# AoI-Based Scheduling of Correlated Sources for Timely Inference

Md Kamran Chowdhury Shisher, *Member, IEEE*, Vishrant Tripathi, *Member, IEEE*,
Mung Chiang, *Fellow, IEEE*, and Christopher G. Brinton, *Senior Member, IEEE*

*Abstract*—We investigate a real-time remote inference system where multiple correlated sources transmit observations over a communication channel to a receiver. The receiver utilizes these observations to infer multiple time-varying targets. Due to limited communication resources, the delivered observations may not be fresh. To quantify data freshness, we employ the Age of Information (AoI) metric. To minimize the inference error, we aim to design a signal-agnostic scheduling policy that leverages AoI without requiring knowledge of the actual target values or the source observations. This scheduling problem is a restless multi-armed bandit (RMAB) problem with a non-separable penalty function. Unlike traditional RMABs, the correlation among sources introduces a unique challenge: the penalty function of each source depends on the AoI of other correlated sources, preventing the problem from decomposing into multiple independent Markov Decision Processes (MDPs), a key step in applying traditional RMAB solutions. To address this, we propose a novel approach that approximates the penalty function for each source and establishes an analytical bound on the approximation error. We then develop scheduling policies for two scenarios: (i) full knowledge of the penalty functions and (ii) no knowledge of the penalty functions. For the case of known penalty functions, we present an upper bound on the optimality gap that highlights the impact of the correlation parameter and the system size. For the case of unknown penalty functions and signal distributions, we develop an online learning approach that utilizes bandit feedback to learn an online Maximum Gain First policy. Simulation results demonstrate the effectiveness of our proposed policies in minimizing inference error and achieving scalability in the number of sources.

*Index Terms*—Age of Information, remote inference, correlated sources, scheduling, restless multi-armed bandit.

## I. INTRODUCTION

{N}{EXT-GENERATION} communications (Next-G) (e.g., 6G) are expected to support many intelligent applications such as environmental forecasting, surveillance, networked control of robot or UAV swarms, communication between connected vehicles, and massive sensing via the Internet of Things (IoTs) [2]. These applications often require timely inference of dynamic targets (e.g., positions of moving objects, changes in environmental conditions).

In this paper, we investigate timely inference of multiple time-varying targets based on observations collected from remote sources (e.g., sensors, cameras, IoT devices, UAVs). These observations are transmitted to a receiver over a capacity-limited communication channel. Furthermore, the source observations can exhibit correlation in their dynamics. For example, in environmental monitoring, the data collected by geographically close sensors measuring the temperature or humidity would be correlated. In scenarios where vehicles communicate with other neighboring vehicles, data collected from them can be spatially correlated. Similarly, in target tracking with UAV-mounted cameras, observations from cameras with overlapping fields of view are correlated. Timely and effective scheduling of correlated sources in wireless communication networks is crucial for minimizing inference errors of dynamic targets and improving real-time performance.

Due to limited communication resources, observations delivered from remote sources may not be fresh. *Age of Information* (AoI), introduced in [3] and [4], provides a convenient measure of information freshness regarding the sources at the receiver. Specifically, consider packets sent from a source to a receiver: if $U(t)$ is the generation time of the most recently received packet by time $t$, then the AoI at time $t$ is the difference between $t$ and $U(t)$. Recent works on remote inference [5], [6], [7], [8] have shown that the inference errors for different tasks can be expressed as functions of AoI. Additionally, AoI can be readily tracked, making it a promising metric for determining how to prioritize resource allocation. Recent works [9], [10] have also shown that AoI-based scheduling is sufficient to obtain near-optimal scheduling of correlated Gauss–Markov sources with linear time-invariant (LTI) system models. While this is promising, many real-world systems exhibit non-linear dynamics and complex correlation structures. Motivated by this, in this work, we pose the following research question:

> *How can we develop AoI-based scheduling of correlated sources with arbitrary correlation structure to minimize inference errors for target processes?*

### A. Outline and Summary of Contributions

In answering the above question, we make the following contributions:

- To minimize the discounted sum of inference errors for multiple targets, we formulate the problem of scheduling correlated sources over a capacity-limited channel. For the set of all causal and signal-agnostic scheduling policies in which (i) the scheduler makes its decision based on the current and the past information available and (ii) the scheduler does not have access to the realization of the actual processes, we show that the problem can be expressed as minimizing the discounted sum of AoI penalty functions (see Lemma 1 and (12)-(13)).
- The scheduling problem is a restless multi-armed bandit (RMAB) problem with a non-separable penalty function. In contrast to the traditional RMAB framework [11], the penalty functions of all arms in this problem are intertwined. This is because the penalty function for each arm depends on the AoI values of all sources. This inter-connectedness prevents the decomposition of the problem into independent Markov Decision Processes (MDPs) even with constraint relaxation, a common approach used for traditional RMABs [11]. We address this problem by establishing an information-theoretic lower bound (see Lemma 2) on inference error and use it to decompose the problem into multiple independent MDPs. Prior work [10] also provided a lower bound of inference error under correlated settings. However, the bound in [10] was limited to Gauss-Markov sources with an LTI system model. Our information-theoretic lower bound extends the bound of [10] to a significantly broader class of systems, including non-linear and non-Gaussian sources.
- After establishing the approximated function, we develop a Maximum Gain First (MGF) policy (see Algorithm 1) when the scheduler has full knowledge of penalty functions and/or the distribution of targets and source observations. We present an upper bound on the optimality gap of our MGF policy (see Theorem 2). The optimality gap highlights the impact of correlation parameter and the system size.
- For unknown penalty functions and arbitrary correlation settings, it is not possible to use the bound developed for the known penalty function. Towards this end, we first solve a Lagrangian relaxed problem with a new method of function approximation. Then, we provide an Online Threshold Policy for solving the relaxed problem in Algorithm 2 by using a bandit feedback structure. Theorem 3 analyzes the behavior of Algorithm 2.
- By using the structure of the Online Threshold Policy (Algorithm 2) and Gain index (Definition 1), we provide a novel "Online Maximum Gain First" (Online-MGF) policy in Algorithm 3. We conduct simulations (see Section VII) by using a correlation model developed in [9]. The simulation results show the effectiveness of our MGF and Online-MGF policies without knowing the exact correlation structure.

### B. Related Works

Over the past decade, there has been a rapidly growing body of research on analyzing AoI for queuing systems [4], [12], [13], [14], [15], [16], using AoI as a metric for scheduling

---

**Algorithm 1** Maximum Gain First Policy

1: *Input* $\lambda_k = 0$
2: **for** *episode* $k = 1, 2 \ldots$ **do**
3:    Evaluate $\alpha_{m,\lambda_{t,k},t}(\delta)$ for all $m$ and $\delta$
4:    Update $\mathrm{Subgrad}(t) \leftarrow 0$
5:    **for** *every time $t$ in episode $k$* **do**
6:       *Update $\Delta_m(t)$ for all $m$*
7:       *Initialize $\pi_m(t) \leftarrow 0$ for all $m$*
8:       $\alpha_m \leftarrow \alpha_{m,\lambda_{t,k},t}(\Delta_m(t))$
9:       Schedule $N$ sources with highest $\alpha_m$
10:      Sources with same gain index are ranked
11:      with a tie-breaker index function $\psi_{t,m}$
12:      $\mathrm{Subgrad}(t) \leftarrow \sum_{m=1}^{M} \gamma^t \mathbf{1}(\alpha_m > \frac{\lambda_{t,k}}{\gamma}) - N$
13:    *Update $\lambda_{k+1}$ using (29)*

---

**Algorithm 2** Online Policy for (36)

1: **for** *episode* $k = 1, 2 \ldots$ **do**
2:    *Update $\tilde{f}_{m,k-1}$ using (41).*
3:    Evaluate $\tilde{J}_{m,\lambda,k}$ for all $m$ and $\delta$ using (43).
4:    **for** *every $t$ in episode $k$* **do**
5:       *Update $\Delta_m(t)$ for all $m$*
6:       Initialize $\pi_m(t) \leftarrow 0$ for all $m$
7:       **for** $m = 1, 2 \ldots, M$ **do**
8:          **if** $\tilde{J}_{m,\lambda,k}(\Delta_m(t)+1) - \tilde{J}_{m,\lambda,k}(1) > \frac{\lambda}{\gamma}$ **then**
9:             $\pi_m(t) \leftarrow 1$; *Schedule Source $m$*

---

policies in networks [17], [18], for monitoring or controlling systems over networks [19], [20], and for optimizing remote estimation [21], [22], [23], [24], [25] and inference systems [5], [6], [7], [8], [26], [27], [28]. For detailed surveys of AoI literature see [29] and [30].

One of the first works to analyze how AoI values of correlated source observations affect remote inference was [5], in which an information-theoretic analysis was provided to understand the impact of AoI of correlated sources on inference error. However, the authors of [5] did not provide any scheduling policies to optimize inference performance under correlation. In [9], the authors studied the scheduling of correlated sources based on AoI values. This work considers a probabilistic model of correlation between the sources, and then proposed an AoI-based scheduling policy for improving information freshness at the receiver. However, the probabilistic model does not reflect the impact of correlation on inference error or estimation of targets. In [10], the authors developed an AoI-based scheduling policy for minimizing estimation error of correlated Gauss-Markov sources with an LTI system model. Our current work is related to these prior works on AoI-based scheduling of correlated sources [9], [10]. In particular, this paper generalizes the work of [9] and [10] to arbitrary systems, loss functions, and target/source processes. In addition, the prior works [9], [10] consider that the distribution of the signal processes and the penalty functions are known to the scheduler. In this paper, we extend our scheduling policy to the setting, where the distribution of the signal processes and the penalty functions are not available to the scheduler.

---

**Algorithm 3** Online Maximum Gain First Policy

1  *Input: Lagrange Multiplier $\lambda_1 = 0$*
2  **for** *episode $k = 1, 2 \ldots$* **do**
3  $\quad$ *Update $\tilde{f}_{m,k-1}$ by using (41)*
4  $\quad$ Evaluate $\tilde{J}_{m,\lambda,k}$ for all $m$ and $\delta$ by using (43)
5  $\quad$ **for** *every $t$ in episode $k$* **do**
6  $\quad\quad$ *Update $\Delta_m(t)$ for all $m$*
7  $\quad\quad$ Compute Gain index
$\quad\quad$ $\alpha_m := \tilde{J}_{m,\lambda,k}(\Delta_m(t)+1) - \tilde{J}_{m,\lambda,k}(1)$
8  $\quad\quad$ Schedule $N$ sources with highest $\alpha_m$
9  $\quad\quad$ Subgrad $\leftarrow$ Subgrad $+ \sum_{m=1}^{M} \gamma^t \mathbf{1}(\alpha_m > \frac{\lambda}{\gamma})$
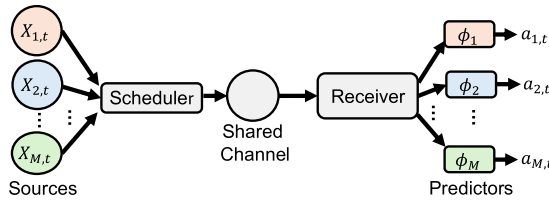10  *Update $\lambda_{k+1} \leftarrow \lambda_k + (\theta/k)(\text{Subgrad} - \frac{N}{1-\gamma})$*

---



Fig. 1. A remote inference system with $M$ correlated sources, scheduler, shared channel, receiver, and $M$ predictors.

The problem of scheduling multiple sources to optimize linear and non-linear functions of AoI can be formulated as a restless multi-armed bandit (RMAB) problem [20], [26], [31], [32], [33], [34]. While Whittle index policies [6], [11], [20], [26], [34] and gain index policies [7], [25], [28], [35] have been shown to achieve good performance and asymptotic optimality in certain RMABs [36], [37], they cannot be directly applied to our problem. This is because the correlation among sources introduces a non-separable penalty function, where the penalty for each arm depends on the AoI of other arms. This interdependence prevents the decomposition of the problem into multiple independent MDPs, a key step in applying traditional RMAB solutions. Our paper addresses this challenge by developing a novel approach to handle correlated sources with a non-separable penalty function.

Furthermore, our paper contributes to the design of online policies for RMAB problems. While existing works like [38], [39], and [40] have explored online policies, they focus on scenarios with independent penalty functions and arms. A recent study [41] investigates RMAB problems with non-separable, monotonic, and sub-modular global penalty function, in which the authors develop an approximate Whittle index-based policy. In contrast, our work considers arbitrary penalty functions and employs a gain index-based approach. Moreover, we focus on a signal-agnostic scheduler and provide scheduling policies for both known and unknown penalty functions, unlike [41], which assumes full system information.

## II. SYSTEM MODEL

Consider $M$ sources communicating over a shared wireless channel to a receiver (Fig. 1). We assume discrete slotted time and at every time slot $t$, each source $m$ observes a time-varying signal $X_{m,t} \in \mathcal{X}_m$, where $\mathcal{X}_m$ represents the finite set of pos-

sible observations from source $m$. A scheduler progressively schedules source observations to the receiver. At each time slot $t$, the receiver uses the received observations to infer $M$ targets $(Y_{1,t}, Y_{2,t}, \ldots, Y_{M,t})$ of interest. Each target $Y_{m,t}$ is drawn from a finite set $\mathcal{Y}_m$ and can be directly inferred from the corresponding source observation $X_{m,t}$. For example, $Y_{m,t}$ can be a function of $X_{m,t}$, such as $Y_{m,t} = X_{m,t}$. Moreover, the target $Y_{m,t}$ can be correlated with the observation $X_{n,t}$ of any other source $n \neq m$. The results of this paper are derived without assuming a specific correlation model. This ensures the general applicability of our results. An information theoretic interpretation on the impact of the correlation is provided in Sec. IV-A.

### A. Communication Model

Due to interference and bandwidth limitations, we assume that at most $N$ out of $M$ sources can be scheduled to transmit their observations at any time slot $t$, where $0 < N < M$. If source $m$ is scheduled at time $t$, it transmits the current observation $X_{m,t}$ to the receiver. For simplicity, we assume reliable channels, i.e., observations sent at time slot $t$ are delivered error-free at time slot $t + 1$. Because of communication constraints, the receiver may not have fresh observations from all sources. Let $X_{m,t-\Delta_m(t)}$ be the most recently delivered signal observation from source $m$ which was generated $\Delta_m(t)$ time-slots ago. We call $\Delta_m(t)$ the age of information (AoI) [4], [6] of source $m$. Let $U_m(t)$ be the generation time of the most recent delivered observation from source $m$. Then, the AoI can be formally defined as:

$$\Delta_m(t) = t - U_m(t). \tag{1}$$

Let $\pi_m(t) \in \{0, 1\}$ be the scheduling decision of source $m$. At time slot $t$, if $\pi_m(t) = 1$, the $m$-th source is scheduled to transmit its observation to the receiver; otherwise, if $\pi_m(t) = 0$, this transmission does not occur. If $\pi_m(t-1) = 0$, AoI $\Delta_m(t) = \Delta_m(t-1)+1$ grows by 1; otherwise, if $\pi_m(t-1) = 1$, AoI drops to $\Delta_m(t) = 1$.

### B. Inference Model

The receiver employs $M$ predictors. The $m$-th predictor $\phi_m$ uses the most recent observations from all sources along with their corresponding AoI values, to produce an inference result $a_{m,t} \in \mathcal{A}_m$ for target $Y_{m,t}$. Specifically, given the most recently delivered source observations and their AoI values

$$(X_{m,t-\Delta_m(t)}, \Delta_m(t))_{m=1}^{M} = (x_m, \delta_m)_{m=1}^{M},$$

the inference result

$$a_{m,t} = \phi_m((x_m, \delta_m)_{m=1}^{M}) \in \mathcal{A}_m$$

minimizes the expected loss function $\mathbb{E}[L(Y_{m,t}, a_m)]$ over all possible inference results $a_m \in \mathcal{A}_m$. $L(y_m, a_m)$ is the loss incurred when the actual target is $Y_{m,t} = y_m$ and the predicted output is $a_{m,t} = a_m$. The loss function $L(\cdot, \cdot)$ and the output space $\mathcal{A}_m$ can be designed according to the goal of the system. For example, $\mathcal{A}_m = \mathcal{Y}_m$ and quadratic loss $\|y - \hat{y}\|^2$ can be used for a regression task. In maximum likelihood estimation, we can use logarithmic loss function and the output space $\mathcal{A}_m$

can be $\mathcal{P}_{\mathcal{Y}_m}$, which is the set of all probability distributions on $\mathcal{Y}_m$. At time $t$, the expected inference error for target $Y_{m,t}$ is given by $\mathbb{E}\left[L(Y_{m,t}, a_{m,t})\right]$.

## III. PROBLEM FORMULATION

We focus on a class of *signal-agnostic* scheduling policies, where scheduling decisions are made without knowledge of the observed process's current signal values, i.e., at time $t$ the centralized scheduler does not have access to $\{(Y_{m,t}, X_{m,t}), m = 1, \ldots, M, t = 0, 1, \ldots\}$. Moreover, we make the following assumption on the observed and the target processes:

*Assumption 1:* The process

$$\{Y_{m,t}, X_{1,t}, X_{2,t}, \ldots, X_{M,t}, t \in \mathbb{Z}\}$$

is jointly stationary, i.e., the joint distributions of

$$(Y_{m,t}, X_{1,t_1}, X_{2,t_2}, \ldots, X_{M,t_M})$$

and

$$(Y_{m,t+\tau}, X_{1,t_1+\tau}, X_{2,t_2+\tau}, \ldots, X_{M,t_M+\tau})$$

are same for all $\tau, t, t_1, t_2, \ldots, t_M$. Moreover, the assumption holds for all target process $Y_{m,t}$ with $m = 1, 2, \ldots, M$.

Assumption 1 implies that the dependence of the inference target $Y_{m,t}$ on the signals of all sources remains stationary over time. Further, the correlation structure among the source observations also remains stationary over time. This allows us to show that inference error is a time-invariant function of AoI, as we will see in Lemma 1. It is practical to approximate time-varying functions as time-invariant functions in the scheduler design. Moreover, the scheduling policy developed for time-invariant AoI functions serves as a valuable foundation for studying time-varying AoI functions [42].

Now, we are ready to formulate our problem. We denote the scheduling policy as

$$\pi = (\pi_m(0), \pi_m(1), \ldots)_{m=1,2,\ldots,M}.$$

We let $\Pi$ denote the set of all signal-agnostic and causal scheduling policies $\pi$ that satisfy (i) the scheduler makes its decision at every time $t$ based on the current and the past information available to the scheduler and (ii) the scheduler does not have access to the realization of the process $\{(Y_{m,t}, X_{m,t}), m = 1, \ldots, M, t = 0, 1, \ldots\}$.

Our goal is to find a policy $\pi \in \Pi$ that minimizes the discounted sum of inference errors:

$$V_{opt} = \inf_{\pi \in \Pi} \mathbb{E}_\pi \left[ \sum_{t=0}^{\infty} \gamma^t \sum_{m=1}^{M} L(Y_{m,t}, a_{m,t}) \right], \quad (2)$$

$$\text{s.t. } \sum_{m=1}^{M} \pi_m(t) = N, t = 0, 1, \ldots, \quad (3)$$

where $0 < \gamma < 1$ is a discount factor, $L(Y_{m,t}, a_{m,t})$ is the inference error for the $m$-th target at time $t$, and $V_{opt}$ is the minimum average inference error. For the simplicity of theoretical analysis, we consider that the loss function is normalized to $[0, 1]$. However, our results can easily be generalized to any bounded loss function.

We also discuss the time-averaged version of the problem in Appendix F.

## IV. AoI-BASED PROBLEM REFORMULATION

In this section, we first use an information-theoretic approach and AoI as tools to interpret how the correlation among the $m$-th target and the observation from $n$-th source affect the inference error for the $m$-th target. To that end, we use the concept of generalized conditional entropy [43], [44] or specifically, the $L$-conditional entropy [6].

### A. An Information-Theoretic Interpretation

For a random variable $Y$, the $L$-entropy is given by

$$H_L(Y) = \min_{a \in \mathcal{A}} \mathbb{E}_{Y \sim P_Y}[L(Y, a)]. \quad (4)$$

The $L$-conditional entropy of $Y$ given $X = x$ is [6], [43], [44]

$$H_L(Y|X = x) = \min_{a \in \mathcal{A}} \mathbb{E}_{Y \sim P_{Y|X=x}}[L(Y, a)] \quad (5)$$

and the $L$-conditional entropy of $Y$ given $X$ is

$$H_L(Y|X) = \mathbb{E}_{X \sim P_X}[H_L(Y|X = x)]. \quad (6)$$

Moreover, an $L$-mutual information among two random variables $Y$ and $X$ is defined as [6], [43], and [44]

$$I_L(Y; X) = H_L(Y) - H_L(Y|X). \quad (7)$$

The $L$-mutual information $I_L(Y; X)$ quantifies the reduction of expected loss in predicting $Y$ by observing $X$. The $L$-conditional mutual information among two random variables $Y$ and $X$ given $Z$ is defined as [6], [43], and [44]

$$I_L(Y; X|Z) = H_L(Y|Z) - H_L(Y|X, Z). \quad (8)$$

If AoI of sources are $(\Delta_1(t), \Delta_2(t), \ldots, \Delta_M(t)) = (\delta_1, \delta_2, \ldots, \delta_M)$, then we can measure the reduction of expected inference error for target $Y_{m,t}$ given the freshest observations possible $X_{n,t-1}$ from source $n$ by using the definition of $L$-conditional mutual information (8):

$$
\begin{aligned}
&I_L(Y_{m,t}; X_{n,t-1}|X_{1,t-\delta_1}, X_{2,t-\delta_2}, \ldots, X_{M,t-\delta_M}) \\
&= H_L(Y_{m,t}|X_{1,t-\delta_1}, X_{2,t-\delta_2}, \ldots, X_{M,t-\delta_M}) \\
&\quad - H_L(Y_{m,t}|X_{1,t-\delta_1}, X_{2,t-\delta_2}, \ldots, X_{M,t-\delta_M}, X_{n,t-1}).
\end{aligned}
\quad (9)
$$

Now, by using the concept of the $L$-conditional entropy, we equivalently express the multi-source scheduling problem in (2)-(3) as the minimization of a penalty function of the AoI values $(\Delta_1(t), \Delta_2(t), \ldots, \Delta_M(t))$.

Lemma 1 is first proved in [5]. We restate the result for the completeness of the paper.

*Lemma 1:* The inference error for $Y_{m,t}$ can be expressed as

$$\mathbb{E}_\pi[L(Y_{m,t}, a_{m,t})] = H_L\left(Y_{m,t}|(X_{n,t-\Delta_n(t)}, \Delta_n(t))_{n=1}^{M}\right), \quad (10)$$

where $\Delta_n(t)$ is the AoI of source $n$ under the scheduling policy $\pi$. Moreover, if Assumption 1 holds, then the $L$-conditional entropy

$$H_L\left(Y_{m,t}|(X_{n,t-\Delta_n(t)}, \Delta_n(t))_{n=1}^{M}\right)$$

is a function of AoI values $(\Delta_1(t), \Delta_2(t), \ldots, \Delta_M(t))$.

Lemma 1 implies that the inference error $\mathbb{E}_\pi[L(Y_{m,t}, a_{m,t})]$ can be represented as an L-conditional entropy of $Y_{m,t}$ given most recently delivered observations from all sources and their

AoI values. Under Assumption 1, the L-conditional entropy is a function of AoI values $(\Delta_1(t), \Delta_2(t), \ldots, \Delta_M(t))$. For the simplicity of presentation, we represent the function as

$$g_m(\Delta_1(t), \Delta_2(t), \ldots, \Delta_M(t))$$
$$= H_L\left(Y_{m,t}|(X_{n,t-\Delta_n(t)}, \Delta_n(t))_{n=1}^M\right). \quad (11)$$

By using Lemma 1 and (11), we can express the problem (2)-(3) as a minimization of the discounted sum of the AoI penalty functions.

$$V_{opt} = \inf_{\pi \in \Pi} \mathbb{E}_\pi \left[ \sum_{t=0}^\infty \gamma^t \sum_{m=1}^M g_m(\Delta_1(t), \Delta_2(t), \ldots, \Delta_M(t)) \right], \quad (12)$$

$$\text{s.t. } \sum_{m=1}^M \pi_m(t) = N, t = 0, 1, \ldots, \quad (13)$$

However, as the number of sources $M$ and the number of channels $N$ increases, problem (12)-(13) becomes intractable. Specifically, problem (12)-(13) can be modeled as a Restless Multi-armed Bandit (RMAB) problem with a global penalty function:

$$\sum_{m=1}^M g_m(\Delta_1(t), \Delta_2(t), \ldots, \Delta_M(t)),$$

where AoI $\Delta_m(t)$ serves as the state of the $m$-th bandit arm. This problem is classified as "restless" because even when a source $m$ is not scheduled for transmission, its AoI $\Delta_m(t)$ continues to evolve, incurring a penalty

$$g_m(\Delta_1(t), \Delta_2(t), \ldots, \Delta_M(t)).$$

In contrast to the traditional RMAB framework [11], the penalty function of each arm in this problem are intertwined. This is because the penalty function $g_m(\cdots)$ for each arm depends on the AoI values of all sources. This interconnectedness prevents the decomposition of the problem into independent Markov Decision Processes (MDPs), a common approach used for traditional RMABs [11], even with constraint relaxation. This inherent complexity distinguishes our problem from traditional RMAB problems and poses a greater analytical challenge.

## V. DESIGN OF SCHEDULING POLICY

In this section, we will explore how to find a low complexity and close to optimal scheduling policy for (12)-(13). We also consider that the scheduler knows the penalty function $g_m(\cdots)$ for any source $m$. In Section VI, we will design scheduling policy with unknown penalty function.

### A. Function Approximation $f_m(\delta_m)$

Like any MDP, our problem (12)-(13) can be solved by using dynamic programming method [45]. However, computing the optimal policy using dynamic programming becomes progressively harder in terms of space and time complexity for larger values of $M$, as the state-space and the action space to be considered grows exponentially with $M$. Whittle index policy [6], [11], [20] and gain index policy [7], [25], [37]

are low-complexity policies that have good performance for RMAB problems. For our problem, we can not directly obtain a scheduling policy like Whittle index policy and gain index policy. This is because the penalty function $g_m(\cdot, \cdot, \cdot)$ depends on the AoI of the other sources. To this end, in this section, we approximate the penalty function $g_m(\cdot, \cdot, \cdot)$ by a function $f_m(\Delta_m(t))$ which only depends on the AoI of the source $m$. Then, we provide an index-based scheduling policy.

We can define a possible option for $f_m(\delta_m)$ as follows:

$$f_m(\delta_m) = H_L(Y_{m,t}|X_{m,t-\delta_m}, Z_{-m,t-1}) \quad (14)$$

where $Y_{m,t}$ is the $m$-th target at time $t$, $X_{m,t-\delta_m}$ is the observation generated from source $m$ at time $t - \delta_m$, $Z_{-m,t-1} = [X_{n,t-1} : \forall n \neq m]$ is a vector containing the signal observations generated at time $t-1$ from all sources except source $m$. This definition of $f_m(\delta_m)$ captures the expected inference error of the current target $Y_{m,t}$ given observation $X_{m,t-\delta_m}$ of source $m$ generated $\delta_m$ time slots ago and the recent observations $Z_{-m,t-1}$ from other sources generated 1 time slot ago.

Since we have knowledge of the penalty function $g_m(\delta_1, \delta_2, \ldots, \delta_M)$, we can evaluate the function $f_m(\delta_m)$, by setting $\delta_n = 1$ for all $n \neq m$. This effectively evaluates the function under the scenario where all other sources have an AoI of 1 except source $m$. In Appendix A, we discuss how to compute the function $f_m(\delta)$ without using $g_m(\cdot, \cdot, \cdot)$.

*Lemma 2:* If the sequence $(Y_{m,t}, X_{n,t-1}, X_{n,t-k})$ forms a Markov chain $Y_{m,t} \leftrightarrow X_{n,t-1} \leftrightarrow X_{n,t-k}$ for all $k \geq 1$, then the following assertions are true.

(a) We have

$$g_m(\delta_1, \delta_2, \ldots, \delta_M) \geq f_m(\delta_m). \quad (15)$$

(b) We have

$$g_m(\delta_1, \delta_2, \ldots, \delta_M) = f_m(\delta_m) + O(\max_{n \neq m} \epsilon_{m,n}^2). \quad (16)$$

where the parameter $\epsilon_{m,n}$ is determined as

$$\epsilon_{m,n}$$
$$= \sqrt{\max_{\delta_1, \delta_2, \ldots, \delta_M} I_L(Y_{m,t}; X_{n,t-1}|X_{1,t-\delta_1}, \ldots, X_{M,t-\delta_M})}, \quad (17)$$

and $L$-conditional mutual information $I_L()$ is defined in (8).

*Proof:* See Appendix B. □

By the definition of Markov chain $Y_{m,t} \leftrightarrow X_{n,t-1} \leftrightarrow X_{n,t-k}$, the target $Y_{m,t}$ does not depend on the older observation $X_{n,t-k}$ given a new observation $X_{n,t-1}$ from source $n$. Lemma 2(a) implies that If the Markov chain $Y_{m,t} \leftrightarrow X_{n,t-1} \leftrightarrow X_{n,t-k}$ holds, then $f_m(\delta_m)$ is a lower bound of $g_m(\delta_1, \delta_2, \ldots, \delta_M)$.

In (17), we measure the maximum reduction of expected inference error for target $m$ after adding a new observation from source $n$ by the parameter $\epsilon_{m,n}$. If the parameter $\epsilon_{m,n}$ is close to zero, the impact of correlation between target $Y_{m,n}$ and a new observation from source $n$ on inference error of $Y_{m,n}$ is nearly negligible. We say if $\epsilon_{m,n}$ is close to zero, correlation is low.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

6                                                                                                              IEEE TRANSACTIONS ON NETWORKING

Lemma (2)(b) implies that if $\epsilon_{m,n}$ tends to zero for all $n \neq m$, then the approximation error goes to zero, i.e., the function $f(\delta_m)$ is a good approximation of inference error function $g(\delta_1, \delta_2, \ldots, \delta_M)$ under low correlation.

*Remark 1:* Our information-theoretic lower bound, $f_m(\delta_m)$, extends the bound in [10] to a significantly broader class of systems, including non-linear and non-Gaussian dynamics. As shown in Appendix A, this bound coincides with the bound in [10, Theorem 2] when specialized to LTI systems with Gaussian processes.

### B. Scheduling Policy

To design a scheduling policy, we first divide entire duration into episodes, where each episode $k$ contains $T$ time slots. Here, we approximate infinite time with finite time $T$. For discounted penalty problem, the approximation error in considering finite time $T$ is negligible if $T$ is a very large number. Next, we use approximated penalty function. Thus, we reformulated the problem (12)-(13) as follows:

$$V_{f,opt}(T) = \inf_{\pi \in \Pi} \mathbb{E}_\pi \left[ \sum_{t=0}^{T-1} \gamma^t \sum_{m=1}^{M} f_m(\Delta_m(t)) \right], \quad (18)$$

$$\text{s.t.} \sum_{m=1}^{M} \pi_m(t) = N, t = 0, 1, \ldots, \quad (19)$$

where $T$ is the truncated time horizon, $V_{f,opt}(T)$ is the optimal objective value of the approximated problem (18)-(19) and $\pi_{f,opt}$ be an optimal policy to (18)-(19).

*Theorem 1:* We have

$$|V_{\pi_{f,opt}}(T) - V_{opt}(T)| \leq \frac{2(1-\gamma^T)}{1-\gamma} \sum_{m=1}^{M} O\left( \max_{n \neq m} \epsilon_{m,n}^2 \right), \quad (20)$$

where $V_{\pi_{f,opt}}(T)$ is the optimal objective value of (18)-(19) and $V_{opt}(T)$ is the optimal objective value of the problem (12)-(13) with finite horizon $T$.

*Proof:* See Appendix C.                                      □

Theorem 1 provides an approximation bound. Theorem 1 implies that if the correlation among the sources is low, i.e., $\epsilon_{m,n}$ is close to 0 for all $n \neq m$, then the optimal policy $\pi_{f,opt}$ to the approximated problem (18)-(19) will yield close to optimal performance for the main problem (12)-(13).

*1) Lagrangian Relaxation:* After establishing the approximation bound in Theorem 1, our goal is to find scheduling policy for the approximated problem (18)-(19). Due to constraint (19), finding an optimal policy for the approximated problem (18)-(19) remains computationally intractable. This problem has been proven to be PSPACE-hard [46]. We relax the constraint and get the relaxed problem:

$$V_{f,opt}(\boldsymbol{\lambda}, T)$$
$$= \inf_{\pi \in \Pi} \mathbb{E}_\pi \left[ \sum_{t=0}^{T-1} \gamma^t \left( \sum_{m=1}^{M} f_m(\Delta_m(t)) + \lambda_t(\pi_m(t) - N) \right) \right], \quad (21)$$

where $\lambda_t \in \mathbb{R}$ is the Lagrangian multiplier, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_T)$, and $V_{f,opt}(\boldsymbol{\lambda}, T)$ is the optimal objective value of the relaxed problem (21).

The dual problem to (21) is given by

$$\max_{\boldsymbol{\lambda} \in \mathbb{R}^T} V_{f,opt,\lambda}, \quad (22)$$

where $\boldsymbol{\lambda}^*$ denotes the optimal Lagrange value to the problem (22).

*2) Gain Index and Threshold Policy:* The problem (21) can be decomposed into $M$ sub-problems, where each sub-problem $m$ is given by

$$\inf_{\pi_m \in \Pi_m} \mathbb{E}_\pi \left[ \sum_{t=0}^{\infty} \gamma^t \left( f_m(\Delta_m(t)) + \lambda_t \pi_m(t) \right) \right]. \quad (23)$$

The optimal value function for each source $m$ at time $t$ is given by

$$J_{m,\boldsymbol{\lambda},t}(\delta)$$
$$= f_m(\delta) + \min \left\{ \lambda_t + \gamma J_{m,\boldsymbol{\lambda},t+1}(1), \gamma J_{m,\boldsymbol{\lambda},t+1}(\delta+1) \right\}, \quad (24)$$

where $\delta$ is the current AoI value and $J_{m,\boldsymbol{\lambda},t}$ is the value function associated with the problem (23). The value function can be obtained by using backward induction method [45].

*Lemma 3:* Given the AoI value $\Delta_m(t) = \delta$, the scheduling decision $\pi_m(t) = 1$ is optimal to (23), if the following holds:

$$J_{m,\boldsymbol{\lambda},t+1}(\delta+1) - J_{m,\boldsymbol{\lambda},t+1}(1) > \frac{\lambda_t}{\gamma}. \quad (25)$$

*Proof:* Lemma 3 holds due to (84).                      □

Lemma 3 implies that if the difference between

$$J_{m,\boldsymbol{\lambda},t+1}(\delta+1) - J_{m,\boldsymbol{\lambda},t+1}(1)$$

exceeds a threshold value $\frac{\lambda_t}{\gamma}$, then scheduling source $m$ at time $t$ is optimal to the problem (23). However, the policy can not be applied to the main problem due to scheduling constraint. Towards this end, following [7], [25], and [28], we define gain index.

*Definition 1:* [**Gain Index**] For AoI value $\delta$ and Lagrangian multiplier $\boldsymbol{\lambda}$, the gain index is given by

$$\alpha_{m,\boldsymbol{\lambda},t}(\delta) := J_{m,\boldsymbol{\lambda},t+1}(\delta+1) - J_{m,\boldsymbol{\lambda},t+1}(1). \quad (26)$$

The gain index $\alpha_{m,\boldsymbol{\lambda},t}(\delta)$ quantifies the total reduction in the discounted expected sum of inference errors from time $t+1$ to time $T-1$, when action $\pi_m(t) = 1$ is chosen at time $t$ over $\pi_m(t) = 0$, where the latter implies $m$-th source is not scheduled at time $t$. This metric enables strategic resource utilization at each time slot to enhance overall system performance. Prior works [7], [25], [28] used action-value function to define the gain index, whereas [47] called it Lagrangian index. In [7], [25], [37], and [47], source $m$ is scheduled if the difference between two action value $Q_{m,\boldsymbol{\lambda},t}(\delta, 0) - Q_{m,\boldsymbol{\lambda},t}(\delta, 1)$ becomes greater than 0, where

$$Q_{m,\boldsymbol{\lambda},t}(\delta, 0) = f_m(\delta) + \gamma J_{m,\boldsymbol{\lambda},t}(\delta+1) \quad (27)$$
$$Q_{m,\boldsymbol{\lambda},t}(\delta, 1) = f_m(\delta) + \lambda_t + \gamma J_{m,\boldsymbol{\lambda},t}(1). \quad (28)$$

Because the instantaneous penalty $f_m(\delta)$ incurred at time $t$ is same for whichever decision we take, the policies derived from $(Q_{m,\boldsymbol{\lambda},t}(\delta, 0) - Q_{m,\boldsymbol{\lambda},t}(\delta, 1))$ and our gain index-based single source scheduling are equivalent.

For the main problem, our algorithm is provided in Algorithm 1. We call it Maximum Gain First (MGF) Policy. At each time $t$, we schedule $N$ sources with highest gain index $\alpha_{m,\boldsymbol{\lambda}_{t,k},t}(\Delta_m(t))$. If multiple sources have same gain index, we use a tie-breaker index function $\psi_{t,m}(\delta)$ that can rank sources with same gain index value at any time $t$. We can use simple random mixing or Markov random mixing policy to design a tie-breaker index function as discussed in [47].

After every episode $k$, we update Lagrange multiplier by using sub-gradient ascent method as follows:

$$\lambda_{t,k+1} = \lambda_{t,k} + \frac{\theta}{k}\left(\gamma^t \sum_{m=1}^{M} \mathbf{1}\left(\alpha_{m,\lambda_{t,k},t}(\Delta_m(t)) > \frac{\lambda_{t,k}}{\gamma}\right) - \gamma^t N\right),\quad (29)$$

where $\mathbf{1}(\cdot)$ is an indicator function and $\theta/k > 0$ is step size.

### C. Performance Analysis

In this section, we analyze the performance of Algorithm 1 with an optimal policy.

Let $\pi_{\text{gain}}$ denote the policy presented in Algorithm 1 and $V_{\text{gain}}(\boldsymbol{\lambda}, T)$ represent the objective value under policy $\pi_{\text{gain}}$ associated with Lagrange multiplier $\boldsymbol{\lambda}$.

*Theorem 2:* Given a discount factor $\gamma \in (0.5, 1)$, an optimal tie-breaker function $\psi$, and the fixed ratio $r = N/M$, if $\boldsymbol{\lambda_k} = \boldsymbol{\lambda^*}$, then in episode $k$, we have

$$\frac{|V_{\text{gain}}(\boldsymbol{\lambda^*}, \boldsymbol{T}) - V_{opt}(T)|}{M}$$
$$\leq \frac{2(1-\gamma^T)}{(1-\gamma)}\frac{\sum_{m=1}^{M} O\left(\max_{n\neq m}\epsilon_{m,n}^2\right)}{M} + O\left(\beta\sqrt{\frac{(1-r)r}{M}}\right),\quad (30)$$

where

$$\beta = \frac{2(2\gamma)^T}{2\gamma - 1}.$$

*Proof:* See Appendix D. □

Theorem 2 shows the normalized optimality gap of our policy depends on the correlation parameter $\epsilon_{m,n}$ and the system size $M$. In general, for large system size i.e., $M \to \infty$, the optimality gap can be finite, specifically we have

$$\lim_{M\to\infty} \frac{|V_{\text{gain}}(\boldsymbol{\lambda^*}, \boldsymbol{T}) - V_{opt}(T)|}{M}$$
$$\leq \lim_{M\to\infty} \frac{2(1-\gamma^T)}{(1-\gamma)}\frac{\sum_{m=1}^{M} O\left(\max_{n\neq m}\epsilon_{m,n}^2\right)}{M}.\quad (31)$$

However, for a special case, if the sum of the correlation parameters satisfies

$$\sum_{m=1}^{M} O\left(\max_{n\neq m}\epsilon_{m,n}^2\right) = o(M),\quad (32)$$

we have

$$\lim_{M\to\infty} \frac{2(1-\gamma^T)}{(1-\gamma)}\frac{\sum_{m=1}^{M} O\left(\max_{n\neq m}\epsilon_{m,n}^2\right)}{M} = 0.\quad (33)$$

For distributed sensor networks with a large number of sensors, where only a few sensors exhibit mutual correlation, the special case (32) can hold.

### D. Algorithm Simplification

In Algorithm 1, we consider a large sequence of Lagrange multiplier $\boldsymbol{\lambda} = (\lambda_0, \lambda_2, \ldots, \lambda_{T-1})$. This leads to a Lagrangian dual problem (22) that is practically difficult to solve in optimality. To address this challenge, we can simplify Algorithm 1 by considering $\lambda_t = \lambda$ for all $t$. For this modification, our Lagrange update rule becomes

$$\lambda_{k+1} = \lambda_k + \frac{\theta}{k}\left(\sum_{t=0}^{T-1}\gamma^t \sum_{m=1}^{M} \mathbf{1}\left(\alpha_{m,\lambda_k}(\Delta_m(t)) > \frac{\lambda_k}{\gamma}\right) - \frac{(1-\gamma^T)N}{1-\gamma}\right).\quad (34)$$

## VI. DESIGN OF ONLINE SCHEDULING POLICY WITH UNKNOWN PENALTY FUNCTION

As discussed in the previous section, Algorithm 1 can be applied when (i) the penalty function $g_m(\delta_1, \delta_2, \ldots, \delta_M)$ or its approximation $f_m(\delta_m)$ for all $m$ are known and (ii) the correlation $\epsilon_{m,n}$ among the sources are low. In this section, we establish an online scheduling policy that does not know the penalty function $g_m(\delta_1, \delta_2, \ldots, \delta_M)$, its approximation $f_m(\delta_m)$, and the distribution of the targets and the source observations. Moreover, we consider an arbitrary correlation structure, i.e., not limited to low correlation like previous section.

If the penalty function or the observation sequences $(Y_{m,t}, (X_{n,t-\mu_n})_{n=1}^{M})$ are not known to the scheduler, it is not possible to get an online estimate of the function $f_m$ defined in (14). This is because it requires AoI values of all $M-1$ sources equal to 1 at some time slots $t$. But, it is impossible to schedule $M-1$ sources at any time $t$ unless $N = M-1$. For the online setting, we consider a new function approximation:

$$f_{m,k}(\delta) = \mathbb{E}_{\pi_k}[L(Y_{m,t}, a_{m,t})|\Delta_m(t) = \delta],\quad (35)$$

which is the expected inference error of target $m$ under policy $\pi_k$ at episode $k$ when the AoI of source $m$ is $\Delta_m(t) = \delta$. The approximated function $f_{m,k}(\delta)$ depends on the policy $\pi_k$ of episode $k$. Because the function $f_{m,k}(\delta)$ depends on the policy, it is not possible to get the function $f_{m,k}(\delta)$ before the episode $k$. Hence, the scheduler can only decide policy $\pi_k$ based on the estimation of the functions $f_{m,k-1}, f_{m,k-2}, \ldots, f_{m,0}$.

For designing online scheduling, we consider the following setting:

- The entire time horizon is divided into episodes

$$k = 0, 1, 2, \ldots,$$

each consisting of $T$ time slots. The value of $T$ is chosen to be sufficiently large such that approximating the infinite horizon discounted penalty problem with a finite horizon of length $T$ introduces negligible error.

- Computing the loss function for all time $t$ and in real-time is computationally infeasible because ground truth target is not immediately available. To address this challenge, we adopt a bandit feedback structure. This structure allows the receiver to compute the loss function only for the scheduled time slots. The receiver stores the following quantities for every source $m$: the predicted output $a_{m,t}$,

the most recent observation $X_{m,t-\Delta_m(t)}$, and $\pi_m(t)$ or equivalently $\Delta_m(t)$. Let's consider $\pi_m(j) = 1$, i.e., source $m$ is scheduled at time $t = j$. In this case, the receiver obtains $X_{m,j}$ at time $t = j + 1$. Since we have assumed the true target $Y_{m,t}$ can be inferred directly from $X_{m,t}$ (e.g., $Y_{m,t}$ can be a function of, or equivalent to, $X_{m,t}$), the receiver can use the stored $X_{m,j}$ to compute $Y_{m,j}$. With both the prediction $a_{m,j}$ and the computed true label $Y_{m,j}$, the receiver can exactly compute the loss $L(Y_{m,j}, a_{m,j})$. Crucially, this loss is only available at the scheduled time slots $j$ and not for all $t$, which defines the partial feedback characteristic of the bandit setting.

- At the beginning of every episode $k$, the scheduler decides a policy $\pi_k \in \Pi$ by using an estimate of the function $f_{m,k-1}(\Delta_m(t))$.

### A. Online Scheduling for Lagrangian Relaxed Problem

At first, we establish an online scheduling policy for Lagrangian relaxed problem. In each episode $k$, we replace $f_m(\Delta_m(t))$ in (23) by the function $f_{m,k-1}(\Delta_m(t))$ and solve the following Lagrangian relaxed problem for each source $m$:

$$\inf_{\pi_m \in \Pi_m} \mathbb{E}_\pi \left[ \sum_{t=0}^{\infty} \gamma^t \left( f_{m,k-1}(\Delta_m(t)) + \lambda \pi_m(t) \right) \right]. \quad (36)$$

Since the true function $f_{m,k-1}(\delta)$ is unknown, we use an empirical estimate, $\hat{f}_{m,k-1}(\delta)$, calculated as the average loss for a given state $\delta$ during episode $k - 1$:

$$\hat{f}_{m,k-1}(\delta)$$
$$= \frac{\sum_{t=T_{k-1}}^{T_k-1} \mathbf{1}(\Delta_m(t) = \delta, \pi_m(t) = 1) L(Y_{m,t}, a_{m,t})}{\mathcal{N}_{m,k}(\delta)}, \quad (37)$$

where $T_k$ is the starting time of episode $k$ and the frequency $\mathcal{N}_{m,k}(\delta)$ is

$$\mathcal{N}_{m,k}(\delta) = \max \left( \sum_{t=T_{k-1}}^{T_k-1} \mathbf{1}(\Delta_m(t) = \delta, \pi_m(t) = 1), 1 \right). \quad (38)$$

However, the estimate $\hat{f}_{m,k-1}(\delta)$ will naturally deviate from the true value. Given a chosen exploration parameter $\eta \in (0, 1)$, we account for the uncertainty by defining the following confidence radius

$$d_{m,k}(\delta) = \sqrt{\frac{\ln\left(\frac{2}{\eta}\right)}{2\mathcal{N}_{m,k}(\delta)}}, \quad (39)$$

which shrinks as we gather more samples for that state. Moreover, the parameter $\eta$ also controls the confidence radius: The confidence radius increases as the parameter $\eta$ decreases.

The confidence radius creates a ball of plausible values, $B_{m,k}(\delta)$:

$$B_{m,k}(\delta)$$
$$= \{ f_{m,k}(\delta) : |f_{m,k}(\delta) - \hat{f}_{m,k}(\delta)| \le d_{m,k}(\delta) \}, \quad (40)$$

Lemma 4 shows that the actual value $f_{m,k}(\delta)$ is in the ball of plausible values, $B_{m,k}(\delta)$ with probability $1 - \eta$. As the parameter $\eta$ decreases, the ball $B_{m,k}(\delta)$ expands and the

uncertainty reduces. However, as we will see next that the optimistic estimate of the actual value deviates away from the empirical estimate as $\eta$ decreases.

Following the principle of "optimism in the face of uncertainty" [48], [49], [50], we select an optimistic estimate, $\tilde{f}_{m,k-1}(\delta)$, given by

$$\tilde{f}_{m,k-1}(\delta) = \max\{\hat{f}_{m,k-1}(\delta) - d_{m,k}(\delta), 0\}. \quad (41)$$

As the parameter $\eta$ decreases, the confidence radius $d_{m,k}(\delta)$ increases and the estimate $\tilde{f}_{m,k-1}(\delta)$ deviates away from empirical value.

This optimistic value $\tilde{f}_{m,k-1}(\delta)$ is used in a value iteration algorithm to update our policy, ensuring we explore efficiently while accounting for statistical uncertainty:

$$J_{m,\text{optimistic},k}(\delta) = \tilde{f}_{m,k-1}(\delta)$$
$$+ \min \{\lambda + \gamma J_{m,\text{optimistic},k}(1), \gamma J_{m,\text{optimistic},k}(\delta + 1)\}, \quad (42)$$

where $J_{m,\text{optimistic},k}(\delta)$ can obtained by using value iteration algorithm [45].

To ensure convergence of the estimated value function, we introduce a weighting parameter $\zeta \in (0, 1)$. This parameter controls the trade-off between relying on the optimistic value function and retaining information from the history. Using $\zeta$, the estimated value function is defined as follows for all $\delta$:

$$\tilde{J}_{m,\lambda,k}(\delta) = \zeta^k J_{m,\text{optimistic},k}(\delta) + (1 - \zeta^k)\tilde{J}_{m,\lambda,k-1}(\delta), \quad (43)$$

where $0 < \zeta < 1$. This step is necessary to ensure convergence to a stable solution.

Following (43), policy $\pi_k$ is provided in Algorithm 2. The actual value function of policy $\pi_k$ associated with the problem (36) can be written as

$$J_{m,\lambda,k}(\delta)$$
$$= f_{m,k}(\delta) + \min \{\lambda + \gamma J_{m,\lambda,k}(1), \gamma J_{m,\lambda,k}(\delta + 1)\}. \quad (44)$$

Let $J^*$ be the optimal value function. Now, we are ready to establish the following results.

*Theorem 3:* The following assertions are true.

(a) There exists a value $J_m \in \mathbb{R}$ such that

$$\lim_{k \to \infty} \tilde{J}_{m,\lambda_k,k}(\delta) = J_m(\delta).$$

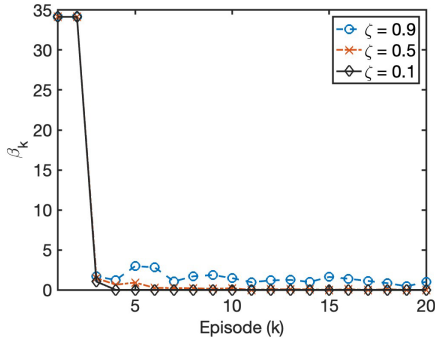(b) Given an exploration parameter $\eta \in (0, 1)$, for any episode $k$ with probability $1 - \eta$, we have

$$\left| J_{m,\lambda,k}(\delta) - J_{m,\lambda,k-1}(\delta) \right|$$
$$\le \frac{1}{1 - \gamma} \left( 2 \max_\delta d_{m,k-1}(\delta) + \beta_k \right), \quad (45)$$

where

$$\beta_k = \max_\delta |f_{m,k-1}(\delta) - f_{m,k}(\delta)| \quad (46)$$

*Proof:* See Appendix E. $\square$

Theorem 3(a) implies that the estimated value function $\tilde{J}_{m,\lambda_k,k}$ converges to a stable value function. As the policy $\pi_k$ depends on $\tilde{J}_{m,\lambda_k,k}$, we can say that the policy also converges to a stable solution, but it is not guaranteed to converge to an optimal policy.

Fig. 2.  $\beta_k$ vs. Iteration $k$.

Theorem 3(b) characterizes the distance between the value function of $k$-th episode and the value function of the $k-1$-th episode. According to the theorem, the distance is upper bounded by a confidence radius $d_{m,k-1}$ and $\beta_k$. The term $d_{m,k-1}$ is related to the uncertainty of the estimation of $f_{m,k-1}$. If number of time slot $T$ in every episode increases, the confidence radius $d_{m,k-1}$ decreases. On the other hand, we show in Fig. 2 that $\beta_k$ decreases to 0 for smaller $\zeta$.

### B. Online Maximum Gain First Policy

Algorithm 2 can not be applied as it does not satisfy the constraint (13). For this reason, similar to Algorithm 1, we utilize the concept of gain index defined in Definition (26) and the structure of the Online Threshold Policy provided in Algorithm 2 and develop a new "Online Maximum Gain First" policy in Algorithm 3.

Initially, we take Lagrange multiplier $\lambda_1 = 0$ as input. Then, the Online Maximum Gain First policy proceeds as follows:

- At the beginning of every episode $k$, we update function $\tilde{f}_{m,k-1}$ by using (41), where we use the feedback of inference loss after episode $k-1$.
- After that by using the function $\tilde{f}_{m,k-1}$, we compute value function $\tilde{J}_{m,\lambda,k}$ by using (43).
- At every time in episode $k$, we update AoI value $\Delta_m(t)$ for all $m$ and compute the gain index

$$\alpha_m = \tilde{J}_{m,\lambda,k}(\Delta_m(t)+1) - \tilde{J}_{m,\lambda,k}(1). \quad (47)$$

- $N$ sources with highest gain index are scheduled at every time $t$.
- After every episode, we update $\lambda_{k+1}$ using sub-gradient ascent method.

At the initial episode $k = 0$, the scheduler can apply Maximum Age First policy.

## VII. SIMULATION RESULTS

Though our theoretical analysis consider finite space $\mathcal{X}_m$ and $\mathcal{Y}_m$, our algorithms are general and can be applied to continuous space. Our algorithms only require countable AoI values. In this section, we will illustrate the performance of our Algorithm 1 and Algorithm 3 on a continuous space $\mathcal{X}_m$ and $\mathcal{Y}_m$ setting.

### A. Simulation Model

We use a model from [9] for designing the correlation structure between sources. At the beginning of every time-slot, each source $m$ collects information about its own state. In addition, with probability $p_{m,n}$, the update collected by source $m$ also contains information about the current state of source $n$. Motivating examples of such correlated sources are cameras with overlapping fields of view and sensors with spatial correlation between the processes being monitored. Here are two examples:

**Cameras with Overlapping Fields of View:** If camera $m$ successfully updates, there is a probability ($p_{m,n}$) that it also provides an updated view of targets associated with an adjacent camera $n$.

**Spatially Correlated Environmental Sensors:** Sensors monitoring processes (e.g., temperature, air quality, fire, rain) often exhibit spatial correlation. The probability ($p_{m,n}$) that one sensor's update contains useful, non-redundant information about an adjacent sensor's state.

We denote the state of source $m$ by $Z_{m,t}$. The state $Z_{m,t}$ is considered to evolve as

$$Z_{m,t} = a_m Z_{m,t-1} + W_{m,t}, \quad (48)$$

where $W_{m,t}$ follows zero mean Gaussian distribution with variance 1 and $W_{m,t}$ are i.i.d. over time $t$ and independent over source $m$. Let $X_{m,t}$ be the update from source $m$. The update $X_{n,t}$ includes state $Z_{n,t}$ of the source $n$ and $Z_{m,t}$ of source $n$ with probability $p_{m,n}$. A value of $p_{m,n} = 0$ suggests that there is never any information at source $n$ about the state of source $m$, while a value of $p_{m,n} = 1$ suggests that source $n$ has complete information about state of source $m$ at all times. In all of our simulation, we consider source 1 gets information about states of other sources with probability $p_{m,1} = p$. Sources except 1 never get any information of other sources, i.e., $p_{m,n} = 0$ for all $n \neq 1$ and $m \neq n$. In this simulation, we have used discount factor $\gamma = 0.7$.

Now, we discuss the relationship of the correlation parameter $\epsilon_{m,n}$ and the correlation probability $p_{m,n}$. By definition (17), we can write for the system discussed above as

$$\epsilon_{m,1}^2 =$$
$$\max_{\delta_1, \delta_2, \ldots, \delta_M} I_L(Z_{m,t}; X_{1,t-1}|X_{1,t-\delta_1}, \ldots, X_{M,t-\delta_M}). \quad (49)$$

We can decompose the conditional mutual information as

$$I_L(Z_{m,t}; X_{1,t-1}|X_{1,t-\delta_1}, \ldots, X_{M,t-\delta_M})$$
$$= H_L(Z_{m,t}|X_{1,t-\delta_1}, \ldots, X_{M,t-\delta_M})$$
$$\quad - p H_L(Z_{m,t}|Z_{m,t-1}, Z_{1,t-1}, X_{1,t-\delta_1}, \ldots, X_{M,t-\delta_M})$$
$$\quad - (1-p) H_L(Z_{m,t}|Z_{1,t-1}, X_{1,t-\delta_1}, \ldots, X_{M,t-\delta_M})$$
$$= p(H_L(Z_{m,t}|X_{1,t-\delta_1}, \ldots, X_{M,t-\delta_M}) - H_L(Z_{m,t}|Z_{m,t-1})), \quad (50)$$

where the last inequality holds because $Z_{m,t}$ and $Z_{1,t-1}$ are independent for all $m \neq 1$ and given $Z_{m,t-1}$, $Z_{m,t}$ is independent of $(X_{1,t-\delta_1}, \ldots, X_{M,t-\delta_M})$ for all $\delta_n \geq 1$. From (49) and (50), it implies that the correlation parameter $\epsilon_{m,1}$ increases as the correlation probability $p$ increases. When $p$ is zero, $\epsilon_{m,1}$ is also zero. Also, $\epsilon_{m,1}$ is maximized when $p$ is 1.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.
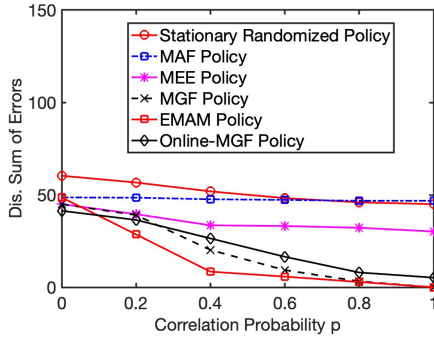
10

IEEE TRANSACTIONS ON NETWORKING



Fig. 3. Dis. Sum of Errors vs. $p$ with $M = 10$ with $a_m^2 = 0.9$ in half of sources and other half with $a_m^2 = 0.7$.
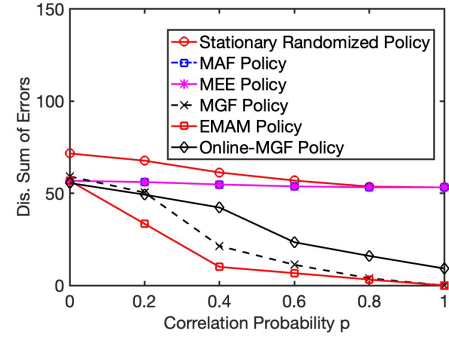


Fig. 4. Dis. Sum of Errors vs. $p$ with $M = 10$ with $a_m^2 = 0.9$ in all sources.

### B. Performance Evaluation

In this section, we first evaluate the convergence of our "Online Maximum Gain First" policy provided in Algorithm 3. Then, we compare the performance of following six policies. In our simulation, we consider $N = 1$.

- *Exponential Moving Average Max-weight (EMAM) Policy:* The policy follows [9, Algorithm 1]. The policy knows the probabilistic correlated structure within sources.
- *Maximum Expected Error (MEE) Policy:* The policy is obtained from [10], where we use that fact that $W_{m,t}$ are independent over sources.
- *Maximum Age First (MAF) Policy:* The policy schedules the source with highest AoI value.
- *Stationary Randomized Policy:* Randomized policy selects client $m$ with probability $\beta_m / \sum_{m=1}^{M} \beta_m$, for every client $m$ with $\beta_m > 0$. We consider $\beta_m = 1$.
- *Maximum Gain First (MGF) Policy:* The policy follows Algorithm 1. The policy has access to the penalty function $f_m(\delta)$ defined in (14).
- *Online Maximum Gain First (Online-MGF) Policy:* The policy follows Algorithm 3. At episode $k = 0$, Algorithm 3 applies MAF policy.

Fig. 2 illustrates the convergence behavior of our online-MGF policy. In Fig. 2, we plot $\beta_k$ defined in (46) versus episode $k$, where $\beta_k$ is evaluated empirically $\beta_k = \max_\delta |\hat{f}_{m,k-1}(\delta) - \hat{f}_{m,k}(\delta)|$. As $\beta_k$ converges to zero, policy $\pi_k$ converges. This is because the policy $\pi_k$ and $\pi_{k-1}$ are different only if $f_{m,k}$ and $f_{m,k-1}$ are different for a given Lagrange multiplier $\lambda$. From Fig. 2, we observe that smaller values of $\zeta$ ensures faster converges.

Figs. 3–4 plot the discounted sum of inference errors versus the correlation probability $p$. The results show that the Online-MGF, MGF, and EMAM policies significantly outperform the MAF, Random, and MEE policies. EMAM becomes the best-performing policy, which is expected since it is specifically designed for this correlation model. A key advantage of the Online-MGF and MGF policies is their generality. Unlike the EMAM policy, which is tailored for a specific probabilistic correlation structure, our policies are derived without assuming any model and are thus applicable to any correlation type, but the policy EMAM can not be applied to other correlation
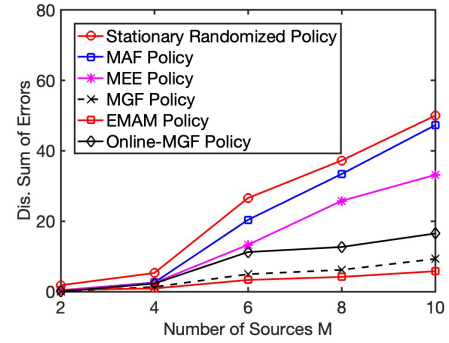


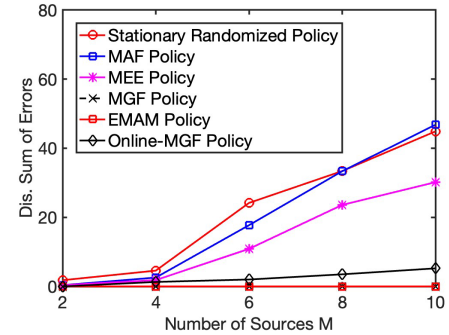Fig. 5. Dis. Sum of Errors vs. $M$ with $p = 0.6$.



Fig. 6. Dis. Sum of Errors vs. $M$ with $p = 1$.

structure. Despite this model-agnostic design, it is noteworthy that both MGF and Online-MGF perform close to EMAM. Furthermore, the Online-MGF policy performs comparably to the MGF policy without requiring knowledge of the penalty function.

Figs. 6–5 plot discounted sum of inference errors vs. number of sources $M$ with $p = 0.6$ and $p = 1$, respectively. Here, we set $a_m^2 = 0.9$ for half of the sources and $a_m^2 = 0.7$ for the other half. Similar to Figs. 3–4, Figs. 6–5 also illustrate that the Online-MGF, MGF, and EMAM policies significantly outperform the MAF, Random, and MEE policies for any number of sources. As expected, EMAM achieves the best performance by leveraging its knowledge of the probabilistic correlation structure. When $p = 1$, both the MGF and EMAM policies achieves the lowest inference error possible by scheduling the source with correlated updates.

## VIII. Conclusion and Future Work

This paper investigated AoI-based signal-agnostic scheduling policies for correlated sources to minimize inference errors for multiple time-varying targets. We developed scheduling policies for two scenarios: (i) when the scheduler has full knowledge of the penalty functions and (ii) when the scheduler has no knowledge of these functions. In the first scenario, we proposed an MGF policy and we showed an upper bound of the optimality gap in an asymptotic region. In the second scenario, we first developed an online threshold-based scheduling policy for a relaxed version of the problem. We then leveraged the structure of this policy to design a novel online-MGF policy for the original problem. Simulation results demonstrate that our online policy effectively learns and exploits the correlation structure and system dynamics to achieve significant performance gains compared to Maximum Age First and random policies.

Signal-aware scheduling problem is an important future direction, where the centralized scheduler has access to the signal values of the sources. While existing research [22], [23], [25] has demonstrated that signal-aware scheduling can improve performance compared to AoI-based scheduling, it often comes at the cost of increased complexity. In contrast, AoI-based scheduling offers the advantages of low complexity and privacy preservation, which are important considerations in many applications. Moreover, distributed scheduling of correlated sources is an interesting future direction.

## Appendix A
### How to Compute Penalty Function $f_m(\delta)$

It is possible to get closed form expressions of $f_m(\delta_m)$ for some well known random processes. An example is provided below:

*Example 1:* Let observation of $m$-th source $X_{m,t}$ and the target $Y_{m,t} = X_{m,t}$ evolves as follows:

$$X_{m,t} = a_m X_{m,t-1} + W_{m,t}, \qquad (51)$$

where the noise vector $\mathbf{W}_t = [W_{1,t}, W_{2,t}, \ldots, W_{M,t}]$ is an i.i.d. multi-variate normal random variable across time, i.e., $\mathbf{W}_t \sim \mathcal{N}(0, Q)$ and $Q$ is the covariance matrix. Let $q_{i,j}$ denote the $(i, j)$-th element of the covariance matrix $Q$. For this example, if $L(y, \hat{y}) = (y - \hat{y})^2$ is a quadratic loss, then we have

$$f_m(\delta) = \begin{cases} \bar{q}_{m,m} \delta, & \text{if } a_m = 1, \\ \bar{q}_{m,m} \dfrac{a_m^{2\delta} - 1}{a_m^2 - 1}, & \text{otherwise}, \end{cases} \qquad (52)$$

where

$$\bar{q}_{m,m} = q_{m,m} - \mathbf{q}_m Q_{-1}^{-1} \mathbf{q}_m^T,$$
$$\mathbf{q}_m = [q_{m,1}, \ldots, q_{m,m-1}, q_{m,m+1}, \ldots, q_{m,M}]$$

is the covariance of the $m$-th noise process with other process, and $Q_{-1}^{-1}$ is the noise covariance sub-matrix of the other processes excluding source $m$.

**Empirical Estimate of** $f_m(\delta)$: We can also estimate the penalty function $f_m(\delta)$ by using dataset collected in offline.

*Example 2:* Let $\{(Y_{m,t}, X_{m,t})_{m=1}^M\}_{t=1}^n$ be the dataset that contains $n$ samples of signal values from time $t = 1$ to $t = n$. Then, the empirical estimate of $f_m(\delta)$ is given by

$$f_m(\delta) = \frac{\sum_{t=\delta+1}^n L(Y_{m,t}, \phi_m(X_{m,t-\delta}, Z_{-m,t-1}))}{n - \delta}. \qquad (53)$$

## Appendix B
### Proof of Lemma 2

For the simplicity of presentation, we prove Lemma 2 for $M = 3$ and $m = 1$.

Part (a): we have

$$\begin{aligned} g_1(\delta_1, \delta_2, \delta_3) &= H_L(Y_{1,t}|X_{1,t-\delta_1}, X_{2,t-\delta_2}, X_{3,t-\delta_3}) \\ &\geq H_L(Y_{1,t}|X_{1,t-\delta_1}, X_{2,t-1}, X_{3,t-\delta_3}) \\ &\geq H_L(Y_{1,t}|X_{1,t-\delta_1}, X_{2,t-1}, X_{3,t-1}) \\ &= f_m(\delta_1), \end{aligned} \qquad (54)$$

where the inequalities hold due to [43, Lemma 12].

Part (b): We have

$$\begin{aligned} g_1(\delta_1, \delta_2, \delta_3) &= H_L(Y_{1,t}|X_{1,t-\delta_1}, X_{2,t-\delta_2}, X_{3,t-\delta_3}) \\ &= H_L(Y_{1,t}|X_{1,t-\delta_1}, X_{2,t-\delta_2}, X_{3,t-\delta_3}) \\ &\quad + H_L(Y_{1,t}|X_{1,t-\delta_1}, X_{2,t-\delta_2}, X_{3,t-\delta_3}, X_{2,t-1}) \\ &\quad - H_L(Y_{1,t}|X_{1,t-\delta_1}, X_{2,t-\delta_2}, X_{3,t-\delta_3}, X_{2,t-1}) \\ &\overset{(a)}{=} H_L(Y_{1,t}|X_{1,t-\delta_1}, X_{2,t-1}, X_{3,t-\delta_3}) \\ &\quad + H_L(Y_{1,t}|X_{1,t-\delta_1}, X_{2,t-\delta_2}, X_{3,t-\delta_3}) \\ &\quad - H_L(Y_{1,t}|X_{1,t-\delta_1}, X_{2,t-\delta_2}, X_{3,t-\delta_3}, X_{2,t-1}) \\ &\overset{(b)}{=} H_L(Y_{1,t}|X_{1,t-\delta_1}, X_{2,t-1}, X_{3,t-\delta_3}) \\ &\quad + I_L(Y_{1,t}; X_{2,t-1}|X_{1,t-\delta_1}, X_{2,t-\delta_2}, X_{3,t-\delta_3}) \\ &\overset{(c)}{=} H_L(Y_{1,t}|X_{1,t-\delta_1}, X_{2,t-1}, X_{3,t-1}) \\ &\quad + I_L(Y_{1,t}; X_{3,t-1}|X_{1,t-\delta_1}, X_{2,t-1}, X_{3,t-\delta_3}) \\ &\quad + I_L(Y_{1,t}; X_{2,t-1}|X_{1,t-\delta_1}, X_{2,t-\delta_2}, X_{3,t-\delta_3}), \\ &\overset{(d)}{\leq} f_m(\delta_1) + \epsilon_{1,3}^2 + \epsilon_{1,2}^2 \leq f_m(\delta_1) + 2 \max_{n \neq 1} \epsilon_{1,n}^2 \\ &\leq f_m(\delta_1) + O(\max_{n \neq 1} \epsilon_{1,n}^2), \end{aligned} \qquad (55)$$

where (a) holds due to Markov chain $Y_{1,t} \leftrightarrow X_{2,t-1} \leftrightarrow X_{2,t-\delta_2}$ [43, Lemma 12], (b) is obtained by definition (8), (c) can be obtained by using similar steps of (a)&(b), and (d) holds because of (17). This completes the proof for $M = 3$ and $m = 1$.

Following the same steps, we can prove it for any values of $M$ and $m$.

## Appendix C
### Proof of Theorem 1

We define

$$\Pi_c = \left\{ \pi \in \Pi : \sum_{m=1}^M \pi_m(t) = N, \forall t = 0, 1, \ldots \right\}. \qquad (56)$$

Then, from (12)-(13) and (18)-(19), we have

$$V_{opt}(T)$$

$$\inf_{\pi \in \Pi_c} \mathbb{E}_\pi \left[ \sum_{t=0}^{T-1} \gamma^t \sum_{m=1}^{M} g_m(\Delta_1(t), \Delta_2(t), \ldots, \Delta_M(t)) \right]$$

$$\stackrel{a}{=} \inf_{\pi \in \Pi_c} \mathbb{E}_\pi \left[ \sum_{t=0}^{T-1} \gamma^t \sum_{m=1}^{M} f_m(\Delta_m(t)) \right]$$

$$+ \sum_{t=0}^{T-1} \gamma^t \sum_{m=1}^{M} O\left( \max_{n=\neq m} \epsilon_{m,n}^2 \right)$$

$$= V_{f,opt}(T) + \frac{1-\gamma^T}{1-\gamma} \sum_{m=1}^{M} O\left( \max_{n \neq m} \epsilon_{m,n}^2 \right), \qquad (57)$$

where (a) is due to (16).

Moreover, we can write

$$V_{\pi_{f,opt}}(T)$$

$$= \mathbb{E}_{\pi_{f,opt}} \left[ \sum_{t=0}^{T-1} \gamma^t \sum_{m=1}^{M} g_m(\Delta_1(t), \Delta_2(t), \ldots, \Delta_M(t)) \right]$$

$$\stackrel{b}{=} \mathbb{E}_{\pi_{f,opt}} \left[ \sum_{t=0}^{T-1} \gamma^t \sum_{m=1}^{M} f_m(\Delta_m(t)) \right]$$

$$+ \sum_{t=0}^{T-1} \gamma^t \sum_{m=1}^{M} O\left( \max_{n=\neq m} \epsilon_{m,n}^2 \right)$$

$$= V_{f,opt}(T) + \frac{1-\gamma^T}{1-\gamma} \sum_{m=1}^{M} O\left( \max_{n \neq m} \epsilon_{m,n}^2 \right), \qquad (58)$$

where (b) is due to (16).

By combining (57) and (58), we obtain (20).

## APPENDIX D
## PROOF OF THEOREM 2

By using triangle inequality, We can write

$$|V_{\text{gain}}(\boldsymbol{\lambda}^*, T) - V_{opt}(T)| \leq |V_{\text{gain}}(\boldsymbol{\lambda}^*, T) - V_{f,\text{gain}}(\boldsymbol{\lambda}^*, T)|$$
$$+ |V_{f,\text{gain}}(\boldsymbol{\lambda}^*, T) - V_{f,opt}(T)|$$
$$+ |V_{f,opt}(T) - V_{opt}(T)|, \quad (59)$$

where we define

$$V_{f,\text{gain}}(\boldsymbol{\lambda}^*, T) = \mathbb{E}_{\pi_{gain}} \left[ \sum_{t=0}^{T-1} \gamma^t \sum_{m=1}^{M} f_m(\Delta_m(t)) \right]. \quad (60)$$

From (57), we obtain

$$|V_{f,opt}(T) - V_{opt}(T)| = \frac{1-\gamma^T}{1-\gamma} \sum_{m=1}^{M} O\left( \max_{n \neq m} \epsilon_{m,n}^2 \right) \quad (61)$$

Similar to (57), we can also get

$$V_{\text{gain}}(\boldsymbol{\lambda}^*, T)$$

$$= \mathbb{E}_{\pi_{gain}} \left[ \sum_{t=0}^{T-1} \gamma^t \sum_{m=1}^{M} g_m(\Delta_1(t), \Delta_2(t), \ldots, \Delta_M(t)) \right]$$

$$= \mathbb{E}_{\pi_{gain}} \left[ \sum_{t=0}^{T-1} \gamma^t \sum_{m=1}^{M} f_m(\Delta_m(t)) \right]$$

$$+ \sum_{t=0}^{T-1} \gamma^t \sum_{m=1}^{M} O\left( \max_{n=\neq m} \epsilon_{m,n}^2 \right)$$

$$= V_{f,\text{gain}}(\boldsymbol{\lambda}^*, T) + \frac{1-\gamma^T}{1-\gamma} \sum_{m=1}^{M} O\left( \max_{n \neq m} \epsilon_{m,n}^2 \right). \quad (62)$$

For an optimal tie-breaker function $\psi$ the gain index-based policy satisfies [47, Proposition 5]:

$$V_{f,\text{gain}}(\boldsymbol{\lambda}^*, T) - V_{f,opt}(T) \leq V_{f,\text{gain}}(\boldsymbol{\lambda}^*, T) - V_{f,opt}(\boldsymbol{\lambda}^*, T)$$
$$\leq \sum_{t=0}^{T-1} \beta_t \sqrt{N(1-r)}, \quad (63)$$

where $\beta_t = \frac{\gamma^t}{2\gamma - 1}((2\gamma)^{T-t} - 1)$ as provided in [47, equation (24)]. Next, we have [47, equation (59)]

$$\sum_{t=0}^{T-1} \beta_t = \frac{1}{2\gamma - 1} \left[ 2\gamma^T(2^T - 1) - \frac{1-\gamma^T}{1-\gamma} \right]. \quad (64)$$

For $\delta > 0.5$, we have $2\gamma - 1 > 0$. Then, we can drop the negative terms

$$-\frac{2\gamma^T}{2\gamma - 1} - \frac{1-\gamma^T}{1-\gamma}$$

and get

$$V_{f,\text{gain}}(\boldsymbol{\lambda}^*, T) - V_{f,opt}(T)$$
$$\leq O(\beta \sqrt{N(1-r)}) = O(\beta \sqrt{Mr(1-r)}), \quad (65)$$

where $\beta = \frac{2(2\gamma)^T}{2\gamma - 1}$.

By substituting (61)–(63) into (59), we obtain Theorem 2.

## APPENDIX E
## PROOF OF THEOREM 3

Part (a): From [51], we can say that our sub-gradient ascent method ensures the convergence of $\lambda_k$. Let $\lambda_k$ converge to $\lambda$. Next, as $k$ goes to $\infty$, $\zeta^k$ goes to 0. Hence, from (43) we can say that

$$\lim_{k \to \infty} \tilde{J}_{m,\lambda_k,k} = \tilde{J}_{m,\lambda_{k-1},k-1}.$$

This proves Theorem 3(a).

Part (b) To prove Theorem 3(b), the following Lemma is important.

*Lemma 4:* For the penalty function $f_{m,k}$, we have

$$\Pr\left( f_{m,k}(\delta) \in B_{m,k}(\delta) \right) = 1 - \eta. \quad (66)$$

Lemma 4 can be directly proven by using the following Hoeffding's inequality.

*Lemma 5 (**Hoeffding's inequality**):* Let $Z_1, Z_2, \ldots, Z_n$ are i.i.d. samples of a random variable $Z \in [0, 1]$. For any $\eta > 0$, we must have

$$\Pr\left\{ \left| \mathbb{E}[Z] - \frac{1}{n} \sum_{i=1}^{n} Z_i \right| \leq \sqrt{\frac{\ln(2/\bar{\eta})}{2n}} \right\} \geq 1 - \eta. \quad (67)$$

Now, using Lemma 4, (42)–(44), we prove Theorem 3(b). By using triangle inequality, we can have

$$\left| J_{m,\lambda,k}(\delta) - J_{m,\lambda,k-1}(\delta) \right|$$
$$\leq |J_{m,\lambda,k}(\delta) - J_{m,\text{optimistic},k}(\delta)|$$
$$+ |J_{m,\text{optimistic},k}(\delta) - J_{m,\lambda,k-1}(\delta)|. \quad (68)$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

SHISHER et al.: AoI-BASED SCHEDULING OF CORRELATED SOURCES FOR TIMELY INFERENCE                                                                13

By using [52, Theorem 12], we can have with probability $1-\eta$:

$$|\tilde{f}_{m,k-1}(\delta) - f_{m,k-1}(\delta)| \leq d_{m,k-1}(\delta). \qquad (69)$$

This yields:

$$
\begin{aligned}
&|J_{m,\text{optimistic},k}(\delta) - J_{m,\lambda,k-1}(\delta)| \\
&\leq \frac{1}{1-\gamma} \max_{\delta'} |\tilde{f}_{m,k-1}(\delta') - f_{m,k-1}(\delta')| \\
&\leq \frac{1}{1-\gamma} \max_{\delta'} d_{m,k-1}(\delta') \qquad (70)
\end{aligned}
$$

and

$$
\begin{aligned}
&|J_{m,\lambda,k}(\delta) - J_{m,\text{optimistic},k}(\delta)| \\
&\leq \frac{1}{1-\gamma} \max_{\delta'} |f_{m,k}(\delta') - \tilde{f}_{m,k-1}(\delta')| \\
&\leq \frac{1}{1-\gamma} \max_{\delta'} |f_{m,k}(\delta') - f_{m,k-1}(\delta')| \\
&\quad + \frac{1}{1-\gamma} \max_{\delta'} d_{m,k-1}(\delta'). \qquad (71)
\end{aligned}
$$

By substituting (71), (70) into (68), we obtain Theorem 3(b).

## APPENDIX F
## MINIMIZATION OF TIME-AVERAGED SUM OF INFERENCE ERRORS

Our goal is to find a policy $\pi \in \Pi$ that minimizes the time-averaged sum of inference errors:

$$\text{L}_{\text{opt}} =$$

$$\inf_{\pi \in \Pi} \limsup_{T \to \inf} \mathbb{E}_\pi \frac{1}{T} \left[ \sum_{t=0}^{T-1} \sum_{m=1}^{M} g_m(\Delta_1(t), \Delta_2(t), \ldots, \Delta_M(t)) \right], \qquad (72)$$

$$\text{s.t.} \sum_{m=1}^{M} \pi_m(t) = N, t = 0, 1, \ldots, \qquad (73)$$

where $g_m(\Delta_1(t), \Delta_2(t), \ldots, \Delta_M(t))$ is the inference error for the $m$-th target at time $t$ and $\text{L}_{\text{opt}}$ is the minimum average inference error.

For $M = 2$ sources and $N = 1$, we are able to find a low-complexity optimal policy to (72)-(73). According to Theorem 4, the optimal policy is a stationary cyclic policy.

*Definition 2 (Stationary Cyclic Policy):* A stationary cyclic policy is a stationary policy that cycles through a finite subset of points in the state space, repeating a fixed sequence of actions in a particular order.

*Theorem 4:* For $M = 2$ source case and $N = 1$, there exists a stationary cyclic policy that is optimal for (72)-(73), where the policy consists of a period $\tau_1^* + \tau_2^*$ and in each period, source 1 is scheduled for $\tau_1^*$ consecutive time slots, immediately followed by source 2 being scheduled for $\tau_2^*$ consecutive time slots. The period $\tau_1^*$ and $\tau_2^*$ minimizes

$$
\begin{aligned}
\text{L}_{\text{opt}} = \min_{\substack{\tau_1=0,1,\ldots \\ \tau_2=0,1,\ldots}} \frac{1}{\tau_1+\tau_2} &\left( \sum_{k=0}^{\tau_1-1} \sum_{m=1}^{2} (g_m(1, 2+k)) \right. \\
&\left. + \sum_{j=0}^{\tau_2-1} \sum_{m=1}^{2} (g_m(2+j, 1)) \right), \qquad (74)
\end{aligned}
$$

where $\text{L}_{\text{opt}}$ is the optimal objective value of (72)-(73).

*Proof:* Given $(\Delta_1(t), \Delta_2(t)) = (\delta_1, \delta_2)$, we can show that there exists a stationary deterministic policy that satisfies the following Bellman optimality equation [45]:

$$
\begin{aligned}
h(\delta_1, \delta_2) = \min_{(\pi_1(t),\pi_2(t))\in\mathcal{A}'} & g(\delta_1, \delta_2) - \text{L}_{\text{opt}} \\
& + \pi_1(t)h(1, \delta_2+1) + \pi_2(t)h(\delta_1+1, 1), \qquad (75)
\end{aligned}
$$

where $g(\delta_1, \delta_2) = \sum_{m=1}^{2} g_m(\delta_1, \delta_2)$, $\mathcal{A}' = \{(0,1),(1,0)\}$, $h(\delta_1, \delta_2)$ is the relative value function for the state $(\delta_1, \delta_2)$, $\text{L}_{\text{opt}}$ is the average inference error under an optimal policy, $(1, \delta_2+1)$ is the next state if source 1 is scheduled, $(\delta_1+1, 1)$ is the next state if source 2 is scheduled.

We can further express the Bellman equation (75) for state $(\delta, 1)$ at time $t$ as follows:

$$h(\delta, 1) = \min_{\tau_2 \in \{0,1,\ldots\}} \sum_{k=0}^{\tau_2} (g(\delta+k, 1) - \text{L}_{\text{opt}}) + h(1, 2), \qquad (76)$$

where $\tau_2 = 0, 1, \ldots$ is the time to keep scheduling source 2 after time $t$. By solving (76), we get that the optimal $\tau_2(\delta)$ satisfies

$$\tau_2(\delta) = \inf \{\tau \in \mathbb{Z}^+ : \gamma_1(\delta+\tau) \geq \text{L}_{\text{opt}}\}, \qquad (77)$$

where $\gamma_1(\delta+\tau)$ is defined as

$$\gamma_1(\delta) = \inf_{k=1,2,\ldots} \frac{1}{k} \sum_{j=0}^{k-1} g(\delta+1+j, 1). \qquad (78)$$

Similarly, we can express the Bellman equation (75) for state $(1, \delta)$ as follows:

$$h(1, \delta) = \min_{\tau_1 \in \{0,1,\ldots\}} \sum_{k=0}^{\tau_1} (g(1, \delta+k) - \text{L}_{\text{opt}}) + h(2, 1), \qquad (79)$$

where $\tau_1 = 0, 1, \ldots$ is the time to keep scheduling source 1 after time $t$. By solving (79), we get that the optimal $\tau_1(\delta)$ satisfies

$$\tau_1(\delta) = \inf \{\tau \in \mathbb{Z}^+ : \gamma_2(\delta+\tau) \geq \text{L}_{\text{opt}}\}, \qquad (80)$$

where $\gamma_2(\delta+\tau)$ is defined as

$$\gamma_2(\delta) = \inf_{k=1,2,\ldots} \frac{1}{k} \sum_{j=0}^{k-1} g(1, \delta+1+j). \qquad (81)$$

Moreover, by using (76) and (79), we get

$$
\begin{aligned}
h(1, 2) &= \sum_{k=0}^{\tau_1(2)} (g(1, 2+k) - \text{L}_{\text{opt}}) + h(2, 1) \\
&= \sum_{k=0}^{\tau_1(2)} (g(1, 2+k) - \text{L}_{\text{opt}}) \\
&\quad + \sum_{k=0}^{\tau_2(2)} (g(2+k, 1) - \text{L}_{\text{opt}}) + h(1, 2), \qquad (82)
\end{aligned}
$$

which yields

$$
\text{L}_{\text{opt}} = \frac{\sum_{k=0}^{\tau_1(2)} (g(1, 2+k)) + \sum_{k=0}^{\tau_2(2)} (g(2+k, 1))}{\tau_1(2) + \tau_2(2)}
$$

$$= \min_{\substack{\tau_1 = 0,1,\dots \\ \tau_2 = 0,1,\dots}} \frac{\sum_{k=0}^{\tau_1} \left(g(1, 2+k)\right) + \sum_{k=0}^{\tau_2} \left(g(2+k, 1)\right)}{\tau_1 + \tau_2}. \tag{83}$$

Notice that the optimal objective $L_{\mathrm{opt}}$ is same as if source 1 is scheduled for $\tau_1^*$ consecutive time slots, followed by source 2 being scheduled for $\tau_2^*$ consecutive time slots. $\square$

*Remark 2:* Theorem 4 presents a low-complexity optimal cyclic scheduling policy for two correlated sources. To our knowledge, this is the first such result. Prior work [53] showed the existence of an optimal cyclic policy but required solving a computationally difficult minimum average cost cycle problem over a large graph. Our result provides a significantly more efficient solution for $M = 2$ sources.

For larger values of $M$, we can have a gain index-based policy similar to discounted version of the problem, but with average cost version. To get the gain index of the average cost version, we determine the value function as follows:

The optimal relative value function for each source $m$ is given by

$$h_{m,\lambda}(\delta) = f_m(\delta) - \mathrm{f}_{\mathrm{opt},\lambda} + \min\left\{\lambda, h_{m,\lambda}(\delta+1)\right\}, \quad (84)$$

where $\delta$ is the current AoI value and $h_{m,\boldsymbol{\lambda}}$ is the relative value function associated with the problem

$$\mathrm{f}_{\mathrm{opt},\lambda} =$$
$$\inf_{\pi_m \in \Pi_m} \limsup_{T \to \inf} \frac{1}{T} \mathbb{E}_\pi \left[ \sum_{t=0}^{T-1} f_m(\Delta_m(t)) + \lambda \pi_m(t) \right], \quad (85)$$

with $h_{m,\lambda}(1) = 0$, $f(\delta)$ is defined in (14). The relative value function can be obtained by using relative value iteration algorithm [45]. The average cost gain index is given by

$$\alpha_{m,\lambda}(\delta) = h_{m,\lambda}(\delta+1) - \lambda. \tag{86}$$

For the average cost problem, the Lagrange update rule becomes

$$\lambda_{k+1} = \lambda_k + \frac{\theta}{k} \left( \frac{1}{T} \sum_{t=0}^{T-1} \sum_{m=1}^{M} \mathbf{1}\left(\alpha_{m,\lambda_k}(\Delta_m(t)) > 0\right) - N \right). \tag{87}$$

## REFERENCES

[1] M. K. C. Shisher, V. Tripathi, M. Chiang, and C. G. Brinton, "AoI-based scheduling of correlated sources for timely inference," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2025, pp. 2101–2107.

[2] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G networks: Use cases and technologies," *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 55–61, Mar. 2020.

[3] X. Song and J. W. S. Liu, "Performance of multiversion concurrency control algorithms in maintaining temporal consistency," in *Proc. 14th Annu. Int. Comput. Softw. Appl. Conf.*, 1990, pp. 132–133.

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

[5] M. K. Chowdhury Shisher, H. Qin, L. Yang, F. Yan, and Y. Sun, "The age of correlated features in supervised learning based forecasting," in *Proc. IEEE INFOCOM Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, May 2021, pp. 1–8.

[6] M. K. C. Shisher, Y. Sun, and I.-H. Hou, "Timely communications for remote inference," *IEEE/ACM Trans. Netw.*, vol. 32, no. 5, pp. 3824–3839, Oct. 2024.

[7] 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.

[8] M. K. C. Shisher and Y. Sun, "On the monotonicity of information aging," in *Proc. IEEE INFOCOM Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, May 2024, pp. 1–6.

[9] V. Tripathi and E. Modiano, "Optimizing age of information with correlated sources," in *Proc. ACM MobiHoc*, 2022, pp. 41–50.

[10] R. V. Ramakanth, V. Tripathi, and E. Modiano, "Monitoring correlated sources: AoI-based scheduling is nearly optimal," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, May 2024, pp. 641–650.

[11] P. Whittle, "Restless bandits: Activity allocation in a changing world," *J. Appl. Probab.*, vol. 25, no. A, pp. 287–298, 1988.

[12] 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.

[13] L. Huang and E. Modiano, "Optimizing age-of-information in a multi-class queueing system," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 1681–1685.

[14] A. M. Bedewy, Y. Sun, and N. B. Shroff, "Minimizing the age of information through queues," *IEEE Trans. Inf. Theory*, vol. 65, no. 8, pp. 5215–5232, Aug. 2019.

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

[16] Z. Jiang and S. Zhou, "Status from a random field: How densely should one update?," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2019, pp. 1037–1041.

[17] I. Kadota, A. Sinha, and E. Modiano, "Optimizing age of information in wireless networks with throughput constraints," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2018, pp. 1844–1852.

[18] 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, Dec. 2018.

[19] M. Klügel, M. H. Mamduhi, S. Hirche, and W. Kellerer, "AoI-penalty minimization for networked control systems with packet loss," in *Proc. IEEE INFOCOM Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Apr. 2019, pp. 189–196.

[20] V. Tripathi and E. Modiano, "A whittle index approach to minimizing functions of age of information," *IEEE/ACM Trans. Netw.*, vol. 32, no. 6, pp. 5144–5158, Dec. 2024.

[21] J. Pan, Y. Sun, and N. B. Shroff, "Sampling for remote estimation of the Wiener channel over an unreliable channel," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 52, no. 1, pp. 31–32, Jun. 2024.

[22] 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.

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

[24] T. Z. Ornee and Y. Sun, "A whittle index policy for the remote estimation of multiple continuous Gauss–Markov processes over parallel channels," in *Proc. 24th Int. Symp. Theory, Algorithmic Found., Protocol Design Mobile Netw. Mobile Comput.*, Oct. 2023, pp. 91–100.

[25] T. Z. Ornee, M. K. C. Shisher, C. Kam, and Y. Sun, "Context-aware status updating: Wireless scheduling for maximizing situational awareness in safety-critical systems," in *Proc. MILCOM IEEE Mil. Commun. Conf. (MILCOM)*. Union City, NJ, USA: Wire, Oct. 2023, pp. 194–200.

[26] M. K. C. Shisher and Y. Sun, "How does data freshness affect real-time supervised learning?," in *Proc. 23rd Int. Symp. Theory, Algorithmic Found., Protocol Design Mobile Netw. Mobile Comput.*, Oct. 2022, pp. 31–40.

[27] C. Ari, M. K. C. Shisher, E. Uysal, and Y. Sun, "Goal-oriented communications for remote inference under two-way delay with memory," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2024, pp. 1179–1184.

[28] M. K. C. Shisher, A. Piaseczny, Y. Sun, and C. G. Brinton, "Computation and communication co-scheduling for timely multi-task inference at the wireless edge," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, May 2025, pp. 1–10.

[29] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and Ş. Ulukuş, "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, May 2021.

[30] Y. Sun, I. Kadota, R. Talak, and E. Modiano, *Age of Information: A New Metric for Information Freshness*. Cham, Switzerland: Springer Nature, 2022.

[31] G. Xiong, X. Qin, B. Li, R. Singh, and J. Li, "Index-aware reinforcement learning for adaptive video streaming at the wireless edge," in *Proc. ACM MobiHoc*, 2022, pp. 81–90.

[32] 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 *Proc. 22nd Int. Symp. Theory, Algorithmic Found., Protocol Design Mobile Netw. Mobile Comput.*, Jul. 2021, pp. 11–20.

[33] Y. Chen and A. Ephremides, "Scheduling to minimize age of incorrect information with imperfect channel state information," *Entropy*, vol. 23, no. 12, p. 1572, Nov. 2021.

[34] 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, Sep. 2022.

[35] 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, Jan. 2024.

[36] I. M. Verloop, "Asymptotically optimal priority policies for indexable and nonindexable restless bandits," *Ann. Appl. Probab.*, vol. 26, no. 4, pp. 1947–1995, Aug. 2016.

[37] N. Gast, B. Gaujal, and C. Yan, "LP-based policies for restless bandits: Necessary and sufficient conditions for (exponentially fast) asymptotic optimality," 2021, *arXiv:2106.10067*.

[38] K. E. Avrachenkov and V. S. Borkar, "Whittle index based Q-learning for restless bandits with average reward," *Automatica*, vol. 139, May 2022, Art. no. 110186.

[39] K. Wang, L. Xu, A. Taneja, and M. Tambe, "Optimistic whittle index policy: Online learning for restless bandits," in *Proc. AAAI*, 2023, vol. 37, no. 8, pp. 10131–10139.

[40] J. A. Killian, A. Biswas, S. Shah, and M. Tambe, "Q-learning Lagrange policies for multi-action restless bandits," in *Proc. 27th ACM SIGKDD Conf. Knowl. Discovery Data Mining*, Aug. 2021, pp. 871–881.

[41] F. Fang, N. Raman, and Z. Shi, "Global rewards in restless multi-armed bandits," in *Proc. Adv. Neural Inf. Process. Syst.*, 2024, pp. 24625–24658.

[42] V. Tripathi and E. Modiano, "An online learning approach to optimizing time-varying costs of AoI," in *Proc. 22nd Int. Symp. Theory, Algorithmic Found., Protocol Design Mobile Netw. Mobile Comput.*, Jul. 2021, pp. 241–250.

[43] A. P. Dawid, "Coherent measures of discrepancy, uncertainty and dependence, with applications to Bayesian predictive experimental design," Dept. Stat. Sci., Univ. College London, Tech. Rep. 139, Aug. 1998. [Online]. Available: http://www.ucl.ac.uk/Stats/research/abs94.html

[44] F. Farnia and D. Tse, "A minimax approach to supervised learning," in *Proc. NIPS*, vol. 29, 2016, pp. 4240–4248.

[45] D. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1. Belmont, MA, USA: Athena Scientific, 2017.

[46] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of optimal queueing network control," in *Proc. IEEE 9th Annu. Conf. Struct. Complex. Theory*, 1994, pp. 318–322.

[47] D. B. Brown and J. E. Smith, "Index policies and performance bounds for dynamic selection problems," *Manage. Sci.*, vol. 66, no. 7, pp. 3029–3050, Jul. 2020.

[48] P. Auer and R. Ortner, "Logarithmic online regret bounds for undiscounted reinforcement learning," in *Proc. NIPS*, 2007, pp. 49–56.

[49] T. Jaksch, R. Ortner, and P. Auer, "Near-optimal regret bounds for reinforcement learning," in *Proc. NIPS*, vol. 11, 2010, pp. 1563–1600.

[50] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Mach. Learn.*, vol. 47, nos. 2–3, pp. 235–256, May 2002.

[51] S. Boyd, L. Xiao, and A. Mutapcic, "Subgradient methods," Stanford Univ., Stanford, CA, USA, Lecture Notes EE392o, 2004, no. 1.

[52] B. C. Csáji and L. Monostori, "Value function based reinforcement learning in changing Markovian environments," in *Proc. JMLR*, 2008, vol. 9, no. 54, pp. 1679–1709.

[53] P. R. Jhunjhunwala and S. Moharir, "Age-of-information aware scheduling," in *Proc. Int. Conf. Signal Process. Commun. (SPCOM)*, Jul. 2018, pp. 222–226.

**Md Kamran Chowdhury Shisher** (Member, IEEE) received the Ph.D. degree in electrical engineering from Auburn University, AL, USA, in 2024. He is currently a Post-Doctoral Researcher with the Department of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA. His research interests include stochastic optimization and sequential decision-making problems in communication networks, machine learning, and multi-agent networked control systems. He received the IEEE Communications Society William R. Bennett Prize in 2025.

**Vishrant Tripathi** (Member, IEEE) received the Ph.D. degree from the Laboratory of Information and Decision Systems (LIDS), Massachusetts Institute of Technology (MIT), in 2023. He is currently an Elmore Scholar of Excellence Assistant Professor with the Electrical and Computer Engineering Department, Purdue University. His research interests include the modeling, analysis, and design of communication networks, with emphasis on wireless and real-time networks. His current research interests include scheduling problems in networked control systems, robotics, online learning, and federated learning. He was a recipient of the ACM MobiHoc 2022 Best Paper Runner-Up Award.

**Mung Chiang** (Fellow, IEEE) is currently the 13th President of Purdue University and the Roscoe H. George Distinguished Professor of electrical and computer engineering. Previously, he was the Arthur LeGrand Doty Professor of electrical engineering with Princeton University, where he founded the Princeton Edge Laboratory in 2009 and co-founded several startups spun out from there. He is a member of the National Academy of Inventors and the Royal Swedish Academy of Engineering Science. He received the 2013 NSF Alan T. Waterman Awardee, the Guggenheim Fellowship, the IEEE Kiyo Tomiyasu Award, and the IEEE INFOCOM Achievement Award. He served as the Science and Technology Adviser for the U.S. Secretary of State.

**Christopher G. Brinton** (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in EE from Princeton University in 2013 and 2016, respectively. He is currently an Elmore Associate Professor of ECE with Purdue University. He was a recipient of the four U.S. Top Early Career Awards from the National Science Foundation (CAREER), the Office of Naval Research (YIP), the Defense Advanced Research Projects Agency (YFA), and the Air Force Office of Scientific Research (YIP). He was a recipient of the Intel Rising Star Faculty Award and the Qualcomm Faculty Award.