Optimal Sampling and Scheduling for Remote Fusion Estimation of Correlated Wiener Processes

Aimin Li and Elif Uysal, *Fellow, IEEE*

Communication Networks Research Group (CNG), EE Dept, METU, Ankara, Turkey

E-mail: liaimingogo@gmail.com; uelif@metu.edu.tr

Abstract—In distributed sensor networks, sensors often observe a dynamic process within overlapping regions. Due to random delays, these *correlated* observations arrive at the fusion center *asynchronously*, raising a central question: *How can one fuse asynchronous yet correlated information for accurate remote fusion estimation?* This paper addresses this challenge by studying the joint design of sampling, scheduling, and estimation policies for monitoring a correlated Wiener process. Though this problem is coupled, we establish a *separation principle* and identify the joint optimal policy: the optimal fusion estimator is a weighted-sum fusion estimator conditioned on Age of Information (AoI), the optimal scheduler is a Maximum Age First (MAF) scheduler that prioritizes the most stale source, and the optimal sampling can be designed given the optimal estimator and the MAF scheduler. To design the optimal sampling, we show that, under the infinite-horizon average-cost criterion, optimizing AoI is equivalent to optimizing MSE under pull-based communications, *despite the presence of strong inter-sensor correlations*. This structural equivalence allows us to identify the MSE-optimal sampler as one that is AoI-optimal. This result underscores an insight: information freshness can serve as a design surrogate for optimal estimation in *correlated* sensing environments.

Index Terms—Age of Information, Markov Decision Process, Information Fusion, Remote Estimation

I. INTRODUCTION

A. Motivation

Minimizing Age of Information (AoI) has been a central objective in status update systems, driven by the intuitive notion that *fresher data is more valuable* [1]. However, the precise impact of AoI on application-layer performance is not yet fully understood. In single-source settings, prior studies have shown that for tasks such as timely estimation [2]–[6] and inference [7]–[10], expected task performance can be analytically expressed using AoI under *pull-based communication*. For instance, when observing a Wiener process, the expected mean-square error (MSE) given a message equals its AoI [2]; for Ornstein-Uhlenbeck (OU) processes, the expected MSE is a monotonic function of AoI [3]. More generally, [7] demonstrates that inference performance may degrade according to a penalty function that is both non-linear and non-monotonic in AoI. These results imply that minimizing AoI or its potentially non-linear penalty can improve the system's estimation/reference performance.
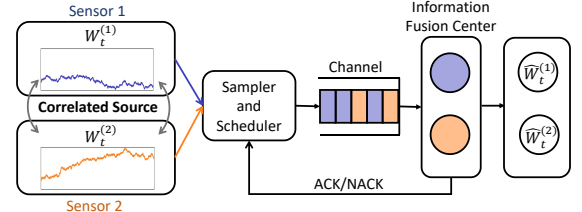
Fig. 1. System Model.

However, in multi-source systems where multiple information streams share a common communication channel, the situation becomes considerably more intricate. New challenges arise. *First*, even simply minimizing the sum of AoI across sources becomes nontrivial, requiring the joint design of sampling and scheduling policies under a shared channel [11]. *Second*, the insight into the relationship between AoI and task performance observed in single-source settings does not necessarily hold for multi-source networks. Different sources may have varying importance, dynamics, or update requirements [12], [13], making it unclear whether minimizing AoI across all sources leads to meaningful application-level gains. *Third*, and more fundamentally, updates from distributed sensors are often *correlated* in practice, *e.g.*, temperature and humidity in environmental monitoring, or video and audio in multimedia systems. In such cases, updates from one source may implicitly convey partial information about others. Consequently, the value of an update is determined not solely by its individual freshness, but by its task-relevant marginal contribution to a certain reference goal [14], [15].

B. Related Work

To address the inter-source correlations in multi-source status update systems, a heuristic metric named *Age of Correlated Information* (AoCI) was proposed in [16]. Under this model, the AoCI resets only when all relevant correlated updates are successfully received. This metric has been well-studied in various sensor network settings [17]–[19]. However, AoCI overlooks the differential contribution of each source to the final task and imposes an ideal "*all-or-nothing*" assumption that penalizes the system until all relevant updates are received—even when partial observations are already sufficient to capture the key characteristics of the system.

To address these limitations, some recent works [20]–[22] propose probabilistic frameworks where the delivery of an update from source i reduces the AoI of source j with prob-

ability p_{ij}. In [23], correlated IoT devices are grouped such that an update from one sensor implies concurrent updates in the same group. Additionally, [24] introduces an attenuation factor $\alpha_n(t)$ to model inter-source correlation, enabling partial AoI reduction based on the statistical overlap between sources.

While the aforementioned models capture information overlap caused by inter-source correlation, they often treat freshness as an abstract metric, largely decoupled from the system's estimation and other inference objectives. Consequently, although heuristic age metrics and AoI dynamics are introduced, their direct impact on application-layer estimation accuracy remains insufficiently explored.

A notable exception that explicitly links information freshness with estimation accuracy for correlated sources is the recent work by Ramakanth et al. [25], which addresses joint scheduling and estimation for monitoring a correlated Wiener process. Their framework, built on a time-slotted system with *idealized*, *zero-delay* communication channel, offers important insights into how information freshness can help improve the estimation error of correlated sources. However, real-world networks are characterized by *unavoidable* random delays, due to factors like buffer overflow, link-layer retransmissions, and routing variability—especially in dynamic environments such as Non-Terrestrial Networks (NTNs). Yet, in scenarios with random transmission delays [26]–[28], this sampling decision becomes far from trivial: the *zero-wait* sampling, while simple, can severely degrade estimation performance.

C. Contributions of This Work

- **System Model**: We investigate the joint design of sampling, scheduling, and estimation policies for *correlated* Wiener processes in the presence of random delays. Unlike the time-slotted model in [25], which considers ideal immediate delivery and focuses only on scheduling, our continuous-time framework incorporates stochastic delays typical in real-world networks. Moreover, we include sampling decisions as part of the design, which determines when each sensor should generate and transmit updates. This is motivated by recent findings showing that non-zero-waiting policies can outperform zero-waiting policies under random delay [26]–[28].

- **Separation-Structured Optimal Design:** We establish a separation principle for the joint design of sampling, scheduling, and estimation in correlated sensing. This principle enables a modular decomposition of the problem, leading to a structured policy that achieves optimal long-term average MSE in estimation and scheduling, and near-optimal performance in sampling. The joint policy comprises: (i) a weighted-sum fusion estimator; (ii) a Maximum Age First (MAF) scheduler; and (iii) a threshold-triggered Water-Filling (WF) sampler.

- **Methodology Design:** Solving the joint optimization problem poses a key technical challenge, as the state space of the formulated infinite-horizon MDP is both *transient* and *unbounded*, rendering classical average-cost approaches such as the *Average Cost Optimality Equation*

(ACOE) and *Relative Value Iteration* (RVI) inapplicable. Existing remedies, including *discounting* and *state truncation* [29], [30], often introduce bias or compromise optimality. In contrast, we construct a new MDP with *bounded* and *recurrent* dynamics. We prove that the transformed MDP is equivalent to the original MDP in terms of average-cost performance. This reduction allows us to shift the analysis from an intractable MDP to a well-behaved MDP, thereby enabling efficient optimization while preserving theoretical optimality.

- **Theoretical Insight:** We reveal that the transformed well-behaved MDP is mathematically equivalent to a classical long-term average sum-AoI minimization problem. This equivalence implies that, under the infinite-horizon average-cost criterion, an AoI-optimal sampling policy for multiple sources [11] also minimizes the MSE, *even in the presence of strong inter-source correlations*.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a system where spatially distributed sensors monitor a dynamic physical process (e.g., temperature or pressure), with correlated observations due to overlapping sensing regions. Sensor samples are transmitted over a shared random-delay channel to a remote estimator, which fuses the collected information and reconstructs the underlying process in real-time. The system model is shown in Fig. 1.

A. Coupled Correlated Sources

The monitored dynamic source is modeled as a correlated bivariate Wiener process $\mathbf{W}_t = (W_t^{(1)}, W_t^{(2)})^\top$, constructed from two independent standard Wiener processes $B_t^{(1)}$ and $B_t^{(2)}$, defined on a common filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$. Specifically, the dynamics are given by:

$$
\begin{aligned}
dW_t^{(1)} &= dB_t^{(1)}, \\
dW_t^{(2)} &= \rho\, dB_t^{(1)} + \sqrt{1 - \rho^2}\, dB_t^{(2)},
\end{aligned}
\tag{1}
$$

where $\rho \in [0, 1]$ quantifies the correlation due to distributed sensors' spatial deployment or sensing modalities.

B. Multi-source Network Model

Consider two distributed sensors where the sensor $m \in \{1, 2\}$ monitors $W_t^{(m)}$. Let S_i denote the sampling time of the i-th sample, and let $a_i \in \{1, 2\}$ indicate the scheduled sensor corresponding to that sample. Each sample $W_{S_i}^{(a_i)}$ is encapsulated into a packet $(S_i, W_{S_i}^{(a_i)})$ and transmitted to a remote data center for remote information fusion estimation. Each packet experiences a random transmission delay $Y_i \in \mathcal{Y}$ due to factors such as network handover, congestion, packet size variability, and retransmissions. The delay support $\mathcal{Y} \subset \mathbb{R}^+$ is assumed to be finite and bounded. We adopt a non-preemptive scheduling policy: a new sample cannot be generated while the communication channel is occupied by an ongoing transmission. Let $D_i = S_i + Y_i$ denote the delivery time of the i-th sample. To enhance remote estimation performance, the sampler is allowed to insert a deliberate

waiting time $Z_i \in \mathcal{Z}$ before generating the next sample[1], where $S_{i+1} = D_i + Z_i$. The waiting time set $\mathcal{Z} \subset \mathbb{R}^+$ is also assumed to be finite and bounded. Without loss of generality, we assume $D_0 = 0$ and $S_0 = 0$.

At any time t, the most recently delivered sample from source m is generated at time:

$$U_m(t) = \max\{S_i : a_i = m, D_i \leq t\}, \quad \forall m \in \{1, 2\}. \quad (2)$$

Accordingly, the AoI of source m at time t is defined as:

$$\Delta_m(t) \triangleq t - U_m(t), \forall m \in \{1, 2\}. \quad (3)$$

C. Sampling, Scheduling, and Estimating Policies

We consider a *pull-based* communication framework in which all decisions, including sampling, scheduling, and fusion-based estimation, are made by the information fusion center. These decisions are made *causally*, relying on past sampling and scheduling actions, as well as previously received sample updates. At any time $t \in \mathbb{R}^+$, the information available for making such *causal* decisions(*history*) is given by:

$$\mathcal{I}_t = \left\{ \left(a_i, S_i, W_{S_i}^{(a_i)}, D_i \right) \Big| D_i \leq t \right\}. \quad (4)$$

Here, $a_i \in \{1, 2\}$ indicates the source index, S_i is the sampling time, D_i is the delivery time, and $W_{S_i}^{(a_i)}$ is the corresponding sample value.

A *causal policy* at the fusion center includes the following components:

- **Estimator** g: This function determines the real-time estimate $\hat{\mathbf{W}}_t$ based on the *history* $\mathcal{I}_t$:

$$\hat{\mathbf{W}}_t = (\hat{W}_t^{(1)}, \hat{W}_t^{(2)})^\top = g(\mathcal{I}_t), \quad \forall t \in \mathbb{R}^+. \quad (5)$$

- **Scheduler** π: A scheduler is a policy function that determines which source is selected for transmission. The selected source during the i-th transmission is based on the *history* $\mathcal{I}_{D_{i-1}}$:

$$a_i = \pi(\mathcal{I}_{D_{i-1}}), \quad \forall i \in \mathbb{N}. \quad (6)$$

  This policy leads to a scheduling order $(a_1, a_2, \cdots)$.

- **Sampler** f: A sampler is a function that determines when each sensor is sampled. The sampling time of the i-th packet is determined based on the *history* $\mathcal{I}_{D_{i-1}}$:

$$S_i = f(\mathcal{I}_{D_{i-1}}), \quad \forall i \in \mathbb{N}. \quad (7)$$

  Equivalently, the sampler can be defined via the waiting time after the previous delivery, where $Z_i = S_i - D_{i-1}$ and $Z_i \in \mathcal{Z}$, leading to the sequence $(Z_1, Z_2, \dots)$.

D. Problem Formulation

We aim to design sampling, scheduling, and estimation policies that jointly minimize the long-term average MSE. For any admissible policy triple (g, π, f), the performance is

evaluated via the long-term average MSE:

$$\overline{\mathrm{MSE}}(g, \pi, f) \triangleq \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_{g, \pi, f} \left[\int_0^T \left\| \mathbf{W}_t - \hat{\mathbf{W}}_t \right\|_2^2 \mathrm{d}t \right], \quad (8)$$

where $\mathbf{W}_t \in \mathbb{R}^2$ is the source process and $\hat{\mathbf{W}}_t \in \mathbb{R}^2$ is its remote fusion estimate. In this work, we focus on the following optimization problem:

Problem 1. *Find a policy triple (g, π, f) that minimizes the long-term average MSE:*

$$\varphi^\star \triangleq \inf_{g, \pi, f} \overline{\mathrm{MSE}}(g, \pi, f), \quad (9)$$

where $\varphi^\star$ denotes the optimal long-term average MSE.

III. OPTIMAL ESTIMATING, SAMPLING, AND SCHEDULING

A. Separation Principle and Optimal Estimator

Under a given scheduling policy π and a sampling policy f, the minimum mean-square error (MMSE) estimator [31, IV.B.8], which minimizes the instantaneous mean-square error in the system, is the conditional expectation of each source process given the available *history*:

$$\hat{W}_t^{(1)} = \mathbb{E}\left[W_t^{(1)} \mid \mathcal{I}_t\right], \quad \hat{W}_t^{(2)} = \mathbb{E}\left[W_t^{(2)} \mid \mathcal{I}_t\right]. \quad (10)$$

The following theorem provides an explicit characterization of the MMSE estimator.

Theorem 1 (*MMSE Fusion Estimator for Correlated Wiener Processes*). *Let $\rho \in [0, 1]$ denote the correlation coefficient between $W_t^{(1)}$ and $W_t^{(2)}$. Let $\Delta_m(t)$ denote the AoI of source m. Given any scheduling policy π and sampling policy f, the MMSE estimate of $W_t^{(1)}$ is:*

$$\hat{W}_t^{(1)} = \begin{cases} W_{t-\Delta_1(t)}^{(1)}, & \Delta_1(t) \leq \Delta_2(t) \\ G_{t,\rho}(\Delta_1(t), \Delta_2(t)) \cdot W_{t-\Delta_1(t)}^{(1)} + \\ Q_{t,\rho}(\Delta_1(t), \Delta_2(t)) \cdot W_{t-\Delta_2(t)}^{(2)}, & \Delta_1(t) > \Delta_2(t) \end{cases} \quad (11)$$

The MMSE estimation of $W_t^{(2)}$ is:

$$\hat{W}_t^{(2)} = \begin{cases} W_{t-\Delta_2(t)}^{(2)}, & \Delta_2(t) \leq \Delta_1(t) \\ G_{t,\rho}(\Delta_2(t), \Delta_1(t)) \cdot W_{t-\Delta_2(t)}^{(2)} + \\ Q_{t,\rho}(\Delta_2(t), \Delta_1(t)) \cdot W_{t-\Delta_1(t)}^{(1)}, & \Delta_2(t) > \Delta_1(t) \end{cases} \quad (12)$$

where the functions $G_{t,\rho}(x, y)$ and $Q_{t,\rho}(x, y)$ are defined by:

$$G_{t,\rho}(x, y) \triangleq \frac{(1 - \rho^2)(t - y)}{t - y - \rho^2(t - x)}, \quad Q_{t,\rho}(x, y) \triangleq \frac{\rho(x - y)}{t - y - \rho^2(t - x)}, \quad (13)$$

and we refer to them as fusion weights.

Proof. See Appendix A. ∎

The results above reveal how the remote data fusion center exploits asynchronous and potentially *stale* information from distributed sensors to perform real-time estimation of correlated sources. Specifically, the MMSE estimator for $W_t^{(m)}$ depends on the relative freshness (i.e., AoI) of the most

[1]As discussed in [2], sampling immediately upon delivery (i.e., *zero-wait* sampling), although *throughput-optimal*, is suboptimal for remote estimation.

recently received samples from both sources. If the latest sample from source m is fresher than the other source m', i.e., $\Delta_m(t) \leq \Delta_{m'}(t)$, then the MMSE estimate of $W_t^{(m)}$ is simply the value of its most recently received sample:

$$\hat{W}_t^{(m)} = W_{t-\Delta_m(t)}^{(m)}. \tag{14}$$

This corresponds to the intuitive case in which fresher information dominates. However, if the sample from the source m' is fresher, i.e., $\Delta_m(t) > \Delta_{m'}(t)$, then the MMSE estimate of $W_t^{(m)}$ must fuse both the staler sample $W_{t-\Delta_m(t)}^{(m)}$ and the fresher correlated sample $W_{t-\Delta_{m'}(t)}^{(m')}$. These two components are weighted by coefficients that depend on the correlation ρ, the AoIs $\Delta_1(t)$ and $\Delta_2(t)$, and the current time t. Moreover, by choosing specific values of ρ, our results recover known estimation structures studied in the existing literature.

- If $\rho = 1$, the fusion coefficients simplify to $G_{t,1}(x,y) \equiv 0$ and $Q_{t,1}(x,y) \equiv 1$. Substituting these into the MMSE estimator (Theorem 1) leads to the degenerate result:

$$\hat{W}_t^{(1)} = \hat{W}_t^{(2)} = \begin{cases} W_{t-\Delta_1(t)}^{(1)}, & \text{if } \Delta_1(t) \leq \Delta_2(t) \\ W_{t-\Delta_2(t)}^{(2)}, & \text{if } \Delta_1(t) > \Delta_2(t). \end{cases} \tag{15}$$

Note that $W_t^{(1)} = W_t^{(2)}$ when $\rho = 1$; in this case, the problem reduces to the classical single-source remote estimation scenario, as studied in [2].

- If $\rho = 0$, the fusion weights reduce to $G_{t,0}(x,y) \equiv 1$ and $Q_{t,0}(x,y) \equiv 0$. Substituting these into the MMSE estimator yields:

$$\hat{W}_t^{(m)} = W_{t-\Delta_m(t)}^{(m)}, \quad \forall m. \tag{16}$$

This corresponds to the classical setting for remote estimation of *independent* sources, where each source can be estimated independently, as discussed in [12] and [13].

B. Maximum Age First Scheduling is All You Need

In pull-based communication systems, the receiver pulls updates based on its own information requirements. Given the available information *history* $\mathcal{I}_t$, the expected MSE at time t is given by:

$$\epsilon(t) \triangleq \mathbb{E}\left[\|\mathbf{W}_t - g_{\text{MMSE}}(\mathcal{I}_t)\|_2^2 \mid \mathcal{I}_t\right] = \Delta_1(t) + \Delta_2(t)$$
$$- \frac{\rho^2 (\Delta_1(t) - \Delta_2(t))^2}{(1-\rho^2)t + \rho^2 \max\{\Delta_1(t), \Delta_2(t)\} - \min\{\Delta_1(t), \Delta_2(t)\}}. \tag{17}$$

This result reveals a significant insight: *the estimation error depends not only on the individual values of $\Delta_1(t)$ and $\Delta_2(t)$, but also their interaction adjusted by ρ.* When $\rho = 0$, the sources are independent, and the estimation error simplifies neatly to the sum of AoI values:

$$\epsilon(t) = \Delta_1(t) + \Delta_2(t), \tag{18}$$

a form that aligns with prior results in uncorrelated multi-source settings, *e.g.*, [12] and [13]. Consequently, optimizing update scheduling in this setting is equivalent to minimizing the sum of AoI [11].

In this work, we deliberately focus on the more complex and practical regime in which the correlation between sources is strictly positive ($\rho > 0$). Let $S_{i,m} = \max\{S_k : a_k = m, S_k < D_i\}$ denote the most recent sampling time from source m prior to the i-th update interval. Since only one source is sampled per interval, the AoI for source m evolves linearly during $[D_i, D_{i+1})$ as:

$$\Delta_m(t) = t - S_{i,m}, \quad \text{for } t \in [D_i, D_{i+1}). \tag{19}$$

Substituting (19) into (17), the instantaneous estimation error $\epsilon(t)$ can be expressed explicitly as a function of $S_{i,1}$ and $S_{i,2}$:

$$\epsilon(t) = (t - S_{i,1}) + (t - S_{i,2}) -$$
$$\frac{\rho^2(S_{i,1} - S_{i,2})^2}{\max\{S_{i,1}, S_{i,2}\} - \rho^2 \min\{S_{i,1}, S_{i,2}\}}, \quad \forall t \in [D_i, D_{i+1}). \tag{20}$$

The next step is to evaluate the cumulative estimation error accumulated over each update interval $[D_i, D_{i+1})$. Let the length of each update interval be $D_{i+1} - D_i = Y_{i+1} + Z_i$. Let $\mu_Y = \mathbb{E}[Y]$ and $\sigma_Y = \mathbb{E}[Y^2]$ denote the first and second moments of the delay distribution. Then, the the expected accumulated MSE within the interval $[D_i, D_{i+1})$ takes the following form:

$$\mathbb{E}\left[\int_{D_i}^{D_i+Y_{i+1}+Z_i} \epsilon(t)dt \,\middle|\, \mathcal{I}_{D_i}\right]$$
$$= \sigma_Y + \mu_Y Z_i + (Z_i + \mu_Y)(2Y_i + Z_i + q_\rho(S_{i,1}, S_{i,2})). \tag{21}$$

Here, the function $q_\rho(x,y)$ is defined as:

$$q_\rho(x,y) = |x - y| \times \left(1 - \frac{\rho^2|x-y|}{\max\{x,y\} - \rho^2 \min\{x,y\}}\right). \tag{22}$$

For notational simplicity, let $\mathcal{C}_\rho(S_{i,1}, S_{i,2}, Y_i, Z_i)$ denote the right-hand side of (21), representing the total expected cost incurred during the i-th interval. Building on this definition, the long-term average MSE minimization problem can be reformulated as minimizing the long-term time-average of per-interval expected costs:

$$\min_{\pi,f} \lim_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\int_0^T \left\|\mathbf{W}_t - \hat{\mathbf{W}}_t\right\|_2^2 dt\right]$$
$$= \min_{\pi,f} \lim_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}\left[\mathbb{E}\left[\int_{D_i}^{D_{i+1}} \epsilon(t)dt \,\middle|\, \mathcal{I}_{D_i}\right]\right]}{\mathbb{E}[D_n]}$$
$$= \min_{\pi,f} \lim_{n\to\infty} \frac{\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}[\mathcal{C}_\rho(S_{i,1}, S_{i,2}, Y_i, Z_i)]}{\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}[Y_{i+1} + Z_i]}. \tag{23}$$

To address this fractional optimization, we adopt the classical Dinkelbach method [32]. For any $\lambda \geq 0$, , we define the following parameterized auxiliary problem:

Problem 2 (*Average–cost reformulation via Dinkelbach*)**.**

$$\Theta(\lambda) = \min_{\pi,f} \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}\left[\mathcal{C}_\rho(S_{i,1}, S_{i,2}, Y_i, Z_i) - \lambda(Y_{i+1} + Z_i)\right]. \tag{24}$$

The following lemma establishes the equivalence between

the fractional problem (23) and its reformulation (24).

Lemma 1 (*Equivalence via Dinkelbach*)**.** *Let $\varphi^\star$ denote the optimal value of the fractional problem in (23). Then the following assertions hold:*

1) $\varphi^\star \gtreqqless \lambda$ *if and only if* $\Theta(\lambda) \gtreqqless 0$.
2) *When* $\Theta(\lambda) = 0$, *then any policy that minimizes (24) also solves the original fractional problem (23).*

Proof. The proof follows standard arguments in fractional programming and is omitted here due to page constraints. ∎

Problem 2 admits a natural *infinite–horizon average–cost MDP* formulation:

MDP 1 (*Original Infinite-Horizon MDP*)**.** *This MDP can be specified by the following components:*
- *State:* $x_i \triangleq \big(S_{i,1}, S_{i,2}, Y_i\big) \in \mathbb{R}_+^3$.
- *Action:* $u_i \triangleq (a_i, Z_i)$ *with scheduling action* $a_i \in \{1,2\}$ *and waiting time* $Z_i \in \mathbb{R}_+$.
- *One–step cost:*

$$
\begin{aligned}
c_\lambda^\rho(x_i, u_i) &\triangleq \mathcal{C}_\rho\big(S_{i,1}, S_{i,2}, Y_i, Z_i\big) - \lambda\big(\mathbb{E}[Y_{i+1}] + Z_i\big) \\
&= \sigma_Y + \mu_Y Z_i - \lambda\mu_Y - \lambda Z_i \\
&\quad + (Z_i + \mu_Y)\left(2Y_i + Z_i + q_\rho\left(S_{i,1}, S_{i,2}\right)\right).
\end{aligned}
\tag{25}
$$

which coincides with the right–hand side of (24).
- *State transition: Conditioned on* (x_i, u_i), *for any source index* m, *the sampling time evolves as:*

$$
S_{i+1,m} = \begin{cases} \max\{S_{i,1}, S_{i,2}\} + Y_i + Z_i, & \text{if } a_i = m, \\ S_{i,m}, & \text{otherwise}, \end{cases}
\tag{26}
$$

while Y_{i+1} *is drawn i.i.d. from the prescribed delay distribution and is independent of* Y_i.

However, in the MDP above, the state x_i may be *transient and unbounded*, with the sampling times $S_{i,1}, S_{i,2}$ growing without bound. As a result, standard dynamic programming tools for average-cost problems, such as the ACOE [33, Eq. 4.1] and its solution via RVI [34] become inapplicable. To address this technical challenge, we propose a two-stage MDP transformation. We first introduce the following auxiliary variables:

$$
\Gamma_i = |S_{i,1} - S_{i,2}|, \quad M_i = \max\{S_{i,1}, S_{i,2}\}, \quad \forall i
\tag{27}
$$

where Γ_i denotes the age gap between the two sources and M_i captures their common upper envelope. We redefine the state of the system as the pair (Γ_i, M_i). Based on this state space transformation, we construct the following equivalent MDP:

MDP 2 (*State Space Transformation*)**.** *This MDP is defined by the following components:*
- *State:* $x_i \triangleq \big(\Gamma_i, M_i, Y_i\big) \in \mathbb{R}_+^3$.
- *Action:* $u_i \triangleq (a_i, Z_i) \in \{1,2\} \times \mathbb{R}_+$.
- *One–step cost:*

$$
\begin{aligned}
c_\lambda^\rho(x_i, u_i) &= \sigma_Y + \mu_Y Z_i - \lambda\mu_Y - \lambda Z_i \\
&\quad + (Z_i + \mu_Y)\left(2Y_i + Z_i + H_\rho\left(\Gamma_i, M_i\right)\right),
\end{aligned}
\tag{28}
$$

where $H_\rho(\Gamma_i, M_i) = q_\rho(S_{i,1}, S_{i,2})$, *with*

$$
H_\rho(\Gamma_i, M_i) = \Gamma_i\left(1 - \frac{\rho^2\Gamma_i}{\Gamma_i + (1-\rho^2)M_i}\right).
\tag{29}
$$

- *State transition:* M_i *evolves as:*

$$
M_{i+1} = M_i + Y_i + Z_i.
\tag{30}
$$

Γ_i *evolves as:*

$$
\Gamma_{i+1} = \begin{cases} \Gamma_i + Y_i + Z_i, & \text{if } a_i = \arg\max_m\{S_{i,m}\}, \\ Y_i + Z_i, & \text{if } a_i = \arg\min_m\{S_{i,m}\}, \end{cases}
\tag{31}
$$

With this MDP, although M_i may still grow unbounded over time, Γ_i can be bounded under a suitable policy. This motivates us to examine a structural property of the one-step cost function, captured in the following lemma.

Lemma 2 (*Monotonicity in the age gap* Γ)**.** *The function $H_\rho(M, \Gamma)$ as well as the one–step cost $c_\lambda^\rho(\Gamma, M, Y, u)$ in (28) are strictly increasing in Γ.*

Proof. For every fixed envelope level M we have

$$
\frac{\partial H_\rho(M, \Gamma)}{\partial \Gamma} = 1 - \frac{\rho^2\Gamma(\Gamma + 2(1-\rho^2)M)}{(\Gamma + (1-\rho^2)M)^2} > 1 - \rho^2 > 0,
\tag{32}
$$

Hence, $H_\rho(M, \Gamma)$ is strictly increasing in Γ. It follows immediately that $c_\lambda^\rho(\Gamma, M, Y, u)$ is also increasing with Γ. ∎

Leveraging the monotonicity property in Lemma 2, we now show that the MAF scheduling policy is optimal for the transformed MDP 2.

Theorem 2 (*Separation principle and optimality of MAF scheduling*)**.** *For every $\lambda \geq 0$ and any given sampling policy f, the stationary policy that always schedules the sensor with the maximum age,*

$$
a_i^{MAF} = \arg\max_{m \in \{1,2\}} \Delta_m(D_i), \quad \forall i \in \mathbb{N},
\tag{33}
$$

is optimal for the average–cost MDP defined by (28)–(31).

Proof. Fix an arbitrary epoch index ℓ. Construct two parallel action sequences: $\mathbf{u} = ((a_1, Z_1), \dots, (a_\ell, Z_\ell), \dots), \mathbf{u}' = ((a_1', Z_1'), \dots, (a_\ell', Z_\ell'), \dots)$. Here $Z_k' = Z_k$, for $\forall k$. The scheduling action a_ℓ is:

$$
a_\ell = \arg\min_{m \in \{1,2\}} \Delta_m(D_\ell),
\tag{34}
$$

while the action a_k' only differs a_k at the epoch ℓ[2]:

$$
a_k' = \begin{cases} a_k & \text{if } k \neq \ell, \\ \arg\max_m \Delta_m(D_\ell) & \text{if } k = \ell, \end{cases}
\tag{35}
$$

Let the exogenous random delay $\{Y_k\}_{k \geq 1}$ be the same for both systems. Denote the resulting state trajectories by $\{\Gamma_k, M_k\}_{k \geq 1}$ and $\{\Gamma_k', M_k'\}_{k \geq 1}$, respectively. Because M_k

---

[2]As $\Delta_m(D_\ell) = D_\ell - S_{\ell,m}$, minimizing(maximizing) $\Delta_m(D_\ell)$ is equivalent to maximizing(minimizing) $S_{\ell,m}$.

evolves regardless of which sensor is scheduled, and the two sequences share the same $\{Y_k, Z_k\}$, from (30) we have

$$M_k = M'_k, \qquad \forall k \in \mathbb{N}. \tag{36}$$

Then, we apply *induction* in Appendix B to establish:

$$\Gamma'_k \leq \Gamma_k, \qquad \forall k \in \mathbb{N}. \tag{37}$$

Given the strict monotonicity of the cost function $c^\rho_\lambda(\Gamma, M, Y, u)$ in Γ (2), it immediately follows that:

$$c^\rho_\lambda(\Gamma'_k, M'_k, Y_k, (a'_k, Z'_k)) \leq c^\rho_\lambda(\Gamma_k, M_k, Y_k, (a_k, Z_k)), \quad \forall k. \tag{38}$$

Consequently, deviating from the MAF policy at any epoch results in a higher total cost, confirming that consistently scheduling the sensor with the maximum age is optimal. ∎

Theorem 2 reveals a key simplification: *scheduling and sampling can be decoupled*, with the optimal scheduling given by the intuitive MAF rule. Consequently, joint policy optimization reduces to designing the sampling policy under fixed g_{MMSE} and π_{MAF}, as addressed in the next subsection.

C. Optimal Sampling Policy

Given the MAF scheduling, the *age gap* dynamics in (31) becomes *recurrent*: $\Gamma_{i+1} = Y_i + Z_i$. However, the state M_i remains *transient* and *unbounded*. To address this, we here eliminate the *tricky* state M_i and formulate a new MDP:

MDP 3 (M_i *State Elimination Given the MAF Scheduling*). *This MDP is expressed by four components:*

- **State:** $x_i \triangleq (\Gamma_i, Y_i) \in \mathbb{R}^2_+$.
- **Action:** $Z_i \in \mathbb{R}_+$.
- **One–step cost:**

$$\begin{aligned}\psi_\lambda(\Gamma_i, Y_i, u_i) = {} & \sigma_Y + \mu_Y Z_i - \lambda\mu_Y - \lambda Z_i \\ & + (Z_i + \mu_Y)(2Y_i + Z_i + \Gamma_i).\end{aligned} \tag{39}$$

- **State transition:** Γ_i *evolves as:*

$$\Gamma_{i+1} = Y_i + Z_i, \tag{40}$$

while Y_{i+1} is independent of Y_i.

The following theorem establishes that, under the *average cost criterion*, it suffices to solve MDP 3 rather than MDP 2 when the MAF scheduling policy is employed.

Theorem 3 (*Average-cost MDP Equivalence*). *Given the MAF scheduling policy, for any sampling policy $(Z_0, Z_1, \ldots)$, the long-term average costs corresponding to $c^\rho_\lambda(\cdot)$ in MDP 2 and $\psi_\lambda(\cdot)$ in MDP 3 coincide exactly. Formally,*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}[\psi_\lambda(\Gamma_i, Y_i, u_i) - c^\rho_\lambda(\Gamma_i, M_i, Y_i, u_i)] = 0. \tag{41}$$

Proof. See Appendix C. ∎

With this transformation, the previously intractable MDP with *transient* and *unbounded* states becomes analytically tractable. Since Γ_i is *recurrent*, standard solution methods,

such as the RVI algorithm, can be applied directly to solve this problem.
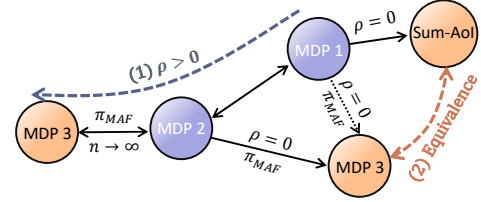
D. Age-Optimal Sampling is All you Need

Fig. 2. Relationship between different MDP formulations.

An insightful special case arises when $\rho = 0$. In this scenario, the function $H_0(M, \Gamma)$ reduces to Γ, which in turn simplifies the one-step cost to $\psi_\lambda(\Gamma, Y, u) = c^0_\lambda(\Gamma, M, Y, u)$. As a result, MDP 2 becomes fully equivalent to MDP 3 under the MAF scheduling policy when $\rho = 0$. On the other hand, it has been established in (18) that the MDP 1 is a classical sum-AoI optimization problem when $\rho = 0$. Combining these two observations leads to conclusion (2) in Fig. 2.

In contrast, for $\rho > 0$, the solution to MDP 1 remains identical to that of MDP 3 under the MAF scheduling policy, as indicated by conclusion (1) in Fig. 2. These structural equivalences lead to the following corollary.

Corollary 1 (*Age-Optimal Sampling is MSE-Optimal in the Infinite Horizon*). *Given the MAF scheduling, the MSE-optimal sampling is asymptotically equivalent to AoI-optimal sampling in the infinite horizon.*

This result allows us to directly leverage age-optimal sampling policies in our estimation framework. In particular, we can adopt the low-complexity threshold-triggered water-filling sampling (WF) policy in [11, Eq. 44], given by:

$$Z_i = \max\left\{0, T - \frac{\Delta_1(D_i) + \Delta_2(D_i)}{2}\right\}, \forall i \in \mathbb{N}, \tag{42}$$

where the threshold T can be obtained by golden-section methods [35]. This policy is shown to achieve near-optimal sum-AoI performance for multiple-sources scenarios.

IV. SIMULATION RESULTS

We evaluate the performance of the proposed joint sampling and scheduling policy under correlated Wiener processes. Unless otherwise specified, the correlation level is fixed at $\rho = 0.9$, and MMSE fusion estimation is implemented at the receiver. As a case study, the transmission delay is modeled as a binary random variable: 0 with probability p, and $Y_{\max}$ with probability $1 - p$.

Benchmark Policies:
- **Zero-Wait (ZW) Sampling:** Each source samples right after the previous transmission completes, i.e., $Z_i = 0$.
- **Constant-Wait Sampling:** A fixed delay $Z_i = d$ is imposed between transmissions.
- **Random (RAND) Scheduling:** Each sensor is selected at random, i.e., $\Pr(a_i = m) = \frac{1}{2}$ for all $m \in \{1, 2\}$.
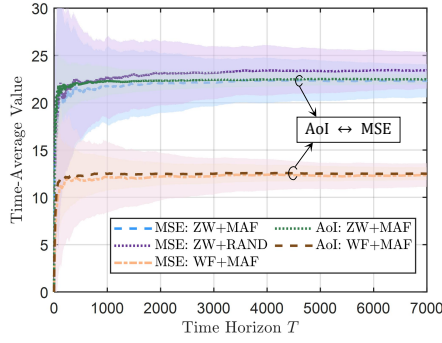
Fig. 3. Finite-horizon time-average MSE and AoI under different scheduling and sampling policies. Here $Y_{\max} = 20$ and $p = 0.95$.

Our proposed MAF scheduling (33) plus Water-filling sampling (42) are jointly referred to by WF+MAF.
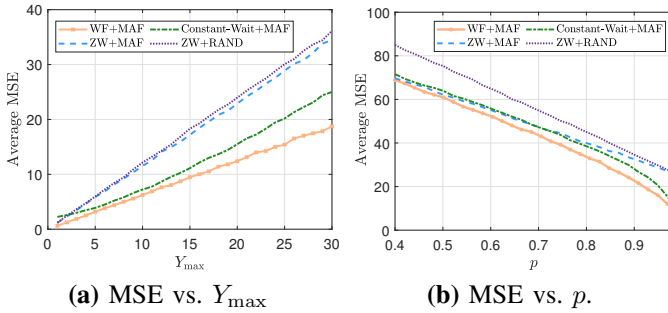

(a) MSE vs. $Y_{\max}$ **(b)** MSE vs. p.

Fig. 4. Long-term average MSE and AoI under different scheduling and sampling policies. We set $d = 1$ for constant-wait sampling. (a) $p = 0.95$; (b) $Y_{\max} = 25$.

Fig. 3 shows that time-average MSE and AoI converge as the horizon T increases, validating the *asymptotic equivalence* in Corollary 1. Notably, AoI serves as a provable upper bound for MSE, as implied by (17), and the gap between the two metrics vanishes as $T \to \infty$.

Fig. 4 compares long-term MSE across policy combinations. WF+MAF consistently achieves the lowest MSE. In Fig. (a), its performance degrades slowly with increasing $Y_{\max}$, unlike ZW-based strategies which exhibit sharp deterioration. In Fig. (b), WF+MAF remains robust across the parameter p. These results demonstrate that the proposed joint sampling and scheduling policy offers strong performance and robustness under various channel delay distributions.

V. CONCLUSION

In this paper, we studied a remote information fusion problem in which a remote fusion center aims to track a multi-dimensional *correlated* Wiener process based on status updates from multiple distributed sources communicating over a shared channel with random delays. We explicitly solved the joint optimization of sampling, scheduling, and estimation policies, and derived a structural characterization of the optimal solution that is both theoretically optimal and practically implementable. Notably, we established an asymptotic equivalence between the time-averaged AoI and MSE under the optimal MAF scheduling and MMSE estimation, even in

the presence of inter-source correlation. Future work includes extending our framework to larger-scale systems with more sensors, and exploring other classes of Markovian sources beyond the correlated Wiener process.

REFERENCES

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

[2] 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, 2019.

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

[4] H. Tang, Y. Sun, and L. Tassiulas, "Sampling of the Wiener process for remote estimation over a channel with unknown delay statistics," *IEEE/ACM Trans. Netw.*, vol. 32, no. 3, pp. 1920–1935, 2024.

[5] Y. Chen, H. Tang, J. Wang, P. Yang, and L. Tassiulas, "Sampling for remote estimation of an Ornstein-Uhlenbeck process through channel with unknown delay statistics," *J. Commun. Networks*, vol. 25, no. 5, pp. 670–687, 2023.

[6] X. Chen, A. Li, and S. Wu, "Optimal sampling for Uncertainty-of-In formation minimization in a remote monitoring system," in *Proc. IEEE ITW*, 2024.

[7] M. K. C. Shisher, Y. Sun, and I.-H. Hou, "Timely communications for remote inference," *IEEE/ACM Trans. Netw.*, 2024.

[8] M. K. C. Shisher and Y. Sun, "How does data freshness affect real-time supervised learning?" in *Proc. ACM Mobihoc*, 2022, pp. 31–40.

[9] 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 Commun.*, 2023.

[10] 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 ISIT*, 2024, pp. 1179–1184.

[11] A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, "Optimal sampling and scheduling for timely status updates in multi-source networks," *IEEE Trans. Inf. Theory*, vol. 67, no. 6, pp. 4019–4034, 2021.

[12] 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. ACM MobiHoc*. ACM, 2023, pp. 91–100.

[13] K. Banawan, A. Arafa, and K. G. Seddik, "Timely multi-process estimation over erasure channels with and without feedback: Signal-independent policies," *IEEE J. Sel. Areas Inf. Theory*, vol. 4, pp. 607–623, 2023.

[14] A. Li, S. Wu, S. Meng, R. Lu, S. Sun, and Q. Zhang, "Toward goal-oriented semantic communications: New metrics, framework, and open challenges," *IEEE Wirel. Commun.*, vol. 31, no. 5, pp. 238–245, 2024.

[15] E. Uysal, O. Kaya, A. Ephremides, J. Gross, M. Codreanu, P. Popovski, M. Assaad, G. Liva, A. Munari, B. Soret, T. Soleymani, and K. H. Johansson, "Semantic communications in networked systems: A data significance perspective," *IEEE Netw.*, vol. 36, no. 4, pp. 233–240, 2022.

[16] Q. He, G. Dán, and V. Fodor, "Joint assignment and scheduling for minimizing Age of Correlated Information," *IEEE/ACM Trans. Netw.*, vol. 27, no. 5, pp. 1887–1900, 2019.

[17] C. Xu, X. Zhang, H. H. Yang, X. Wang, N. Pappas, D. Niyato, and T. Q. S. Quek, "Optimal status updates for minimizing Age of Correlated Information in IoT networks with energy harvesting sensors," *IEEE Trans. Mob. Comput.*, vol. 23, no. 6, pp. 6848–6864, 2024.

[18] B. Zhou and W. Saad, "On the Age of Information in Internet of Things systems with correlated devices," in *Proc. IEEE GLOBECOM*. IEEE, 2020, pp. 1–6.

[19] L. Wang, J. Sun, Y. Sun, S. Zhou, Z. Niu, M. Jiang, and L. Geng, "Grouping-based cyclic scheduling under age of correlated information constraints," *IEEE Transa. Inf. Theory*, vol. 71, no. 3, pp. 2218–2244, 2025.

[20] L. Crosara and L. Badia, "Strategic age of information under different correlation of sources," in *Proc. IEEE WCNC*, 2025, pp. 1–6.

[21] L. M. Hoang, J. Doncel, and M. Assaad, "Age-oriented scheduling of correlated sources in multi-server system," in *Proc. IEEE ISWCS*. IEEE, 2021, pp. 1–6.

[22] V. Tripathi and E. Modiano, "Optimizing Age of Information with correlated sources," in *Proc. ACM Mobihoc*, 2022, pp. 41–50.

[23] A. E. Kalør and P. Popovski, "Timely monitoring of dynamic sources with observations from multiple wireless sensors," *IEEE/ACM Trans. Netw.*, vol. 31, no. 3, pp. 1263–1276, 2022.

[24] J. Tong, L. Fu, and Z. Han, "Age-of-Information oriented scheduling for multichannel iot systems with correlated sources," *IEEE Trans. Wirel. Commun.*, vol. 21, no. 11, pp. 9775–9790, 2022.

[25] R. V. Ramakanth, V. Tripathi, and E. Modiano, "Monitoring correlated sources: AoI-based scheduling is nearly optimal," *IEEE Trans. Mob. Comput.*, vol. 24, no. 2, pp. 1043–1054, 2025.

[26] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.

[27] J. Pan, A. M. Bedewy, Y. Sun, and N. B. Shroff, "Optimal sampling for data freshness: Unreliable transmissions with random two-way delay," *IEEE/ACM Trans. Netw.*, vol. 31, no. 1, pp. 408–420, 2023.

[28] A. Li, S. Wu, G. C. F. Lee, X. Chen, and S. Sun, "Sampling to achieve the goal: An age-aware remote markov decision process," in *Proc. IEEE ITW*, 2024, pp. 121–126.

[29] S. Leng and A. Yener, "Age of Information minimization for an energy harvesting cognitive radio," *IEEE Trans. Cogn. Commun. Netw.*, vol. 5, no. 2, pp. 427–439, 2019.

[30] Q. Wang, H. Chen, and Z. Dong, "Optimizing information freshness in uplink multiuser simo systems: Low-complexity scheduling algorithms," *IEEE Trans. Commun.*, pp. 1–1, 2024.

[31] H. V. Poor, *An introduction to signal detection and estimation.* Springer Science & Business Media, 2013.

[32] W. Dinkelbach, "On nonlinear fractional programming," *Management science*, vol. 13, no. 7, pp. 492–498, 1967.

[33] R. A. Howard, *Dynamic programming and markov processes.* Cambridge, MA; MIT Press, 1960.

[34] D. J. White, "Dynamic programming, markov chains, and the method of successive approximations," *J. Math. Anal. Appl*, vol. 6, no. 3, pp. 373–376, 1963.

[35] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Golden section search in one dimension," *Numerical Recipes in C: The Art of Scientific Computing*, p. 2, 1992.

[36] T. W. Anderson, T. W. Anderson, T. W. Anderson, T. W. Anderson, and E.-U. Mathématicien, *An introduction to multivariate statistical analysis.* Wiley New York, 1958, vol. 2.

APPENDIX A
PROOF OF THEOREM 1

Consider the following zero-mean Gaussian vector:

$$\mathbf{Z} \triangleq \big(X, \underbrace{Y_1, Y_2}_{Y}\big)^\top = \big(W_t^{(1)}, W_{t_1}^{(1)}, W_{t_2}^{(2)}\big)^\top. \quad (43)$$

Its block covariance matrix, obtained from the coupled Wiener process construction, is:

$$\Sigma = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix} = \begin{pmatrix} t & [t_1 \ \rho t_2] \\ [t_1 \ \rho t_2]^\top & \begin{pmatrix} t_1 & \rho t_{\min} \\ \rho t_{\min} & t_2 \end{pmatrix} \end{pmatrix}, \quad (44)$$

where $t_{\min} = \min\{t_1, t_2\}$. Because X, Y_1, Y_2 are jointly Gaussian random variables, one can calculate the following conditional expectation [36, Chap. 2]:

$$\mathbb{E}[X \mid Y_1, Y_2] = \Sigma_{XY} \Sigma_{YY}^{-1} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}. \quad (45)$$

A direct inversion of the matrix Σ_{YY} yields:

$$\Sigma_{YY}^{-1} = \frac{1}{\Delta} \begin{pmatrix} t_2 & -\rho t_{\min} \\ -\rho t_{\min} & t_1 \end{pmatrix}, \quad \Delta = t_1 t_2 - \rho^2 t_{\min}^2, \quad (46)$$

and therefore

$$\Sigma_{XY} \Sigma_{YY}^{-1} = \frac{1}{\Delta} \big[t_2(t_1 - \rho^2 t_{\min}), \ \rho t_1(t_2 - t_{\min}) \big]. \quad (47)$$

Substituting $X = W_t^{(1)}$, $Y_1 = W_{t_1}^{(1)}$ and $Y_2 = W_{t_2}^{(2)}$ into (45)

gives the conditional expectation. Setting $t_1 = t - \Delta_1(t)$ and $t_2 = t - \Delta_2(t)$ immediately yields (11). Symmetrically, we can also establish (12).

APPENDIX B
PROOF OF (37)

For $k \leq \ell$, since $a_k = a'_k$ for all $k \leq \ell$, by the update rule (31), we have $\Gamma_k = \Gamma'_k$ for these steps.

For $k = \ell + 1$, at step $\ell + 1$, the actions differ, so

$$\Gamma_{\ell+1} = \Gamma_\ell + Y_\ell + Z_\ell, \qquad \Gamma'_{\ell+1} = Y_\ell + Z_\ell. \quad (48)$$

which gives $\Gamma_{\ell+1} > \Gamma'_{\ell+1}$ since $\Gamma_\ell > 0$.

For $k \geq \ell+2$, We use induction. Suppose for some $n \geq \ell+1$ we have $\Gamma_n \geq \Gamma'_n$. Since $a_n = a'_n$ for $n > \ell$, we have:

$$\Gamma_{n+1} - \Gamma'_{n+1} = \begin{cases} \Gamma_n - \Gamma'_n, & \text{if } a_n = \arg\max_j\{S_{n,j}\}, \\ 0 & \text{if } a_n = \arg\min_j\{S_{n,j}\}. \end{cases} \quad (49)$$

Both cases preserve the inequality $\Gamma_{n+1} \geq \Gamma'_{n+1}$. Thus, by induction, $\Gamma_k \geq \Gamma'_k$ for all k. ∎

APPENDIX C
PROOF OF THEOREM 3

From the update rule (30), we have

$$M_n = M_0 + \sum_{i=1}^n (Y_{i-1} + Z_{i-1}). \quad (50)$$

By the *Kolmogorov strong large number law*,

$$\lim_{n\to\infty} \frac{M_n}{n} > \lim_{n\to\infty} \frac{M_0}{n} + \frac{\sum_{i=1}^n Y_{i-1}}{n} \xrightarrow{\text{a.s.}} \mu_Y. \quad (51)$$

Therefore, there exists k_0 such that for all $i \geq k_0$,

$$M_i > \frac{\mu_Y}{2} i, \qquad \forall i \geq k_0. \quad (52)$$

Denote $y^{(u)}$ and $z^{(u)}$ as the maximum value of Y_i and Z_i, respectively. From (40), we obtain that:

$$\Gamma_i^2 = (Y_{i-1} + Z_{i-1})^2 \leq 2(y^{(u)})^2 + 2(z^{(u)})^2. \quad (53)$$

Substitute the above bounds into the left-hand side of (41):

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}[\psi_\lambda(\Gamma_i, Y_i, u_i) - c_\lambda^\rho(\Gamma_i, M_i, Y_i, u_i)]$$

$$= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} (Z_i + \mu_Y) \left(\frac{\rho^2 \Gamma_i^2}{\Gamma_i + (1-\rho^2)M_i} \right)$$

$$< \rho^2(z^{(u)} + \mu_Y) \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{k_0-1} \frac{\Gamma_i^2}{(1-\rho^2)M_i} + \sum_{i=k_0}^n \frac{\Gamma_i^2}{(1-\rho^2)M_i}$$

$$< \frac{2\rho^2(z^{(u)} + \mu_Y)}{\mu_Y(1-\rho^2)} \lim_{n\to\infty} \frac{1}{n} \sum_{i=k_0}^n \frac{\Gamma_i^2}{i}$$

$$\leq \frac{4\rho^2(z^{(u)} + \mu_Y)((y^{(u)})^2 + (z^{(u)})^2)}{\mu_Y(1-\rho^2)} \lim_{n\to\infty} \frac{1}{n} \sum_{i=k_0}^n \frac{1}{i} = 0,$$

$$(54)$$

where the last equality follows since $\frac{1}{n} \sum_{i=k_0}^n \frac{1}{i} \to 0$ as $n \to \infty$. By the squeeze theorem, we complete the proof. ∎