infty). \tag{14}$$

Next, we take the Lagrangian dual decomposition of the relaxed problem (13)-(14), which produces the following problem with a dual variable $\lambda \in \mathbb{R}$, also known as the activation cost [47]:

$$\inf_{\pi \in \Pi} \limsup_{T\to\infty} \mathbb{E}_\pi \left[\frac{1}{T} \int_0^T \sum_{n=1}^N w_n \varepsilon_n^2(t) + \lambda \left(\sum_{n=0}^N c_n(t) - L \right) dt \right]. \tag{15}$$

The term $\frac{1}{T} \int_0^T \sum_{n=0}^N \lambda L dt$ in (15) does not depend on policy π and hence can be removed. Then, Problem (15) can be decomposed into $(N + 1)$ separated sub-problems. The sub-problem associated with source n is

$$\bar{m}_{n,\text{opt}} = \inf_{\pi_n \in \Pi_n} \limsup_{T\to\infty} \mathbb{E}_{\pi_n} \left[\frac{1}{T} \int_0^T w_n \varepsilon_n^2(t) + \lambda c_n(t) dt \right], \tag{16}$$

where $\bar{m}_{n,\text{opt}}$ is the optimum value of (16) and $n = 1, 2, \ldots, N$. On the other hand, the sub-problem associated with the *dummy bandits* is given by

$$\inf_{\pi_0 \in \Pi_0} \limsup_{T\to\infty} \mathbb{E}_{\pi_n}\left[\frac{1}{T}\int_0^T \lambda c_0(t)dt\right], \tag{17}$$

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

4.1.2 Indexability. We now establish the indexability of the RMAB in (10)-(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 (16) (or (17) when $n = 0$) is to take a passive action at time t.

Definition 1. ***(Indexability).*** *[44] 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., $\lambda_1 \le \lambda_2$ implies $\Psi_n(\lambda_1) \subseteq \Psi_n(\lambda_2)$. The RMAB* (10)-(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 (16) is a continuous-time MDP with a continuous state space, determining the indexability of (16) appears to be quite formidable. In the sequel, we will utilize the techniques developed in our previous work [28] to solve (16) precisely and analytically characterize the set $\Psi_n(\lambda)$. This analysis will allow us to demonstrate that (16) is indeed indexable.

Define

$$G(x) = \frac{\sqrt{\pi}}{2}\frac{e^{x^2}}{x}\mathrm{erf}(x), \tag{18}$$

$$K(x) = \frac{\sqrt{\pi}}{2}\frac{e^{-x^2}}{x}\mathrm{erfi}(x), \tag{19}$$

where erf(x) and erfi(x) are the error function and imaginary error function, respectively, determined by [8, Sec. 8.25]

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt, \tag{20}$$

$$\mathrm{erfi}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{t^2}dt. \tag{21}$$

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 $G(0) = 1$ [28]. On the other hand, $K(x)$ is strictly decreasing on $x \in [0,\infty)$ and $K(0) = 1$ [27]. 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 [27]

$$K(x) = G(jx), \tag{22}$$

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

Proposition 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* (16)*, where*

$$S_{n,i+1}(\beta_n) = \inf_t \left\{t \ge D_{n,i}(\beta_n) : |\varepsilon_n(t)| \ge v_n(\beta_n)\right\}, \tag{23}$$

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

$$v_n(\beta_n) = \begin{cases} \frac{\sigma_n}{\sqrt{\theta_n}} G^{-1}\left(\frac{w_n \frac{\sigma_n^2}{2\theta_n}\mathbb{E}[e^{-2\theta_n Y_{n,i}}]}{w_n \frac{\sigma_n^2}{2\theta_n} - \beta_n}\right), & \text{if } \theta_n > 0,\\ \frac{1}{\sqrt{w_n}}\sqrt{3(\beta_n - w_n\sigma_n^2\mathbb{E}[Y_{n,i}])}, & \text{if } \theta_n = 0,\\ \frac{\sigma_n}{\sqrt{-\theta_n}} K^{-1}\left(\frac{w_n \frac{\sigma_n^2}{2\theta_n}\mathbb{E}[e^{-2\theta_n Y_{n,i}}]}{w_n \frac{\sigma_n^2}{2\theta_n} - \beta_n}\right), & \text{if } \theta_n < 0, \end{cases} \tag{24}$$

$G^{-1}(\cdot)$ and $K^{-1}(\cdot)$ are the inverse functions of $G(x)$ in (18) *and $K(x)$ in* (19)*, respectively, defined in the region of $x \in [0,\infty)$, and β_n is the unique root of*

$$\mathbb{E}\left[\int_{D_{n,i}(\beta_n)}^{D_{n,i+1}(\beta_n)} w_n\varepsilon_n^2(t)dt\right] - \beta_n\mathbb{E}[D_{n,i+1}(\beta_n) - D_{n,i}(\beta_n)] + \lambda\mathbb{E}[Y_{n,i+1}] = 0. \tag{25}$$

The optimal objective value to (16) *is given by*

$$\bar{m}_{n,\mathrm{opt}} = \frac{\mathbb{E}\left[\int_{D_{n,i}(\beta_n)}^{D_{n,i+1}(\beta_n)} w_n\varepsilon_n^2(t)dt\right] + \lambda\mathbb{E}[Y_{n,i+1}]}{\mathbb{E}[D_{n,i+1}(\beta_n) - D_{n,i}(\beta_n)]}. \tag{26}$$

Furthermore, β_n is exactly the optimal objective value of (16)*, i.e., $\beta_n = \bar{m}_{n,\mathrm{opt}}$.*

The proof is provided in our technical report [29].

Proposition 1 complements earlier optimal sampling results for the remote estimation of the Wiener process (i.e., the case of $\theta_n = 0$ and $\lambda = 0$) [36] and stable OU process (i.e., $\theta_n > 0$ and $\lambda = 0$) [28], 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 1, we can analytically characterize the set $\Psi_n(\lambda)$. To that end, we first show that the threshold $v_n(\beta_n)$ in (23) is a function of the activation cost λ. For any given λ, β_n is the unique root of equation (25). Hence, β_n can be expressed as an implicit function $\beta_n(\lambda)$ of λ, defined by equation (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 (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 (23) for taking a new sample is not satisfied. By this, an analytical expression of set $\Psi_n(\lambda)$ is derived as

$$\Psi_n(\lambda) = \{(\varepsilon,\gamma) \in \mathbb{R}\times[0,\infty) : \gamma > 0 \text{ or } |\varepsilon| < v_n(\beta_n(\lambda))\}. \tag{27}$$

Using (27), we can prove the first key result of the present paper:

Theorem 1. *The RMAB problem* (10)-(11) *is indexable.*

Proof sketch. According to Proposition 1, for any λ, the optimal solution to (16) is a threshold policy. Using this, we can show that the unique root $\beta_n(\lambda)$ of (25) is a strictly increasing function of λ. In addition, $v_n(\beta_n)$ in (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 (27), if $\lambda_1 \le \lambda_2$, then $\Psi_n(\lambda_1) \subseteq \Psi_n(\lambda_2)$. For the *dummy bandits*, it is optimal in (17) to activate a bandit when $\lambda < 0$. Hence, *dummy bandits* are always indexable. The details are provided in our technical report [29]. □

4.1.3 Whittle Index Policy. Next, we introduce the definition of the Whittle index.

Definition 2. *[47] If bandit n is indexable, then the Whittle index $W_n(\varepsilon, \gamma)$ of bandit n at state (ε, γ) is defined by*

$$W_n(\varepsilon, \gamma) = \inf_{\lambda}\{\lambda \in \mathbb{R} : (\varepsilon, \gamma) \in \Psi_n(\lambda)\}, \tag{28}$$

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

Theorem 2. *The following assertions are true for the Whittle index $W_n(\varepsilon, \gamma)$ of problem (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*

$$W_n(\varepsilon, 0) = \frac{w_n}{\mathbb{E}[Y_{n,i}]}\Bigg\{\mathbb{E}[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)]\frac{\sigma_n^2}{2\theta_n}\Bigg(1 - \frac{\mathbb{E}[e^{-2\theta_n Y_{n,i}}]}{G(\frac{\sqrt{\theta_n}}{\sigma_n}\varepsilon)}\Bigg) - \mathbb{E}\Bigg[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds\Bigg]\Bigg\}, \tag{29}$$

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

$$W_n(\varepsilon, 0) = \frac{w_n}{\mathbb{E}[Y_{n,i}]}\Bigg\{\mathbb{E}[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)]\Bigg(\frac{\varepsilon^2}{3} + \sigma_n^2\mathbb{E}[Y_{n,i}]\Bigg) - \mathbb{E}\Bigg[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds\Bigg]\Bigg\}, \tag{30}$$

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

$$W_n(\varepsilon, 0) = \frac{w_n}{\mathbb{E}[Y_{n,i}]}\Bigg\{\mathbb{E}[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)]\frac{\sigma_n^2}{2\theta_n}\Bigg(1 - \frac{\mathbb{E}[e^{-2\theta_n Y_{n,i}}]}{K(\frac{\sqrt{-\theta_n}}{\sigma_n}\varepsilon)}\Bigg) - \mathbb{E}\Bigg[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds\Bigg]\Bigg\}, \tag{31}$$

where $G(\cdot)$ and $K(\cdot)$ are defined in (18) *and* (19)*, respectively.*

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

$$W_n(\varepsilon, \gamma) = -\infty. \tag{32}$$

Proof sketch. When $\gamma = 0$, by (27), (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

$$|\varepsilon| = v_n(\beta_n(\lambda)). \tag{33}$$

Hence, $W_n(\varepsilon, 0) = \beta_n^{-1}(v_n^{-1}(|\varepsilon|)$. By substituting (24) and (25) into (33) and using the fact that $G(\cdot)$ and $K(\cdot)$ are even functions, statement (a) in Theorem 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 (28), $W_n(\varepsilon, \lambda) = -\infty$. By this, statement (b) in Theorem 2 is proven. The details are provided in our technical report [29]. □

In Theorem 2, the delivery time $D_{n,i}(\varepsilon)$ is expressed as a function of ε for the following reason: in the optimal solution to (16), the sample delivery time is a function of the activation cost λ. If the activation cost λ in (16) is chosen as $\lambda = W_n(\varepsilon, \gamma)$, then the sample delivery time in the optimal solution to (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 (29)-(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 (29)-(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 1 provided below, which is obtained by using Dynkin's formula [24, 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

$$O_{n,t} = \begin{cases} \frac{\sigma_n}{\sqrt{2\theta_n}} W_{n,1-e^{-2\theta_n t}}, & \text{if } \theta_n > 0, \\ \sigma_n W_{n,t}, & \text{if } \theta_n = 0, \\ \frac{\sigma_n}{\sqrt{-2\theta_n}} W_{n,e^{-2\theta_n t}-1}, & \text{if } \theta_n < 0. \end{cases} \tag{34}$$

By comparing (7) with (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 (29), (30), and (31).

Lemma 1. *In Theorem 2, it holds that*

$$\mathbb{E}[D_{n,i+1}(\varepsilon) - D_{n,i}(\varepsilon)] = \mathbb{E}\big[R_{n,1}\big(\max\big\{|\varepsilon|, |O_{n,Y_{n,i}}|\big\}\big)\big], \tag{35}$$

$$\mathbb{E}\Bigg[\int_{D_{n,i}(\varepsilon)}^{D_{n,i+1}(\varepsilon)} \varepsilon_n^2(s)ds\Bigg] = \mathbb{E}\big[R_{n,2}\big(\max\big\{|\varepsilon|, |O_{n,Y_{n,i}}|\big\} + O_{n,Y_{n,i+1}}\big)\big] - \mathbb{E}\big[R_{n,2}\big(O_{n,Y_{n,i}}\big)\big], \tag{36}$$

where if $\theta_n \neq 0$, then

$$R_{n,1}(\varepsilon) = \frac{\varepsilon^2}{\sigma_n^2}\, {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta_n}{\sigma_n^2}\varepsilon^2\right), \tag{37}$$

$$R_{n,2}(\varepsilon) = -\frac{\varepsilon^2}{2\theta_n} + \frac{\varepsilon^2}{2\theta_n}\, {}_2F_2\left(1, 1; \frac{3}{2}, 2; \frac{\theta_n}{\sigma_n^2}\varepsilon^2\right); \tag{38}$$

if $\theta_n = 0$, then

$$R_{n,1}(\varepsilon) = \frac{\varepsilon^2}{\sigma_n^2}, \tag{39}$$

$$R_{n,2}(\varepsilon) = \frac{\varepsilon^4}{6\sigma_n^2}. \tag{40}$$

In (37) and (38), we have used the generalized hypergeometric function, which is defined by [26, Eq. 16.2.1]

$$ {}_pF_q(a_1, a_2, \cdots, a_p; b_1, b_2, \cdots b_q; z) = \sum_{n=0}^{\infty} \frac{(a_1)_n (a_2)_n \cdots (a_p)_n}{(b_1)_n (b_2)_n \cdots (b_p)_n} \frac{z^n}{n!}, \tag{41}$$

where

$$(a)_0 = 1, \tag{42}$$

$$(a)_n = a(a+1)(a+2)\cdots(a+n-1),\ n \geq 1. \tag{43}$$

Lemma 1 is more general than Lemma 1 in [28], because Lemma 1 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 [28] was only shown for $\theta_n > 0$. Moreover, (35)-(36) in Lemma 1 are neater than (22)-(23) in Lemma 1 of [28]. Due to space limitation the proof of Lemma 1 is relegated to our technical report [29].

The expectations in (35) and (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.

Algorithm 1 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) and (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 (6) and (29), (30), (31), (32), (35), and (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

Lemma 2. *The Whittle index of the dummy bandits is 0, i.e.,* $W_0(\varepsilon, \gamma) = 0$.

The proof of Lemma 2 and the Whittle index policy for the RMAB (10)-(11) is provided in our technical report [29]. This policy activates the L bandits with the highest Whittle index at any given time t. As stated in Lemma 2, each dummy bandit has a Whittle index of $W_0(\varepsilon_0(t), \gamma_0(t)) = 0$. Consequently, if a bandit n (for $n = 1, 2, \ldots, N$) 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 Whittle index scheduling policy for solving the original sampling and scheduling problem (8)-(9) is illustrated in Algorithm 1. Because RMAB (8)-(9) and the RMAB (10)-(11) are equivalent to each other, the Whittle index policy for RMAB (10)-(11) (provided in [29]) and the Whittle index policy in Algorithm 1 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 (13)-(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 (11) will be satisfied.

In Algorithm 1, 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. Algorithm 1 can be either used as an event-driven algorithm, or be executed on discretized time slots $t = 0, T_s, 2T_s, \ldots$. When T_s is sufficiently small, the performance degradation caused by time discretization can be omitted.

4.1.4 Unity of Whittle Index-based Scheduling and Threshold-based Sampling. 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 (8)-(9) reduces to

$$\bar{m}_{1,\text{opt}} = \inf_{\pi \in \Pi} \limsup_{T \to \infty} \mathbb{E}_\pi \left[\frac{1}{T} \int_0^T \varepsilon_1^2(t) dt \right]. \tag{44}$$

The single-source, single-channel sampling and scheduling problem (44) is a special case of Proposition 1 with $n = 1$ and $\lambda = 0$. A threshold-based optimal solution to (44) is provided by the following corollary of Proposition 1.

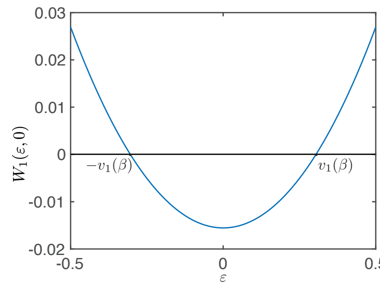

Figure 2: Illustration of the Whittle index $W_1(\varepsilon, \gamma)$ and the optimal threshold $v_1(\beta_1)$, where the parameters of the Gauss-Markov process are $\sigma_1 = 1$ and $\theta_1 = 0.1$ and the *i.i.d.* transmission times follow an exponential distribution with mean $\mathbb{E}[Y_{1,i}] = 2$.

Corollary 1. *If the $Y_{1,i}$'s are i.i.d. with $0 < \mathbb{E}[Y_{1,i}] < \infty$, then $(S_{1,1}(\beta_1), S_{1,2}(\beta_1), \ldots)$ with a parameter β_1 is an optimal solution to* (44)*, where*

$$S_{1,i+1}(\beta_1) = \inf_t \left\{ t \geq D_{1,i}(\beta_1) : |\varepsilon_1(t)| \geq v_1(\beta_1) \right\}, \tag{45}$$

$D_{1,i}(\beta_1) = S_{1,i}(\beta_1) + Y_{1,i}$, *$v_1(\beta_1)$ is defined by*

$$v_1(\beta_1) = \begin{cases} \frac{\sigma_1}{\sqrt{\theta_1}} G^{-1}\left(\frac{\frac{\sigma_1^2}{2\theta_1} \mathbb{E}[e^{-2\theta Y_{1,i}}]}{\frac{\sigma_1^2}{2\theta_1} - \beta_1} \right), & \text{if } \theta_1 > 0, \\ \sqrt{3(\beta_1 - \sigma_1^2 \mathbb{E}[Y_{1,i}])}, & \text{if } \theta_1 = 0, \\ \frac{\sigma_1}{\sqrt{-\theta_1}} K^{-1}\left(\frac{\frac{\sigma_1^2}{2\theta_1} \mathbb{E}[e^{-2\theta_1 Y_{1,i}}]}{\frac{\sigma_1^2}{2\theta_1} - \beta_1} \right), & \text{if } \theta_1 < 0, \end{cases} \tag{46}$$

$G^{-1}(\cdot)$ and $K^{-1}(\cdot)$ are the inverse functions of $G(x)$ in (18) *and $K(x)$ in* (19)*, respectively, for the region $x \in [0, \infty)$, and β_1 is the unique root of*

$$\mathbb{E}\left[\int_{D_{1,i}(\beta_1)}^{D_{1,i+1}(\beta_1)} \varepsilon_1^2(t) dt \right] - \beta_1 \mathbb{E}[D_{1,i+1}(\beta_1) - D_{1,i}(\beta_1)] = 0. \tag{47}$$

The optimal objective value to (44) *is given by*

$$\bar{m}_{1,\text{opt}} = \frac{\mathbb{E}\left[\int_{D_{1,i}(\beta_1)}^{D_{1,i+1}(\beta_1)} \varepsilon_1^2(t) dt \right]}{\mathbb{E}[D_{1,i+1}(\beta_1) - D_{1,i}(\beta_1)]}. \tag{48}$$

Furthermore, β_1 is exactly the optimal objective value of (44)*, i.e., $\beta_1 = \bar{m}_{1,\text{opt}}$.*

Corollary 1 follows directly from Proposition 1. For the cases of the Wiener process ($\theta_1 = 0$) and stable OU process ($\theta_1 > 0$), the threshold-based policy in Corollary 1 were earlier reported in [28]. The case of unstable OU process ($\theta_1 < 0$) is new.

It is important to note that the threshold-based policy in Corollary 1 and the Whittle index policy in the following theorem are equivalent.

Theorem 3. *If the $Y_{1,i}$'s are i.i.d. with $0 < \mathbb{E}[Y_{1,i}] < \infty$, then $(S_{1,1}(\varepsilon), S_{1,2}(\varepsilon), \ldots)$ with a parameter ε is an optimal solution to* (44)*, where*

$$S_{1,i+1}(\varepsilon) = \inf_t \left\{ t \geq S_{1,i}(\varepsilon) : W_1(\varepsilon_1(t), \gamma_1(t)) \geq 0 \right\}, \tag{49}$$

and $W_1(\varepsilon_n(t), \gamma_n(t))$ is the Whittle index of source 1, defined by (29), (30), (31)*, and* (32).

Proof sketch. Because (i) Corollary 1 provides an optimal solution to (44) and (ii) (49) is equivalent to the solution in Corollary 1, (49)

Algorithm 2 Whittle Index Policy for Signal-agnostic Remote Estimation

Initialize the set A of passive bandits $A = \{1, 2, \ldots, N\}$.
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 $n = 1, 2, \ldots, N$ using (3), (57), (59), (60), and (61).
  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 $\max_{n\in A} W_{n,\text{age}}(\Delta_n(t), \gamma_n(t)) \geq 0$ **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

is also an optimal solution to (44). The details are provided in our technical report [29]. □

Corollary 1 and Theorem 3 reveal a unification of threshold-based sampling and scheduling policy developed in [28] and the Whittle index policy developed in the present paper. 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 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 1 and [28] are applicable specifically to the single-source, single-channel scenario. Nevertheless, our exploration in Sections 4.1.1-4.1.3 illustrates the methodology for utilizing these findings to establish indexability and evaluate the Whittle index in the multi-source, multi-channel scenario.

4.2 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

$$\Pi_{\text{agnostic}} = \{\pi \in \Pi : \pi \text{ is independent of } \{X_{n,t}, t \geq 0\}_{n=1}^{N}\}. \quad (50)$$

In a signal-agnostic policy, the mean-squared estimation error of the process $X_{n,t}$ at time t is [36], [28]

$$\mathbb{E}[\varepsilon_n^2(t)] = p_n(\Delta_n(t)) = \begin{cases} \frac{\sigma_n^2}{2\theta_n}(1 - e^{-2\theta_n \Delta_n(t)}), & \text{if } \theta_n \neq 0, \\ \sigma_n^2 \Delta_n(t), & \text{if } \theta_n = 0, \end{cases} \quad (51)$$

where $\Delta_n(t)$ is the AoI and $p_n(\cdot)$ is an increasing function defined in (51). By using (51), for any policy $\pi \in \Pi_{\text{agnostic}}$

$$\mathbb{E}\left[\int_0^T \varepsilon_n^2(t)dt\right] = \mathbb{E}\left[\int_0^T p_n(\Delta_n(t))dt\right]. \quad (52)$$

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

$$\inf_{\pi \in \Pi_{\text{agnostic}}} \limsup_{T\to\infty} \sum_{n=1}^{N} w_n \mathbb{E}_\pi\left[\frac{1}{T}\int_0^T p_n(\Delta_n(t))dt\right] \quad (53)$$

$$\text{s.t.} \quad \sum_{n=1}^{N} c_n(t) \leq L, c_n(t) \in \{0, 1\}, t \in [0, \infty). \quad (54)$$

Problem (53)-(54) 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. By following the standard relaxation and Lagrangian dual decomposition procedure as explained in Section 4.1.1, we obtain the following sub-problem associated with bandit n:

$$\bar{m}_{n,\text{age-opt}} = \inf_{\pi_n \in \Pi_{n,\text{agnostic}}} \limsup_{T\to\infty} \mathbb{E}_{\pi_n}\left[\frac{1}{T}\int_0^T w_n p_n(\Delta_n(t)) + \lambda c_n(t)dt\right], \quad (55)$$

where $\bar{m}_{n,\text{age-opt}}$ is the optimum value of (55), $\pi_n = (S_{n,1}, S_{n,2}, \ldots)$ 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.

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 (55) is to take a passive action at time t.

Definition 3. ***(Indexability).*** *[44] 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* (53)-(54) *is said to be indexable if all* $(N+1)$ *bandits are indexable.*

An optimal solution to problem (55) is provided in of our technical report [29, Proposition 2], where we show that a parameter $\beta_{n,\text{age}}$ is equal to the optimum value of (55), i.e., $\beta_{n,\text{age}} = \bar{m}_{n,\text{age-opt}}$ and $\beta_{n,\text{age}}$ is a function of λ. By using the solution of (55), the set $\Psi_{n,\text{age}}(\lambda)$ in Definition 3 can be simplified as

$$\Psi_{n,\text{age}}(\lambda) = \{(\delta, \gamma) \in [0, \infty) \times [0, \infty) : \gamma > 0 \text{ or } \mathbb{E}[p_n(\delta + Y_{n,i+1})] < \beta_{n,\text{age}}(\lambda)\}. \quad (56)$$

Following the techniques developed in Section 4.1, we can obtain

Theorem 4. *If $p_n(\delta)$ is a strictly increasing function of δ, the RMAB problem* (53)-(54) *is indexable.*

Theorem 5. *In the RMAB problem* (53)-(54)*, 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

$$W_{n,age}(\delta, 0) = \frac{w_n}{\mathbb{E}[Y_{n,i}]}\Bigg\{\mathbb{E}[D_{n,i+1}(\delta) - D_{n,i}(\delta)]\mathbb{E}[p_n(\delta + Y_{n,i+1})] - \mathbb{E}\left[\int_{D_{n,i}(\delta)}^{D_{n,i+1}(\delta)} p_n(s)ds\right]\Bigg\}, \quad (57)$$

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

$$S_{n,i+1}(\delta) = D_{n,i}(\delta) + \max\{\delta - Y_{n,i}, 0\}. \quad (58)$$

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

$$W_{n,age}(\delta, \gamma) = -\infty. \quad (59)$$

The expectations in (57) can be easily evaluated using the following lemma:

Lemma 3. *In Theorem 5, it holds that*

$$\mathbb{E}[D_{n,i+1}(\delta) - D_{n,i}(\delta)] = \mathbb{E}[\max\{\delta, Y_{n,i}\}], \quad (60)$$

$$\mathbb{E}\left[\int_{D_{n,i}(\delta)}^{D_{n,i+1}(\delta)} p_n(s)ds\right] = \mathbb{E}[R_{n,3}(\max\{\delta, Y_{n,i}\} + Y_{n,i+1})] - \mathbb{E}[R_{n,3}(Y_{n,i})], \quad (61)$$

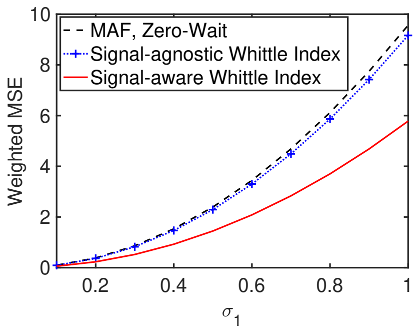

Figure 3: Total time-average MSE vs the parameter σ_1 of the Gauss-Markov source 1, where the number of sources is $N = 4$ and the number of channels is $L = 2$. The transmission times are *i.i.d*, following a normalized log-normal distribution with parameter $\rho = 1.5$, and $\mathbb{E}[Y_{n,i}] = 1$. The other parameters of the Gauss-Markov sources are $\sigma_2 = 0.8$, $\sigma_3 = 0.9$, $\sigma_4 = 1$, and $\theta_1 = -0.1$, $\theta_2 = \theta_3 = \theta_4 = 0.1$.

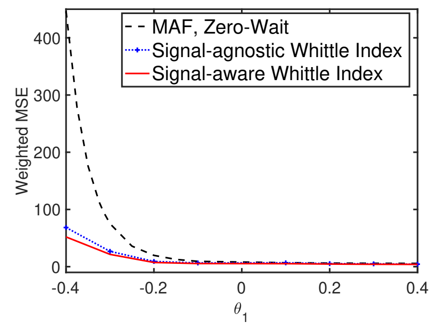

Figure 4: Total time-average MSE vs the parameter θ_1 of the Gauss-Markov source 1, where the number of sources is $N = 4$ and the number of channels is $L = 2$. The transmission times are *i.i.d.*, following a normalized log-normal distribution with parameter $\rho = 1.5$, and $\mathbb{E}[Y_{n,i}] = 1$. The other parameters for the Gauss-Markov sources are $\sigma_1 = \sigma_2 = \sigma_3 = \sigma_4 = 1$, and $\theta_2 = 0.2$, $\theta_3 = 0.3$, $\theta_4 = 0.1$.

where

$$R_{n,3}(\delta) = \int_0^\delta p_n(s)ds. \tag{62}$$

Theorems 4-5 and Lemma 3 hold for all increasing functions $p_n(\delta)$ of the AoI δ, not necessarily the mean-square estimation error function in (51). Due to space limitation, the proofs of Theorems 4-5 and Lemma 3 are relegated to our technical report [29].

The Whittle index scheduling policy for solving the sampling and scheduling problem (53)-(54) is illustrated in Algorithm 2.

Theorems 4-5, Lemma 3, and Algorithm 2 generalize prior studies on AoI-based Whittle index policies, e.g., [10, 12, 41]. More specifically, the Whittle index policies detailed in [10, 12, 41] were derived for the scenario of constant transmission times where the zero-wait sampling policy [39, 48] is an optimal solution for the sub-problem (55), 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 (55), resulting in the potential for both positive and negative values for the Whittle index derived in Theorem 5.

4.2.1 Unity of Whittle Index-based Scheduling and Threshold-based Sampling. For single-source, single-channel special case with $w_1 = 1$, problem (53)-(54) reduces to

$$\bar{m}_{1,\text{age-opt}} = \inf_{\pi \in \Pi_{\text{agnostic}}} \limsup_{T\to\infty} \mathbb{E}_\pi \left[\frac{1}{T}\int_0^T p_1(\Delta_1(t))dt\right] \tag{63}$$

Theorem 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 $(S_{1,1}(\delta), S_{1,2}(\delta), \ldots)$ with a parameter δ is an optimal solution to* (63)*, where*

$$S_{1,i+1}(\delta) = \inf_t \{t \geq S_{1,i}(\delta) : W_{1,age}(\Delta_1(t), \gamma_1(t)) \geq 0\}, \tag{64}$$

where $W_{1,age}(\Delta_1(t), \gamma_1(t))$ is the Whittle index of source 1, defined by (57) *and* (59).

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 [35, Theorem 1] and the Whittle index policy based in Theorem 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 [35, Theorem 1] at time t.

5 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 2.
- Signal-aware, Whittle Index policy: The policy that we proposed in Algorithm 1.

Figure 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 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 4 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 3. In Figure 4, 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.

6 CONCLUSION

In this paper, 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.

REFERENCES

[1] Sahand Haji Ali Ahmad, Mingyan Liu, Tara Javidi, Qing Zhao, and Bhaskar Krishnamachari. 2009. Optimality of myopic sensing in multichannel opportunistic access. *IEEE Trans. Inf. Theory* 55, 9 (2009), 4040–4050.

[2] Krzysztof Bartoszek, Sylvain Glémin, Ingemar Kaj, and Martin Lascoux. 2017. Using the Ornstein–Uhlenbeck process to model the evolution of interacting populations. *Journal of theoretical biology* 429 (2017), 35–45.

[3] Ahmed M Bedewy, Yin Sun, Sastry Kompella, and Ness B Shroff. 2021. Optimal sampling and scheduling for timely status updates in multi-source networks. *IEEE Trans. Inf. Theory* 67, 6 (2021), 4019–4034.

[4] Ahmed M Bedewy, Yin Sun, Rahul Singh, and Ness B Shroff. 2020. Optimizing information freshness using low-power status updates via sleep-wake scheduling. In *ACM MobiHoc*. 51–60.

[5] Gongpu Chen, Soung Chang Liew, and Yulin Shao. 2022. Uncertainty-of-Information Scheduling: A Restless Multiarmed Bandit Framework. *IEEE Trans. Inf. Theory* 68, 9 (2022), 6151–6173.

[6] J. L. Doob. 1942. The Brownian Movement and Stochastic Equations. *Annals of Mathematics* 43, 2 (1942), 351–369.

[7] L.T. Evans, S.P. Keef, and J. Okunev. 1994. Modelling real interest rates. *Journal of Banking and Finance* 18, 1 (1994), 153 – 165.

[8] I.S. Gradshteyn and I.M. Ryzhik. 2007. Table of Integrals, Series, and Products (7th ed.). *Academic Press* (2007).

[9] Qing He, Di Yuan, and Anthony Ephremides. 2017. Optimal link scheduling for age minimization in wireless systems. *IEEE Trans. Inf. Theory* 64, 7 (2017), 5381–5394.

[10] Yu-Pin Hsu. 2018. Age of information: Whittle index for scheduling stochastic arrivals. In *IEEE ISIT*. 2634–2638.

[11] Varun Jog, Richard J La, and Nuno C Martins. 2019. Channels, learning, queueing and remote estimation systems with a utilization-dependent component. *CoRR* (2019), 1–22.

[12] Igor Kadota, Abhishek Sinha, and Eytan Modiano. 2019. Scheduling algorithms for optimizing age of information in wireless networks with throughput constraints. *IEEE/ACM Trans. Netw.* 27, 4 (2019), 1359–1372.

[13] Igor Kadota, Abhishek Sinha, Elif Uysal-Biyikoglu, Rahul Singh, and Eytan Modiano. 2018. Scheduling policies for minimizing age of information in broadcast wireless networks. *IEEE/ACM Trans. Netw.* 26, 6 (2018), 2637–2650.

[14] S. Kaul, R. D. Yates, and M. Gruteser. 2012. Real-Time Status: How Often Should One Update?. In *IEEE INFOCOM*.

[15] H. Kim, J. Park, M. Bennis, and S. Kim. 2018. Massive UAV-to-Ground Communication and its Stable Movement Control: A Mean-Field Approach. In *IEEE SPAWC*. 1–5.

[16] Antzela Kosta, Nikolaos Pappas, Anthony Ephremides, and Vangelis Angelakis. 2017. Age and value of information: Non-linear age case. In *IEEE ISIT*. 326–330.

[17] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. 2015. Continuous control with deep reinforcement learning. *arXiv* (2015).

[18] G. M. Lipsa and N. C. Martins. 2011. Remote State Estimation With Communication Costs for First-Order LTI Systems. *IEEE Trans. Auto. Control* 56, 9 (Sept. 2011), 2013–2025.

[19] Ross A. Maller, Gernot Müller, and Alex Szimayer. 2009. Ornstein-Uhlenbeck Processes and Extensions. In *Handbook of Financial Time Series*, Thomas Mikosch, Jens-Peter Kreiß, Richard A. Davis, and Torben Gustav Andersen (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 421–437.

[20] Peter Morters and Yuval Peres. 2010. Brownian Motion. *Cambridge University Press* (2010).

[21] K. Nar and T. Başar. 2014. Sampling multidimensional Wiener processes. In *IEEE CDC*. 3426–3431.

[22] Johannes Nauta, Yara Khaluf, and Pieter Simoens. 2019. Using the Ornstein-Uhlenbeck process for random exploration. In *COMPLEXIS2019*. SciTePress, 59–66.

[23] Yutian Nie and Vadim Linetsky. 2020. Sticky reflecting Ornstein-Uhlenbeck diffusions and the Vasicek interest rate model with the sticky zero lower bound. *Stochastic Models* 36, 1 (2020), 1–19.

[24] Bernt Øksendal. 2000. Stochastic Differential Equations: An Introduction with Applications (5th ed.). *Springer-Verlag Berlin Heidelberg* (2000).

[25] Bernt Oksendal. 2013. *Stochastic differential equations: an introduction with applications (5th ed.)*. Springer Science & Business Media.

[26] Frank WJ Olver, Daniel W Lozier, Ronald F Boisvert, and Charles W Clark. 2010. NIST handbook of mathematical functions hardback and CD-ROM. *Cambridge University Press* (2010).

[27] Tasmeen Zaman Ornee and Yin Sun. 2021. Performance Bounds for Sampling and Remote Estimation of Gauss-Markov Processes over a Noisy Channel with Random Delay. In *IEEE SPAWC*. 1–5.

[28] Tasmeen Zaman Ornee and Yin Sun. 2021. Sampling and Remote Estimation for the Ornstein-Uhlenbeck Process Through Queues: Age of Information and Beyond. *IEEE/ACM Trans. Netw.* 29, 5 (2021), 1962–1975.

[29] Tasmeen Zaman Ornee and Yin Sun. 2023. Technical Report of "A Whittle Index Policy for the Remote Estimation of Multiple Continuous Gauss-Markov Processes over Parallel Channels". *arXiv preprint arXiv:2305.04809* (2023).

[30] Goran Peskir and A. N. Shiryaev. 2006. Optimal Stopping and Free-Boundary Problems. *Birkhäuswer Verlag* (2006).

[31] Maben Rabi, George V. Moustakides, and John S. Baras. 2012. Adaptive Sampling for Linear State Estimation. *SIAM Journal on Control and Optimization* 50, 2 (2012), 672–702.

[32] Md Kamran Chowdhury Shisher and Yin Sun. 2022. How does data freshness affect real-time supervised learning?. In *ACM MobiHoc*. 31–40.

[33] Touraj Soleymani, Sandra Hirche, and John S. Baras. 2016. Optimal Information Control in Cyber-Physical Systems. *IFAC-PapersOnLine* 49, 22 (2016), 1 – 6.

[34] X. Song and J. W. S. Liu. 1990. Performance of multiversion concurrency control algorithms in maintaining temporal consistency. In *IEEE COMPSAC*. 132–139.

[35] Yin Sun and Benjamin Cyr. 2019. Sampling for data freshness optimization: Non-linear age functions. *Journal of Communications and Networks* 21 (2019), 204–219.

[36] Y. Sun, Y. Polyanskiy, and E. Uysal. 2020. Sampling of the Wiener Process for Remote Estimation Over a Channel With Random Delay. *IEEE Trans. Inf. Theory* 66, 2 (2020), 1118–1135.

[37] Yin Sun, Yury Polyanskiy, and Elif Uysal-Biyikoglu. 2017. Remote estimation of the Wiener process over a channel with random delay. In *IEEE ISIT*. 321–325.

[38] Yin Sun, Elif Uysal-Biyikoglu, Roy Yates, C. Emre Koksal, and Ness B. Shroff. 2016. Update or wait: How to keep your data fresh. In *IEEE INFOCOM*. 1–9.

[39] Yin Sun, Elif Uysal-Biyikoglu, Roy D Yates, C Emre Koksal, and Ness B Shroff. 2017. Update or wait: How to keep your data fresh. *IEEE Trans. Inf. Theory* 63, 11 (2017), 7492–7508.

[40] Zhifeng Tang, Zhuo Sun, Nan Yang, and Xiangyun Zhou. 2021. Whittle index-based scheduling policy for minimizing the cost of age of information. *IEEE Communications Letters* 26, 1 (2021), 54–58.

[41] Vishrant Tripathi and Eytan Modiano. 2019. A Whittle index approach to minimizing functions of age of information. In *IEEE Allerton*. 1160–1167.

[42] Cho-Hsin Tsai and Chih-Chun Wang. 2020. Unifying AoI Minimization and Remote Estimation: Optimal Sensor/Controller Coordination with Random Two-way Delay. In *IEEE INFOCOM*.

[43] Marcos M Vasconcelos and NC Martins. 2020. A survey on remote estimation problems. *Principles of Cyber-Physical Systems: An Interdisciplinary Approach* (2020), 81–103.

[44] Ina Maria Verloop. 2016. Asymptotically optimal priority policies for indexable and nonindexable restless bandits. *The Annals of Applied Probability* 26, 4 (2016).

[45] Jiazheng Wang, Xiaoqiang Ren, Yilin Mo, and Ling Shi. 2020. Whittle Index Policy for Dynamic Multichannel Allocation in Remote State Estimation. *IEEE Trans. Auto. Control* 65, 2 (2020), 591–603.

[46] Richard R Weber and Gideon Weiss. 1990. On an index policy for restless bandits. *Journal of applied probability* 27, 3 (1990), 637–648.

[47] P. Whittle. 1988. Restless Bandits: activity allocation in a changing world. *Journal of Applied Probability* 25A (1988), 287–298.

[48] Roy D Yates. 2015. Lazy is timely: Status updates by an energy harvesting source. In *IEEE ISIT*. IEEE, 3008–3012.

[49] Roy D Yates, Yin Sun, D Richard Brown, Sanjit K Kaul, Eytan Modiano, and Sennur Ulukus. 2021. Age of information: An introduction and survey. *IEEE J. Sel. Areas Commun.* 39, 5 (2021), 1183–1210.

[50] Bo Yin, Shuai Zhang, Yu Cheng, Lin X Cai, Zhiyuan Jiang, Sheng Zhou, and Zhisheng Niu. 2019. Only those requested count: Proactive scheduling policies for minimizing effective age-of-information. In *IEEE INFOCOM*. 109–117.