Goal-oriented Communications for Multimodal Remote Inference: A Correlated Two-Modality Case

Keyuan Zhang
Department of CS
Virginia Tech
Blacksburg, VA, USA
keyuanz@vt.edu

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

Bo Ji
Department of CS
Virginia Tech
Blacksburg, VA, USA
boji@vt.edu

***Abstract*—In this paper, we study a multimodal remote inference system, where a multimodal machine learning (ML) model at the receiver performs real-time inference tasks (e.g., estimating a robot's pose) using time-sensitive data from remote sources (e.g., video streams, robotic states). The expected inference quality of the model is characterized as a joint function of the Age of Information (AoI) for each modality. Due to the bandwidth constraints, a transmitter must decide which modality to transmit when the channel is available. Focusing on the two-modality case, we formulate a modality selection problem aimed at minimizing the inference error of the ML model. The problem introduces two key challenges: (i) the two sources are coupled, as the objective function depends jointly on the AoI vector and may be non-additive and non-monotonic; and (ii) the modalities have different data sizes, leading to heterogeneous transmission times. We model the problem as an infinite-horizon average-cost Semi-Markov Decision Process (SMDP), from which we further derive a low-complexity index-based threshold policy and prove its optimality theoretically. Our numerical results show that by jointly considering inference error and the AoI of both modalities, the proposed policy effectively reduces inference error by up to 55% compared to baseline methods.**

***Index Terms*—Scheduling; Age of Information; Remote Inference; Multimodal**

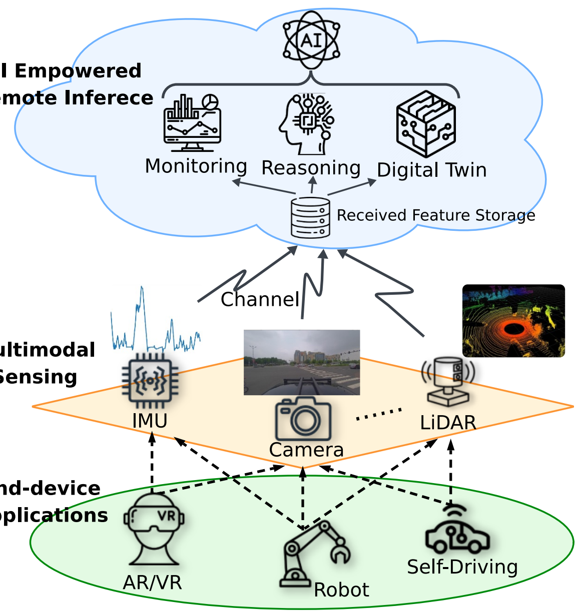


Fig. 1: Illustration of multimodal remote inference system.

I. Introduction

The advent of sixth-generation (6G) technology opens up vast opportunities for intelligent systems—such as autonomous transportation, unmanned mobility, industrial automation, and extended reality—by enabling reliable connectivity between distributed sensing nodes and remote processing units. Fig. 1 illustrates a brief overview of such systems. Machine learning (ML) models are often deployed at remote destinations to process received features for complex real-time tasks such as monitoring, reasoning, or decision-making, and have demonstrated promising performance. However, timely delivery of input data is essential for these models to produce accurate estimations and informed decisions in dynamic environments. Despite this, most communication studies continue to focus on optimizing network-level metrics such as throughput, delay, or data freshness, without explicitly accounting for the performance of the downstream task. Recent work [1] has shown that optimizing these metrics may not be sufficient: both theoretical and empirical results demonstrate that the freshness of correlated features can have a complicated impact on the performance of ML models.

To fill this gap, the concept of goal-oriented communication has recently been proposed [2]. Rather than optimizing network metrics, goal-oriented communication aims to maximize downstream task performance. While this concept aligns communication more closely with task objectives, it also introduces new challenges in communication system design. A key challenge is that the task performance is often difficult to characterize, especially for some black-box ML models such as deep neural networks. As a result, developing scheduling

This work was supported in part by the National Science Foundation (NSF) under Grant CNS-223967.

policies for goal-oriented communication is non-trivial.

Moreover, as depicted in Fig. 1, many complex tasks often involve multiple *data modalities*, such as visual, auditory, spatial (e.g., LiDAR), and motion (e.g., IMU) data. Typically, each modality provides *complementary* information, enhancing the overall task performance. For instance, to track the location of vehicles and pedestrians, color images (RGB) provide information about the shape and appearance of objects, while LiDAR offers precise depth information [3]. Similarly, for teleoperated robotics, tactile sensors provide additional information about the surface properties of objects [4]. Meanwhile, many multimodal ML models have been proposed to fully leverage information from these modalities. However, limited bandwidth and other practical network constraints often prevent the model from consistently accessing fresh data from all modalities simultaneously. This leads to a natural research question: *How can we wisely schedule modality transmissions to manage data freshness and ultimately optimize ML model performance under limited bandwidth resources?*

To that end, we study a modality scheduling problem for a time-slotted multimodal remote inference problem: at each time slot, the scheduler can only select one modality from multiple sensors to a remote receiver. In this paper, we focus on the two-modality case, which is common in many multimodal learning applications (e.g., vision and audio, RGB and LiDAR), yet remains underexplored. To account for differing feature sizes, we also allow transmission times to vary across modalities. On the receiver side, a pre-trained ML model takes the most recently delivered data from all modalities as inputs and infers the value of a target signal. To measure the inference performance, we adopt the approach of recent work [5]–[8] and model the expected inference error at each time slot as a joint function of the Age of Information (AoI) [9] of the two modalities (see Section II-B for a detailed definition). The benefit of this approach is that, instead of directly optimizing an intractable inference performance function, we minimize an AoI-based penalty function, which has been extensively studied and is closely aligned with real-time applications. Nonetheless, the problem remains challenging because the AoI-based penalty function may be non-monotonic and non-additive, due to the inherent complexity of the ML model. Our goal is to find the modality scheduling policy that minimizes the time-averaged inference error of the ML model.

For multi-source scheduling under bandwidth constraints, a common approach is to seek asymptotically optimal solutions by first decomposing the problem, often following paradigms like Whittle index [10] or Net-gain Maximization [6]. Such approaches work well when the objective function for each individual source is complex, but the overall objective is additively separable across sources. However, in our setting, the inference error may not be additive, and thus the two modalities cannot be decomposed as before. Surprisingly, we still derive an efficient and optimal policy. The key idea of our approach is to formulate the problem as a Semi-Markov Decision Process (SMDP) [11] by treating consecutive transmissions for each modality as actions and the AoI at each time slot as the state, and then directly solving its Bellman optimality equations. Moreover, an optimal scheduling policy for two-modality scheduling was known only in the case of monotonic objective functions [12]. We extend this result to the non-monotonic case, offering deeper insight into the conditions for optimal policy existence. We summarize our contribution as follows:

- We formulate a two-modality scheduling problem for the multimodal remote inference system. *To the best of our knowledge, this is the first work to study transmission scheduling for two modalities with the goal of minimizing inference error.* The key challenge in this problem lies in the fact that the inference error is modeled as a general function of the AoI from two modalities, which may be non-monotonic, non-additive, and non-convex. This generality makes the problem more difficult and renders existing scheduling policies, which are designed for monotonic and additive AoI penalty functions, suboptimal in our setting. Our model also incorporates heterogeneous transmission times across modalities, capturing practical conditions in real-world applications.
- We propose an optimal policy for the formulated problem. The policy has a simple cyclic structure: The transmitter switches cyclically between two modalities. Each modality m is consecutively selected by the transmitter for θ transmissions, until its index value $\gamma^{(m)}(\theta + 1)$ exceeds a threshold. Notably, we derive a *closed-form expression* for the index function of each modality, and both modalities share *a common threshold* that can be *efficiently computed*. A similar approach has proven effective in handling non-monotonic inference error in single-source scenarios, but whether it can be extended to the scheduling of two modalities has remained unclear.
- We conduct numerical experiments using a robot state prediction task to evaluate the proposed algorithm. Results show that it reduces inference error by up to 55% compared to baseline methods.

A. Related Work

AoI penalty function. The Age of Information (AoI) has been widely studied in a large body of work (see eg., [1], [5]–[7], [9], [13]–[23] and a recent survey [24]). Over time, its concept has progressively become more general, encompassing a wider range of applications. Early research efforts primarily focused on analyzing and optimizing non-decreasing linear AoI, such as average AoI and peak AoI [9], [13], [17], [21]. Later, as surveyed in [14], several non-decreasing non-linear utility functions of AoI were suggested to characterize data freshness for various applications. Meanwhile, many studies also considered various AoI-related metrics, such as Age of Incorrect Information (AoII) [25], Age of Outdated Information (Ao^{2}I) [26], and Uncertainty of Information [27].

More recently, it was shown in [1] that the impact of AoI from multiple correlated features on inference error can be non-monotonic A similar conclusion was drawn for a Gaussian autoregressive AR(p) process [15]. Since then, a line of research has investigated scheduling under non-monotonic AoI loss functions [5]–[8], [28], [29] (see the next paragraph for more details). On the other hand, AoI scheduling for correlated sources has also been studied in [30]–[33], where the AoI penalty function remains non-decreasing, but the sources are coupled due to correlation. However, scheduling remains underexplored when the inference error is a joint function, potentially non-monotonic and non-additive, of the AoI from multiple coupled sources. This paper partially fills this gap by characterizing an optimal policy for the two-modality case under such conditions.

Scheduling for remote inference systems. Recently, a line of research has focused on scheduling for the remote inference problem, where the inference error is a non-monotonic function of AoI. In the seminal work [5], [7], transmission scheduling policies were proposed to minimize inference error under both single-source and multi-source settings. Subsequently, [6] studied the case where feature lengths are time-varying and proposed a joint feature length selection and transmission scheduling policy. Since then, follow-up work has addressed other settings, such as two-way delay channels [8] and multi-task scenarios [8], [29]. It has also extended to related scenarios, including remote state estimation [34] and channel estimation [28]. However, these results do not directly extend to our two-modality case, as they either assume a single source or weakly coupled multiple sources—that is, they assume that the inference error is an additive function of the AoIs from different sources.

Remote state estimation. This work also relates to the field of remote state estimation [16], [18], [32], where the scheduler needs to select a subset of sources to update their state and then estimate the current state of the system based on all previously received status updates. Our model distinguishes from the remote state estimation literature as follows: (i) The additive objective function is also a common assumption in remote state estimation. (ii) In remote state estimation, sources are often assumed to follow Gaussian and Markovian processes [18], [32], whereas our model applies to more general processes.

II. System Model and Problem Formulation

A. System Model

We consider a two-modality remote inference system depicted in Fig. 2, where a transmitter selects features from two modality sources and forwards them to a receiver through a communication channel. Let the time be slotted, with time slot index $t \in \{0, \ldots, T\}$. At each time slot t, each modality source $m \in \{1, 2\}$ regularly samples one feature, denoted as $X_t^{(m)} \in \mathcal{X}^{(m)}$. The feature can either be a collection of video frames or a sequence of robotic state signals within the time slot, depending on the type of sensor used. Meanwhile, on the receiver side, a predictor (e.g., a pre-trained ML model) uses the received features from the two modality sources to infer the real-time value of a target signal $Y_t \in \mathcal{Y}$ of interest.

The source and the receiver are connected via one communication channel. In many multi-modal remote inference

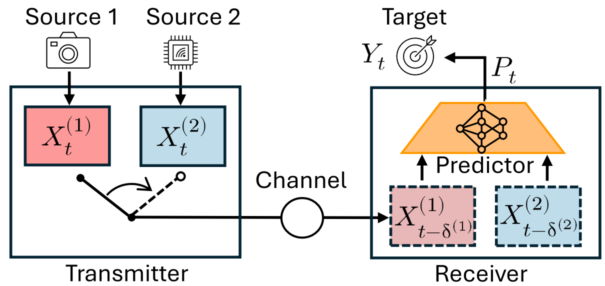


Fig. 2: Overview of system model.

applications, modality sources are typically co-located to collect information from a single object, and therefore share the same communication channel. Given the bandwidth constraint, we assume that only one modality can be transmitted at a time over the shared channel. When the channel is free, the transmitter can select one modality m and start transmitting its freshest feature, i.e., $X_t^{(m)}$. We assume that the channel is reliable, so the receiver will successfully receive the transmitted feature after a few time slots. Since the feature size of each modality can vary, let $T^{(m)}$ denote the transmission time for one feature from modality m. We assume that $T^{(m)}$ is a known constant, but may differ across modalities.

Furthermore, we do not consider idling or preemption in this paper, meaning the transmitter will not initiate a new transmission while the channel is busy and will begin the next transmission immediately after the previous feature is received. Formally, suppose the $(n-1)$-th transmitted feature is delivered at the beginning of time slot D_{n-1}. Then, the transmitter selects the next modality to transmit at the same time slot. Let $a_n \in \{1, 2\}$ denote the binary decision variable indicating the modality of the n-th feature: $a_n = 1$ if modality 1 is selected, and $a_n = 2$ if modality 2 is selected. The next delivery time D_n is then determined by

$$D_n = D_{n-1} + T^{(m)}, \quad \text{if } a_n = m, \ m \in \{1, 2\}. \qquad (1)$$

Our goal is to find the optimal modality scheduling policy that minimizes the predictor's inference error under the given bandwidth constraint. Next, we formally define the metric used to quantify the inference error and introduce the structure of the scheduling policy under consideration.

B. Inference Error: a function of AoI

We aim to measure the predictor's inference performance based on the freshness of both modalities. To achieve this, we first use the Age of Information (AoI) to quantify feature freshness on the receiver side [9], defined as the difference between the current time and the generation time of the most recently received feature. Let $\Delta^{(m)}(t) \in \mathbb{Z}_+$ denote the AoI of modality m on the receiver side at time slot t. For $D_n \leq t < D_{n+1}$, given the scheduler's decisions a_{n+1}, the AoI $\Delta^{(m)}(t)$ is represented by

$$\Delta^{(m)}(t) := t - \max_{j \leq n} \{D_{j-1} : a_j = m\}, \qquad (2)$$

where D_{j-1} is the generation time of the j-th feature. As we are primarily interested in the AoI at the decision time slot D_n, the following lemma provides a more direct characterization of the AoI evolution.

Lemma 1. *Let $m \in \{1,2\}$ denote one modality, and define $m' = \{1,2\}\backslash\{m\}$ as the complementary modality. Then, for each n-th transmitted feature, we have*

$$\Delta^{(m)}(D_n) = \begin{cases} T^{(m)}, & \textit{if } a_n = m \\ \Delta^{(m)}(D_{n-1}) + T^{(m')}, & \textit{otherwise} \end{cases} \tag{3}$$

Proof: Eq. (3) follows directly from Eq. (1) and Eq. (2). If $a_n = m$, then by Eq. (2) we have $\Delta^{(m)}(D_n) = D_n - D_{n-1}$, and by Eq. (1), this simplifies to $\Delta^{(m)}(D_n) = T^{(m)}$. Similarly, when $a_n = m'$, we have $\Delta^{(m)}(D_n) - \Delta^{(m)}(D_{n-1}) = T^{(m')}$. ■

Lemma 1 shows that upon delivery of the n-th feature, the AoI of modality m drops to $T^{(m)}$ if $a_n = m$; otherwise, it continues to increase. During any transmission, the AoI of both modalities increases by one per slot.

We next model the predictor's inference performance at each time slot based on the AoI. Specifically, let $\mathbf{X}_{t-\mathbf{\Delta}(t)} \in \mathcal{X}$ denote the feature vector associated with the AoI vector $\mathbf{\Delta}(t)$ at time slot t, defined as

$$\mathbf{X}_{t-\mathbf{\Delta}(t)} := \left(X^{(1)}_{t-\Delta^{(1)}(t)}, X^{(2)}_{t-\Delta^{(2)}(t)}\right),$$

where $\mathbf{\Delta}(t) := \left(\Delta^{(1)}(t), \Delta^{(2)}(t)\right)$ denote the AoI vector at time slot t. Then, the predictor can be represented as a function $\phi : \mathcal{X} \times \mathbb{Z}_+^2 \mapsto \mathcal{P}$, which takes the latest received feature vector $\mathbf{X}_{t-\mathbf{\Delta}(t)}$ along with the corresponding AoI vector $\mathbf{\Delta}(t)$ as inputs and produces a prediction $P_t \in \mathcal{P}$. The prediction performance is evaluated using a real-valued loss function $\ell : \mathcal{Y} \times \mathcal{P} \mapsto \mathbb{R}$, where $\ell(y, p)$ denotes the value of inference error at time slot t when the target is $Y_t = y$ and the prediction is $P_t = p$. Note that prediction space $\mathcal{P}$ may differ from the target space $\mathcal{Y}$, as the task may not be predicting the exact real-time value of the target Y_t. Depending on the goal of applications, the choice of loss function ℓ is quite flexible, e.g., 0-1 loss, quadratic loss, and logarithmic loss.

In this paper, we focus on the predictor's expected inference error at each time slot. We assume that the process $\{(Y_t, \mathbf{X}_t), t = 0, 1, 2, \dots\}$ is stationary, and that the processes $\{(Y_t, \mathbf{X}_t), t = 0, 1, 2, \dots\}$ and $\{\mathbf{\Delta}(t), t = 0, 1, 2, \dots\}$ are independent. Under these assumptions, the expected inference error can be expressed as a function of the AoI vector $\mathbf{\Delta}(t)$ [5]. That is, if $\mathbf{\Delta}(t) = \boldsymbol{\delta}$, we have

$$L(\boldsymbol{\delta}) := \mathbb{E}_{Y,\mathbf{X} \sim \mathbb{P}(Y_t, \mathbf{X}_{t-\boldsymbol{\delta}})}\left[\ell(Y, \phi(\mathbf{X}, \boldsymbol{\delta}))\right], \tag{4}$$

where $\mathbb{P}(Y_t, \mathbf{X}_{t-\boldsymbol{\delta}})$ denotes the joint distribution of the target Y_t and the feature $\mathbf{X}_{t-\boldsymbol{\delta}}$. In this paper, we require the expected inference error to be uniformly bounded, as stated below:

Assumption 1 (Uniformly Bounded)**.** *There exists a finite constant M such that $|L(\mathbf{\Delta})| \leq M$ for all AoI vectors $\mathbf{\Delta}$.*

Assumption 1 is essential for ensuring the existence of an optimal policy in our following analysis and is reasonable in practice because (i) extremely large AoI values are rare in most real-time applications, and (ii) many ML models apply preprocessing techniques such as cropping and normalization to both inputs and outputs, which helps keep the inference error within a bounded range.

C. Structure of the Scheduling Policy

The scheduling policy π can be represented as a sequence of modality choices, $\pi := (a_1, a_2, \dots)$. In this paper, we consider scheduling policies that satisfy the following three conditions. (i) *Signal-agnostic*: the scheduler has no access to the realization of the process $\{(Y_t, \mathbf{X}_t), t = 0, 1, 2, \dots\}$, implying independence between the processes $\{(Y_t, \mathbf{X}_t), t = 0, 1, 2, \dots\}$ and $\{\mathbf{\Delta}(t), t = 0, 1, 2, \dots\}$. (ii) *Causal*: the modality choice a_n at time slot D_n is determined based on the current and the historical information available at the scheduler (i.e., $\mathbf{\Delta}(0), \mathbf{\Delta}(1), \dots, \mathbf{\Delta}(D_n)$), without knowing the realizations of future information (i.e., $\mathbf{\Delta}(t)$ for $t > D_n$). (iii) The scheduler knows the inference error function $L(\cdot)$. Let Π denote the set of all policies satisfying these conditions.

D. Problem Formulation

We aim to find a scheduling policy in Π that minimizes the summation of the time-averaged expected inference error over an infinite horizon. We define this problem as Problem **OPT**:

$$\textbf{OPT} \quad \bar{L}_{\text{opt}} := \inf_{\pi \in \Pi} \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_\pi \left[\sum_{t=0}^{T-1} L\left(\mathbf{\Delta}(t)\right)\right],$$

where T denotes the horizon, and $\bar{L}_{\text{opt}}$ represents the optimum objective value of **OPT**. The expected inference error is defined in Eq. (4), and the AoI $\Delta^{(m)}(t)$ evolves as in Eq. (2).

III. Optimal Scheduling Policy Design

In this section, we present the optimal policy for **OPT**. We begin by reformulating the problem to introduce a new decision variable, which is key to deriving the optimal policy. With this new decision, the problem can be cast as a Semi-Markov Decision Process (SMDP). Finally, we derive a closed-form expression for the optimal solution by directly solving the Bellman optimality equations.

A. Problem reformulation

In our original formulation, the scheduling policy is represented as a sequence of modality scheduling decisions. We will show that the problem can be reformulated by grouping consecutive transmissions of the same modality as one decision. This reformulation of the decision variables is a crucial step toward deriving the optimal policy.

Specifically, any given sequence of modality scheduling decisions $(a_t)_{t=1,2,\dots}$ can be represented by another sequence

$$\left((\tau_1^{(1)}, \tau_1^{(2)})(\tau_2^{(1)}, \tau_2^{(2)}) \dots (\tau_i^{(1)}, \tau_i^{(2)}) \dots\right), \tag{7}$$

where $\tau_i^{(m)} \in \mathbb{Z}_+$ denotes the number of consecutive transmissions of modality m before switching to the other modality in the i-th cycle; each pair $(\tau_i^{(1)}, \tau_i^{(2)})$ is referred to as one

$$h(T^{(1)}+\tau_{i-1}^{(2)}T^{(2)},T^{(2)}) = \inf_{\tau_i^{(1)}\in\{1,2,\dots\}} \Big[\mathcal{L}^{(1)}(T^{(1)}+\tau_{i-1}^{(2)}T^{(2)},T^{(2)}) + \sum_{j=1}^{\tau_i^{(1)}-1} \mathcal{L}^{(1)}(T^{(1)},jT^{(1)}+T^{(2)}) - \tau_i^{(1)}T^{(1)}\bar{L}_{\text{opt}} + h(T^{(1)},T^{(2)}+\tau_i^{(1)}T^{(1)})\Big], \quad \forall i=1,2,\dots \tag{5}$$

$$h(T^{(1)},\tau_i^{(1)}T^{(1)}+T^{(2)}) = \inf_{\tau_i^{(2)}\in\{1,2,\dots\}} \Big[\mathcal{L}^{(2)}(T^{(1)},\tau_i^{(1)}T^{(1)}+T^{(2)}) + \sum_{j=1}^{\tau_i^{(2)}-1} \mathcal{L}^{(2)}(T^{(1)}+jT^{(2)},T^{(2)}) - \tau_i^{(2)}T^{(2)}\bar{L}_{\text{opt}} + h(T^{(1)}+\tau_i^{(2)}T^{(2)},T^{(2)})\Big], \quad \forall i=1,2,\dots \tag{6}$$

cycle. Rather than selecting a modality for each transmission, the transmitter now determines the number of consecutive transmissions for one modality; once these transmissions are completed, it selects the number of consecutive transmissions for the other modality, and so on. For instance, the sequence $(1,1,1,2,1,1,2)$ can be represented as $((3,1)(2,1))$, meaning that the transmitter selects modality 1 for three consecutive transmissions, then modality 2 for one transmission, followed by two transmissions from modality 1, and finally one transmission from modality 2. Note that although the two policies have different decision times, they are actually *equivalent* for the following two reasons: (i) any sequence $(a_t)_{t=1,2,\dots}$ can be derived from $(\tau_i^{(1)},\tau_i^{(2)})_{i=1,2,\dots}$ and vice versa; (ii) the decisions are made based on the same available information because the transmission times $T^{(1)}$ and $T^{(2)}$ are deterministic. Thus, with a slight abuse of notation, we denote the new policy as $\pi=(\tau_i^{(1)},\tau_i^{(2)})_{i=1,2,\dots}$ and refer to the new policy class as Π in the following.

Now, with the new policy, we cast **OPT** as a Semi-Markov Decision Process (SMDP). In the SMDP, the system state at time t is given by the AoI vector $\boldsymbol{\Delta}(t)$. The transmitter takes an action $\tau_i^{(m)}$, transmitting modality m for $\tau_i^{(m)}$ consecutive times, which takes $\tau_i^{(m)}T^{(m)}$ time slots. Upon completing the transmissions, it proceeds to the next action. In the SMDP, the action is the number of consecutive transmissions $\tau_i^{(m)}$, and the system state at time slot t is represented by the AoI vector $\boldsymbol{\Delta}(t)$. The Bellman optimality equations are provided in Eq. (5) and Eq. (6), while a detailed description of this SMDP and the derivation of the Bellman optimality equations can be found in Appendix A. In Eq. (5) and Eq. (6), we use $h(\Delta^{(1)},\Delta^{(2)})$ to denote the relative value function at state $(\Delta^{(1)},\Delta^{(2)})$, and define $\mathcal{L}^{(m)}(\cdot)$ as follows,

$$\mathcal{L}^{(m)}(\Delta^{(1)},\Delta^{(2)}) = \sum_{t=0}^{T^{(m)}-1} L(\Delta^{(1)}+t,\Delta^{(2)}+t), \tag{8}$$

which represents the total expected inference error incurred when transmitting modality m from state $(\Delta^{(1)},\Delta^{(2)})$ until the transmission is completed. Although the Bellman optimality equation is typically solved using dynamic programming algorithms, such as policy iteration and value iteration, we next derive an optimal solution with lower complexity.

B. Optimal Policy

We aim to identify a special subclass of policy within the class Π, referred to as a cyclic scheduler. In such a policy, there exists constants $\tau^{(1)}$ and $\tau^{(2)}$ such that, in each cycle i, the transmitter always sends features from modality 1 consecutively $\tau^{(1)}$ times, then switches to modality 2 for another $\tau^{(2)}$ consecutive transmissions. We seek to answer two key questions about the optimal policy:

(i) Does an optimal cyclic scheduler exist?
(ii) If (i) is true, let $\tau_{\text{opt}}^{(1)}$ and $\tau_{\text{opt}}^{(2)}$ denote the optimal number of consecutive transmissions. How can we determine these optimal values?

Next, we answer these two questions by solving the Bellman optimality equations. Note that Eq. (5) and Eq. (6) are coupled due to the shared variable $\bar{L}_{\text{opt}}$ and the relative value function. To address this, we first consider a simplified problem in which the decision for one modality is fixed. Specifically, without loss of generality, we assume that modality 2 has a longer transmission time than modality 1 (i.e., $T^{(1)}\geq T^{(2)}$) and fix the number of consecutive transmissions for modality 2. Let Π_τ represent the set of all such scheduling policies:

$$\Pi_\tau := \left\{(\tau_i^{(1)},\tau_i^{(2)})\in\Pi : \tau_1^{(2)}=\tau_2^{(2)}=\cdots=\tau\right\}. \tag{9}$$

Given τ, the simplified problem is then expressed as

$$\textbf{OPT-S} \quad \bar{L}_{\text{opt}}(\tau) := \inf_{\pi\in\Pi_\tau} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_\pi\left[\sum_{t=0}^{T-1} L\left(\boldsymbol{\Delta}(t)\right)\right],$$

where $\bar{L}_{\text{opt}}(\tau)$ denotes the optimal objective value of **OPT-S** when the number of consecutive transmissions for modality 2, $\tau_i^{(2)}$, is fixed to a given value τ for each cycle i. Solving this simplified problem is crucial to deriving the optimal solution to the original problem **OPT**. Before presenting the solution to **OPT-S**, we first define the index function of modality 1:

$$\gamma^{(1)}(\theta+1) := \inf_{k\in\{1,2,\dots\}} \Bigg[\mathcal{L}^{(2)}(T^{(1)},(\theta+k)T^{(1)}+T^{(2)}) + \sum_{t=T^{(2)}}^{T^{(1)}-1} L(T^{(1)}+t,\theta T^{(1)}+T^{(2)}+t) + \sum_{j=1}^{k-1}\mathcal{L}^{(1)}(T^{(1)},(\theta+j)T^{(1)}+T^{(2)})\Bigg]\frac{1}{kT^{(1)}}, \tag{10}$$

where $\theta \in \mathbb{Z}_+$. Intuitively, a modality is selected only if its transmission yields low inference error. Suppose modality 1 has been selected for θ consecutive transmissions; then $\gamma^{(1)}(\theta+1)$ denotes the minimum average error if it continues for another k transmissions, taking the infimum over all k. See [5] for more physical interpretations of the index function. With the index function, we have the following theorem.

Theorem 1. *Under Assumption 1, for any fixed number of consecutive transmissions τ, the optimal policy to **OPT-S** is given by $\tau_i^{(1)} = \tau_{\text{opt}}^{(1)}(\beta(\tau))$, where*

$$\tau_{\text{opt}}^{(1)}(\beta(\tau)) = \inf\{\theta \in \mathbb{Z}_+ : \gamma^{(1)}(\theta+1) \geq \beta(\tau)\}. \tag{11}$$

The index function $\gamma^{(1)}(\theta)$ is defined in (10) *and the value of $\beta(\tau)$ can be obtained by solving the following equation:*

$$\begin{aligned}
&\sum_{j=1}^{\tau_{\text{opt}}^{(1)}(\beta(\tau))-1} \mathcal{L}^{(1)}(T^{(1)}, jT^{(1)} + T^{(2)}) \\
&+ \mathcal{L}^{(1)}(\tau T^{(1)} + T^{(2)}, T^{(2)}) - \tau_{\text{opt}}^{(1)}(\beta(\tau)) \cdot T^{(1)}\beta(\tau) \\
&+ \mathcal{L}^{(2)}(T^{(1)}, \tau_{\text{opt}}^{(1)}(\beta(\tau)) \cdot T^{(1)} + T^{(2)}) \\
&+ \sum_{j=1}^{\tau-1} \mathcal{L}^{(2)}(T^{(1)} + jT^{(2)}, T^{(2)}) - \tau T^{(2)}\beta(\tau) = 0.
\end{aligned} \tag{12}$$

*Furthermore, the threshold $\beta(\tau)$ is exactly the optimal objective value of **OPT-S**, i.e., $\beta(\tau) = \bar{L}_{\text{opt}}(\tau)$.*

Proof sketch: When the number of consecutive transmissions for modality 2 is fixed, the Bellman optimality equations reduce to a single-variable equation with respect to the only decision $\tau_i^{(1)}$ for modality 1. The idea is to show that the solution in Eq. (11) satisfies the reduced Bellman optimality equation. First, we show that **OPT-S** boils down to a single-state SMDP, allowing the subscript i to be dropped. Then, we use induction to show that $\tau^{(1)} = \tau_{\text{opt}}^{(1)}(\bar{L}_{\text{opt}}(\tau))$ if and only if Eq. (11) is satisfied. Finally, by showing that Eq. (12) has a finite and unique root, we prove the existence of the optimal objective value $\bar{L}_{\text{opt}}(\tau)$, ensuring that Eq. (11) is well-defined. The detailed proof is provided in Appendix B. ■

The optimal scheduling policy for **OPT-S**, characterized in Proposition 1, exhibits a nice structure. With the number of consecutive transmissions for modality 2 fixed as τ, it is optimal to select modality 1 consecutively for θ times, until the index function $\gamma^{(1)}(\theta+1)$ exceeds a threshold $\beta(\tau)$. The index function $\gamma^{(1)}(\cdot)$, defined in Eq. (10), depends only on the known parameter $L(\cdot)$, $T^{(1)}$, $T^{(2)}$, and the input θ. Therefore, it can be readily computed independently of modality 2. The threshold $\beta(\tau)$, which captures the impact of decisions for modality 2, is exactly the optimal objective value of **OPT-S** and can be obtained by solving Eq. (12). In Eq. (12), the only variable is $\beta(\tau)$; variable $\tau_{\text{opt}}^{(1)}(\beta(\tau))$ can be determined given $\beta(\tau)$, and all other terms are given parameters. There exist low-complexity algorithms, such as bisection search, that can efficiently solve this equation [16, Algorithm 1-3].

Remark 1. *Although similar index-based policies have been recently explored for different remote inference systems [6]–[8], our index function reflects a setting with two coupled sources that have heterogeneous transmission times. Particularly, for uniform transmission times (i.e., $T^{(1)} = T^{(2)}$), the index function in Eq.* (10) *reduces to*

$$\gamma^{(1)}(\theta) = \inf_{k\in\{1,2,\dots\}} \frac{1}{kT^{(1)}} \sum_{j=1}^{k} \mathcal{L}^{(1)}(T^{(1)}, (\theta+k)T^{(1)}),$$

which resembles that of the single-source case.

We now proceed to solve the original problem **OPT**. Analogous to the index function $\gamma^{(1)}(\cdot)$ defined for modality 1, we provide the index function $\gamma^{(2)}(\cdot)$ for modality 2 as follows.

$$\begin{aligned}
\gamma^{(2)}(\theta+1) \coloneqq \inf_{k\in\{1,2,\dots\}} \Bigg[&\mathcal{L}^{(1)}(T^{(1)} + (\theta+k)T^{(2)}, T^{(2)}) \\
&- \sum_{t=T^{(2)}}^{T^{(1)}-1} L(T^{(1)} + \theta T^{(2)} + t, T^{(2)} + t) \\
&+ \sum_{j=1}^{k-1} \mathcal{L}^{(2)}(T^{(1)} + (\theta+j)T^{(2)}, T^{(2)})\Bigg] \frac{1}{kT^{(2)}}.
\end{aligned} \tag{13}$$

Theorem 2. *Under Assumption 1, an optimal policy for **OPT** $\{(\tau_i^{(1)}, \tau_i^{(2)})\}$ is determined as follows:*

$$\tau_1^{(m)} = \tau_2^{(m)} = \cdots = \tau_{\text{opt}}^{(m)}(\beta), \tag{14}$$

where $\tau_{\text{opt}}^{(m)}(\beta)$ is given by

$$\tau_{\text{opt}}^{(m)}(\beta) = \inf\{\theta \in \mathbb{Z}_+ : \gamma^{(m)}(\theta+1) \geq \beta\},\ m \in \{1,2\}. \tag{15}$$

The index function $\gamma^{(1)}(\theta)$ and $\gamma^{(2)}(\theta)$ are defined in Eq. (10) *and Eq.* (13)*, respectively; the threshold β is given by*

$$\beta = \inf_{\tau\in\{1,2,\dots\}} \bar{L}_{\text{opt}}(\tau), \tag{16}$$

*where $\bar{L}_{\text{opt}}(\tau)$ is the optimal objective value of **OPT-S**.*

*Furthermore, the threshold β is exactly the optimal objective value of **OPT**, i.e., $\beta = \bar{L}_{\text{opt}}$. Hence, the optimal number of consecutive transmissions is given by $\tau_{\text{opt}}^{(m)} = \tau_{\text{opt}}^{(m)}(\bar{L}_{\text{opt}})$.*

Proof sketch: The proof consists of two steps. Step 1: Show that the decisions $\tau_{\text{opt}}^{(m)}(\bar{L}_{\text{opt}})$ is independent of i. Step 2: Show that $\tau_{\text{opt}}^{(m)}(\bar{L}_{\text{opt}})$ satisfies the Bellman optimality equations if Eq. (15) holds, and that the optimal objective value can be obtained by Eq. (16). Step 1 can be shown by substituting Eq. (5) into Eq. (6) and applying Theorem 1. For Step 2, we first prove Eq. (16). The intuition is that a policy achieves the optimal objective value if and only if it satisfies the Bellman optimality equations. Based on this intuition, it suffices to show that the policy $\tau_{\text{opt}}^{(m)}(\bar{L}_{\text{opt}})$ satisfies the Bellman optimality equation if and only if it achieves the infimum in Eq. (16). By applying Theorem 1 with Eq. (16) for both modalities, we obtain Eq. (15). The detailed proof is provided in Appendix C. ■

Theorem 2 answers the two questions posed above. First, it establishes the existence of an optimal cyclic scheduling policy. Second, it provides a simple method to determine the optimal decisions $\tau_{\rm opt}^{(m)}$: *the transmitter should stop transmitting the current modality when its index function exceeds a threshold.* The index function is pre-computed independently for each modality, and the threshold is exactly the optimal objective value of **OPT**. According to Eq. (16), this optimal objective value can be found by performing a one-dimensional search over τ and solving Eq. (12) at each iteration.

So far, the key insight of Theorem 2 is that obtaining the optimal solutions $\tau_{\rm opt}^{(1)}$ and $\tau_{\rm opt}^{(2)}$ requires only the computation of the optimal objective value $\bar{L}_{\rm opt}$ of **OPT**. While Eq. 16 suggests a straightforward way to compute $\bar{L}_{\rm opt}$, but *do we have a more efficient way to compute it?*

Intuitively, since the optimal solutions $\tau_{\rm opt}^{(m)}(\bar{L}_{\rm opt})$ depends on the optimal objective value $\bar{L}_{\rm opt}$, we can obtain a single-variable equation for $\bar{L}_{\rm opt}$, by substituting $\tau_{\rm opt}^{(m)}(\bar{L}_{\rm opt})$ into the Bellman optimality equations Eq. (5) and Eq. (6). Our earlier question now becomes: can we efficiently solve this equation? Motivated by this intuition, we present the following theorem.

Theorem 3. *Define $f(z) \coloneqq f_1(z) - zf_2(z)$, where*

$$\begin{aligned} f_1(z) \coloneqq & \mathcal{L}^{(1)}(T^{(1)} + \tau_{\rm opt}^{(2)}(z)T^{(2)}, T^{(2)}) \\ & + \sum_{j=1}^{\tau_{\rm opt}^{(1)}-1} \mathcal{L}^{(1)}(T^{(1)}, jT^{(1)} + T^{(2)}) \\ & + \mathcal{L}^{(2)}(T^{(1)}, \tau_{\rm opt}^{(1)}(z)T^{(1)} + T^{(2)}) \\ & + \sum_{j=1}^{\tau_{\rm opt}^{(2)}-1} \mathcal{L}^{(2)}(T^{(1)} + jT^{(2)}, T^{(2)}), \end{aligned}$$

and $f_2(z)$ is defined as $f_2(z) \coloneqq \tau_{\rm opt}^{(1)}(z)T^{(1)} + \tau_{\rm opt}^{(2)}(z)T^{(2)}$. The optimal objective value $\bar{L}_{\rm opt}$ is the root of the equation

$$f(z) = 0. \tag{17}$$

Furthermore, The function $f(z)$ has the following properties:
(i) $f(z)$ is concave, continuous, and strictly decreasing in z,
(ii) $\lim_{z\to-\infty} f(z) = -\infty$ and $\lim_{z\to\infty} f(z) = \infty$.

Proof sketch: Function $f(z)$ is derived from the Bellman optimality equations with Eq. (15). It can be shown that $f(z)$ is essentially the infimum of multiple linear and strictly decreasing functions of z, which leads to the properties stated above. The complete proof is similar to that of [7, Lemma 9] and is therefore omitted. ■

Given these properties of $f(z)$, we can efficiently solve the equation $f(z) = 0$ using algorithms such as bisection search, exactly as we do for Eq. (12). By replacing z with $\bar{L}_{\rm opt}$ in the equation $f(z) = f_1(z) - zf_2(z) = 0$, we obtain

$$\bar{L}_{\rm opt} = \frac{f_1(\bar{L}_{\rm opt})}{f_2(\bar{L}_{\rm opt})}. \tag{18}$$

Eq. (18) appears to be a more fundamental expression than Eq. (12) and has a clear physical meaning: $f_1(\bar{L}_{\rm opt}))$ represents the total expected inference error incurred in one cycle, and $f_1(\bar{L}_{\rm opt}))$ represents the total number of time slots spent in that cycle; thus, their ratio corresponds to the average expected inference error per time slot. It is surprising that the optimal objective value can still be efficiently computed, even with two coupled decisions. This result stems from our derivation of closed-form expressions for the two decisions in Eq. (15).

Remark 2. *Our results can be viewed as a generalization of the two-source scheduling with a monotonic increasing loss function (with respect to AoI) and uniform transmission times. Specifically, when the loss function $L(\cdot)$ is monotonic increasing and the two modalities have equal transmission times (i.e., $T^{(1)} = T^{(2)}$), the index function simplifies to*

$$\gamma^{(1)}(\theta) = \mathcal{L}^{(1)}(T^{(1)}, (\theta+1)T^{(1)})/T^{(1)},$$

which can be shown non-decreasing. In this case, it can be further shown that the index-based threshold policy reduces to the optimal periodic schedule as established in [12]. Overall, our index function handles general loss functions and heterogeneous transmission times, and it can still be pre-computed separately for each modality.

IV. Numerical Case Study: Robot State Prediction

In this section, we examine our multimodal remote inference system through a case study on robot state prediction and conduct a trace-driven evaluation to assess the performance of the proposed algorithm in Theorem 2.

A. System Description

Tracking the robot's state in real time is essential for various tasks, such as digital twin applications, safety monitoring, and remote robot control. Consider a robot performing tasks such as lifting objects or walking. It gathers both environmental and self-state information using multimodal sensors, such as LiDAR, cameras, and onboard sensors. On the receiver side, a predictor aims to continuously track the robot's state Y_t, such as its pose and velocity. To achieve this, a transmitter sends the collected information from the robot to the receiver, where an ML model is used to predict the robot's state based on the received data. Due to the high dimensionality and varying sizes of the data, transmission often spans multiple time slots and may differ across modalities. Next, we describe the experimental setup in detail.

B. Experimental Setup

We consider the OpenAI Bipedal Walker as our robot task, where a four-joint robot must run over stumps, pitfalls, and other obstacles (see [35] for details). We use the TD3-FORK reinforcement learning algorithm [36] to control the robot, as it achieves state-of-the-art performance on this task. After training the control model, we generate a time-series dataset in the OpenAI Gymnasium simulation environment, consisting of LiDAR rangefinder measurements, robot state information, and joint control signals.

For the state prediction task, we aim to predict the robot's joint velocities using sequential LiDAR measurements and joint control signals as two distinct modalities. We adopt the

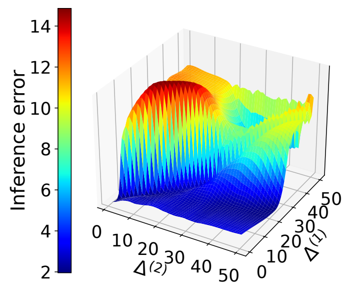
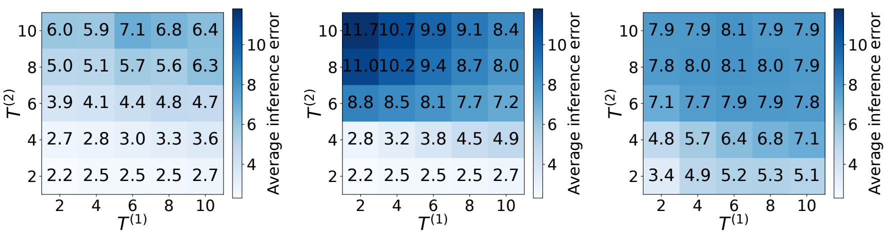


Fig. 3: Inference error vs. AoI.



(a) Index policy **(ours)** (b) Round-robin (c) Random

Fig. 4: Performance comparison under varying transmission times.

Long Short-Term Memory (LSTM) neural network as the predictor model, due to its widespread use and effectiveness in time-series forecasting. The network architecture includes an input layer, a hidden layer with 20 LSTM cells, and a fully connected output layer. We use 80% of the dataset for training. To incorporate AoI into model training, we augment the dataset as follows: given any AoI vector $(\delta^{(1)}, \delta^{(2)})$, we construct the feature-label pairs $(X^{(1)}_{t-\delta^{(1)}}, X^{(2)}_{t-\delta^{(2)}}; Y_t)$ for all data points, where the input features are aligned with their corresponding AoI values. The LSTM network takes the AoI vector as two additional input features and is trained on the whole augmented training dataset.

All experiments were run on a server with an AMD EPYC 7313 CPU (16 cores) and a single NVIDIA A2 GPU.

C. The Impact of AoI on Inference Error

Fig. 3 illustrates how the inference error varies with the Age of Information (AoI) for two modalities, with each AoI ranging from 1 to 50. The color intensity or surface height represents the expected inference error on the testing dataset. We observe that, although the inference error generally increases with the AoI of either modality, the relationship is not strictly monotonic. Furthermore, the impact of each modality differs significantly: as the AoI of modality 1 (LiDAR) increases, the inference error grows significantly faster than it does for modality 2 (control signal), indicating that LiDAR data is more strongly correlated with the target signal. Overall, the inference error is affected by both the AoI of each modality and its correlation to the target, which increases the complexity of the loss function and highlights the importance of addressing the modality transmission problem.

D. Scheduling Policies Evaluation

We now use the results obtained in Section IV-C as the empirical expected inference error function with respect to AoI, and evaluate the following scheduling policies:

(i) Index-based threshold policy: This is our proposed policy, as described in Theorem 2.

(ii) Round-robin policy: This policy alternates between the two modalities. Notably, it is optimal for minimizing the sum of AoI, i.e., $L(\mathbf{\Delta}(t)) = \sum_{m=1}^{2} \Delta^{(m)}(t)$.

(iii) Simple randomized policy: The policy randomly selects one of the two modalities with equal probability at each time slot, regardless of the current AoI state.

In practice, the transmission time for a given modality may depend on the sensor's sampling rate or resolution. To reflect this, we vary the transmission time for each modality using values from $2, 4, 6, 8, 10$, respectively. Fig. 4 presents the performance of each scheduling policy under varying transmission times. Our proposed index policy consistently achieves the best performance across all cases, reducing the inference error by up to 53% compared to the randomized policy and up to 55% compared to the round-robin policy. Although the round-robin policy is effective for minimizing AoI, it fails to capture the complex relationship between inference error and AoI, leading to suboptimal performance in certain scenarios (e.g., when $T^{(1)} = 2$ and $T^{(2)} = 10$). The simple randomized policy exhibits similar limitations, as it ignores both inference error and AoI. These results underscore the importance of jointly considering both modalities and their impact on inference error when designing scheduling policies.

V. Limitations

While we present an optimal index-based threshold policy for the coupled two-modality case, extending it to general multimodal settings remains open. A natural question arises: Can an index-based threshold policy with a closed-form expression still exist for more than two modalities? Investigating such generalizations would deepen our theoretical understanding of the power of index functions.

Moreover, we assume reliable communication with deterministic transmission times, which simplifies both analysis and policy derivation. Considering unreliable communication or random transmission delays would provide more realistic and robust solutions, but would also introduce additional complexity and lead to more intricate policy design.

Our empirical analysis is limited to a simulated robot with basic simulated modalities. However, many real-world tasks involve more complex modalities, such as high-dimensional data (e.g., video streams) and require more sophisticated ML models. Evaluation in such settings would further validate the practical effectiveness of the proposed scheduling policy.

VI. CONCLUSION

In this paper, we considered a problem of goal-oriented communication scheduling for a multimodal remote inference system, where a multimodal ML model at the receiver relies on time-sensitive data from remote sources to perform real-time inference tasks. Focusing on the two-modality scenario, we addressed the modality selection problem with the goal of minimizing the inference error at the receiver. We derived an optimal scheduling policy under the setting where the two modalities may have different transmission times. Finally, we conducted experiments on robot state prediction to evaluate our multimodal remote inference system. Numerical results show that our policy reduces inference error by up to 55% compared to baseline methods.

REFERENCES

[1] M. K. Chowdhury Shisher, H. Qin, L. Yang, F. Yan, and Y. Sun, "The age of correlated features in supervised learning based forecasting," in *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2021, pp. 1–8.

[2] E. Uysal, "Goal-oriented communications for interplanetary and non-terrestrial networks," *arXiv preprint arXiv:2409.14534*, 2024.

[3] A. Asvadi, P. Girão, P. Peixoto, and U. Nunes, "3d object tracking using rgb and lidar data," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2016, pp. 1255–1260.

[4] S. Duan, Q. Shi, and J. Wu, "Multimodal sensors and ml-based data fusion for advanced robots," *Advanced Intelligent Systems*, vol. 4, no. 12, p. 2200213, 2022.

[5] M. K. C. Shisher and Y. Sun, "How does data freshness affect real-time supervised learning?" in *ACM MobiHoc*, New York, NY, USA, 2022, p. 31–40.

[6] 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 Journal on Selected Areas in Information Theory*, pp. 524–538, 2023.

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

[8] C. Ari, M. K. C. Shisher, E. Uysal, and Y. Sun, "Goal-oriented communications for remote inference under two-way delay with memory," in *2024 IEEE International Symposium on Information Theory (ISIT)*, 2024, pp. 1179–1184.

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

[10] P. Whittle, "Restless bandits: Activity allocation in a changing world," *Journal of Applied Probability*, vol. 25, pp. 287–298, 1988.

[11] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 4th ed. Athena Scientific, 2017, vol. I.

[12] L. Shi and H. Zhang, "Scheduling two gauss–markov systems: An optimal solution for remote state estimation under bandwidth constraint," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 2038–2042, 2012.

[13] 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 Transactions on Information Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.

[14] Y. Sun and B. Cyr, "Sampling for data freshness optimization: Non-linear age functions," *Journal of Communications and Networks*, vol. 21, no. 3, pp. 204–219, 2019.

[15] M. K. Chowdhury Shisher and Y. Sun, "On the monotonicity of information aging," in *IEEE INFOCOM 2024 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2024, pp. 01–06.

[16] T. Z. Ornee and Y. Sun, "Sampling and remote estimation for the ornstein-uhlenbeck process through queues: Age of information and beyond," *IEEE/ACM Transactions on Networking*, vol. 29, no. 5, pp. 1962–1975, 2021.

[17] I. Kadota, A. Sinha, and E. Modiano, "Optimizing age of information in wireless networks with throughput constraints," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, 2018, pp. 1844–1852.

[18] T. Z. Ornee and Y. Sun, "A whittle index policy for the remote estimation of multiple continuous gauss-markov processes over parallel channels," in *ACM MobiHoc*, New York, NY, USA, 2023, p. 91–100.

[19] 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 *MILCOM 2023 - 2023 IEEE Military Communications Conference (MILCOM)*, 2023, pp. 194–200.

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

[21] Z. Liu, Y. Sang, B. Li, and B. Ji, "A worst-case approximate analysis of peak age-of-information via robust queueing approach," in *IEEE INFOCOM 2021-IEEE Conference on Computer Communications*. IEEE, 2021, pp. 1–10.

[22] F. Li, Y. Sang, Z. Liu, B. Li, H. Wu, and B. Ji, "Waiting but not aging: Optimizing information freshness under the pull model," *IEEE/ACM Transactions on Networking*, vol. 29, no. 1, pp. 465–478, 2020.

[23] Z. Liu, K. Zhang, B. Li, Y. Sun, Y. T. Hou, and B. Ji, "Learning-augmented online minimization of age of information and transmission costs," in *IEEE INFOCOM 2024-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2024, pp. 01–08.

[24] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1183–1210, 2021.

[25] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides, "The age of incorrect information: A new performance metric for status updates," *IEEE/ACM Trans. Netw.*, vol. 28, no. 5, p. 2215–2228, 2020.

[26] Q. Liu, C. Li, Y. T. Hou, W. Lou, J. H. Reed, and S. Kompella, "Ao2i: Minimizing age of outdated information to improve freshness in data collection," in *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*. IEEE, 2022, pp. 1359–1368.

[27] G. Chen, S. C. Liew, and Y. Shao, "Uncertainty-of-information scheduling: A restless multiarmed bandit framework," *IEEE Transactions on Information Theory*, vol. 68, no. 9, pp. 6151–6173, 2022.

[28] S. Chakraborty and Y. Sun, "Send pilot or data? leveraging age of channel state information for throughput maximization," in *IEEE INFOCOM 2025-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2025, p. accepted.

[29] 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 *IEEE INFOCOM 2025 - IEEE Conference on Computer Communications*. IEEE, 2025, p. accepted.

[30] J. Tong, L. Fu, and Z. Han, "Age-of-information oriented scheduling for multichannel iot systems with correlated sources," *IEEE Transactions on Wireless Communications*, vol. 21, no. 11, pp. 9775–9790, 2022.

[31] V. Tripathi and E. Modiano, "Optimizing age of information with correlated sources," *IEEE/ACM Transactions on Networking*, vol. 32, no. 6, pp. 4660–4675, 2024.

[32] R. V. Ramakanth, V. Tripathi, and E. Modiano, "Monitoring correlated sources: Aoi-based scheduling is nearly optimal," *IEEE Transactions on Mobile Computing*, vol. 24, no. 2, pp. 1043–1054, 2025.

[33] B. Zhou and W. Saad, "On the age of information in internet of things systems with correlated devices," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.

[34] S. Chakraborty and Y. Sun, "Timely remote estimation with memory at the receiver," in *2024 58th Asilomar Conference on Signals, Systems, and Computers*, 2024, pp. 111–115.

[35] M. Towers, A. Kwiatkowski, J. Terry, J. U. Balis, G. De Cola, T. Deleu, M. Goulão, A. Kallinteris, M. Krimmel, A. KG *et al.*, "Gymnasium: A standard interface for reinforcement learning environments," *arXiv preprint arXiv:2407.17032*, 2024.

[36] H. Wei and L. Ying, "Fork: A forward-looking actor for model-free reinforcement learning," in *2021 60th IEEE Conference on Decision and Control (CDC)*, 2021, pp. 1554–1559.

[37] Y. Sun, Y. Polyanskiy, and E. Uysal, "Sampling of the wiener process for remote estimation over a channel with random delay," *IEEE Transactions on Information Theory*, vol. 66, no. 2, pp. 1118–1135, 2020.

Appendix A
The SMDP formulation of Problem **OPT**

Here, we provide a detailed description of the components of the SMDP [11] for Problem **OPT**, along with the derivation of the Bellman optimality equations.

(i) **Action:** For each cycle i, the transmitter determines the number of consecutive transmissions $\tau_i^{(m)}$ for modality m, meaning it will continuously transmit the freshest feature of modality m continuously for $\tau_i^{(m)}$ times.

(ii) **Decision Time:** Let $S_i^{(m)}$ denote the decision time slot for action $\tau_i^{(m)}$. Then, the next decision time slot will be $\tau_i^{(m)}T^{(m)}$ later. For modality 1, we have $S_{i+1}^{(1)} = S_i^{(2)} + \tau_i^{(2)}T^{(2)}$, and for modality 2, we have $S_i^{(2)} = S_i^{(1)} + \tau_i^{(1)}T^{(1)}$.

(iii) **State and State Transitions:** The system state at each decision time $S_i^{(m)}$ is represented by the AoI at that time slot, i.e., $\boldsymbol{\Delta}(S_i^{(m)})$, which can be determined using Lemma 1. At time slot $S_i^{(1)}$, the AoI of modality 2 is $T^{(2)}$ because a feature from modality 2 has just been delivered, while the AoI of modality 1 is $T^{(1)} + \tau_{i-1}^{(2)}T^{(2)}$ because the last feature of modality 1 was delivered $\tau_{i-1}^{(2)}T^{(2)}$ time slots ago. Therefore, we have

$$\boldsymbol{\Delta}(S_i^{(1)}) = (T^{(1)} + \tau_{i-1}^{(2)}T^{(2)}, T^{(2)}).$$

Similarly, for $S_i^{(2)}$, we have

$$\boldsymbol{\Delta}(S_i^{(2)}) = (T^{(1)}, \tau_i^{(1)}T^{(1)} + T^{(2)}).$$

The state transition simply follows the sequence: $(\dots, \boldsymbol{\Delta}(S_i^{(1)}), \boldsymbol{\Delta}(S_i^{(2)}), \boldsymbol{\Delta}(S_{i+1}^{(1)}), \dots)$.

(iv) **Transition Time and Cost:** As explained in **Decision Time**, the transition time from $S_i^{(1)}$ to $S_i^{(2)}$ is $\tau_i^{(1)}T^{(1)}$, and the transition time from $S_i^{(2)}$ to $S_{i+1}^{(1)}$ is $\tau_i^{(2)}T^{(2)}$. From $S_i^{(1)}$ to $S_i^{(2)}$, the transmitter sends $\tau_i^{(1)}$ features from modality 1. Each time a feature is delivered, the AoI of modality 1 drops to $T^{(1)}$, while the AoI of modality 2 continues to increase. The initial state is $(T^{(1)} + \tau_{i-1}^{(2)}T^{(2)}, T^{(2)})$. Hence, the total transition cost is given by

$$\begin{aligned}&\sum_{j=1}^{\tau_i^{(1)}-1}\sum_{t=0}^{T^{(1)}-1} L(T^{(1)}+t, jT^{(1)}+T^{(2)}+t)\\&+\sum_{t=0}^{T^{(1)}-1} L(T^{(1)}+\tau_{i-1}^{(2)}T^{(2)}+t, T^{(2)}+t)\\&=\sum_{j=1}^{\tau_i^{(1)}-1}\mathcal{L}^{(1)}(T^{(1)}, jT^{(1)}+T^{(2)})\\&\quad+\mathcal{L}^{(1)}(T^{(1)}+\tau_{i-1}^{(2)}T^{(2)}, T^{(2)}).\end{aligned} \tag{19}$$

Similarly, the transmission cost from $S_i^{(2)}$ to $S_{i+1}^{(1)}$ is

$$\begin{aligned}&\sum_{j=1}^{\tau_i^{(2)}-1}\sum_{t=0}^{T^{(2)}-1} L(T^{(1)}+jT^{(2)}+t, T^{(2)}+t)\\&+\sum_{t=0}^{T^{(2)}-1} L(T^{(1)}+t, \tau_i^{(1)}T^{(1)}+T^{(2)}+t)\\&=\sum_{j=1}^{\tau_i^{(2)}-1}\mathcal{L}^{(2)}(T^{(1)}+jT^{(2)}, T^{(2)})\\&\quad+\mathcal{L}^{(2)}(T^{(1)}, \tau_i^{(1)}T^{(1)}+T^{(2)}).\end{aligned} \tag{20}$$

Given state, action, transmission duration and cost, the Bellman optimality equations in Eq.(5) and Eq.(6) follow directly.

Appendix B
Proof for Proposition 1

Proof: In this proof, we focus on the case when the continuation time for modality 2 is fixed, i.e., $\tau_i^{(2)} = \tau$ for all i. Accordingly, the Bellman optimality equation of Problem **OPT-S** can be written as Eq. (21). Our goal is to obtain the optimal $\tau_i^{(1)}$ that minimizes the relative value function $h(T^{(1)} + \tau T^{(2)}, T^{(2)})$, which is on the left-hand side (LHS) of Eq. (21). From the last equation of (21), we can see that the SMDP always stays in the same state $(T^{(1)} + \tau T^{(2)}, T^{(2)})$, regardless of the choice of $\tau_i^{(1)}$. That is, the term $h(T^{(1)} + \tau T^{(2)}, T^{(2)})$ can be pulled out of the infimum and canceled from both sides. Then, we can drop the subscript i and focus on solving the following problem:

$$\begin{aligned}\inf_{\tau^{(1)}\in\{1,2,\dots\}}\Big[&\sum_{j=1}^{\tau^{(1)}-1}\mathcal{L}^{(1)}(T^{(1)}, jT^{(1)}+T^{(2)})\\&-\tau^{(1)}T^{(1)}\bar{L}_{\text{opt}}(\tau)+\mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}T^{(1)}+T^{(2)})\Big].\end{aligned} \tag{22}$$

Let $\tau_{\text{opt}}^{(1)}(\bar{L}_{\text{opt}}(\tau))$ be an optimal solution to Eq. (22). We will use induction to show that $\tau_{\text{opt}}^{(1)}(\bar{L}_{\text{opt}}(\tau))$ follows Eq. (11) in Theorem 1. First, one equivalent condition for $\tau_{\text{opt}}^{(1)}(\bar{L}_{\text{opt}}(\tau)) = 1$ is given as follows.

$$\begin{aligned}\inf_{\tau^{(1)}\in\{2,3,\dots\}}\Big[&\sum_{j=1}^{\tau^{(1)}-1}\mathcal{L}^{(1)}(T^{(1)}, jT^{(1)}+T^{(2)})\\&-\tau^{(1)}T^{(1)}\bar{L}_{\text{opt}}(\tau)+\mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}T^{(1)}+T^{(2)})\Big]\\\geq{}&\mathcal{L}^{(2)}(T^{(1)}, T^{(1)}+T^{(2)})-T^{(1)}\bar{L}_{\text{opt}}(\tau),\end{aligned} \tag{23}$$

where the right-hand side (RHS) comes from substituting $\tau^{(1)} = 1$ into Eq. (22). So far, the relationship between Eq.(23)

$$
\begin{aligned}
&h(T^{(1)}+\tau T^{(2)},T^{(2)})\\
&=\inf_{\tau_i^{(1)}\in\{1,2,\dots\}}\Big[\mathcal{L}^{(1)}(T^{(1)}+\tau T^{(2)},T^{(2)})+\sum_{j=1}^{\tau_i^{(1)}-1}\mathcal{L}^{(1)}(T^{(1)},jT^{(1)}+T^{(2)})-\tau_i^{(1)}\bar{L}_{\text{opt}}(\tau)+h(T^{(1)},T^{(2)}+\tau_i^{(1)}T^{(1)})\Big]\\
&=\inf_{\tau_i^{(1)}\in\{1,2,\dots\}}\Big[\mathcal{L}^{(1)}(T^{(1)}+\tau T^{(2)},T^{(2)})+\sum_{j=1}^{\tau_i^{(1)}-1}\mathcal{L}^{(1)}(T^{(1)},jT^{(1)}+T^{(2)})-\tau_i^{(1)}T^{(1)}\bar{L}_{\text{opt}}(\tau)\\
&\qquad+\mathcal{L}^{(2)}(T^{(1)},\tau_i^{(1)}T^{(1)}+T^{(2)})+\sum_{j=1}^{\tau-1}\mathcal{L}^{(2)}(T^{(1)}+jT^{(2)},T^{(2)})-\tau T^{(2)}\bar{L}_{\text{opt}}(\tau)+h(T^{(1)}+\tau T^{(2)},T^{(2)})\Big]\\
&=\inf_{\tau_i^{(1)}\in\{1,2,\dots\}}\Big[\sum_{j=1}^{\tau_i^{(1)}-1}\mathcal{L}^{(1)}(T^{(1)},jT^{(1)}+T^{(2)})-\tau_i^{(1)}T^{(1)}\bar{L}_{\text{opt}}(\tau)+\mathcal{L}^{(2)}(T^{(1)},\tau_i^{(1)}T^{(1)}+T^{(2)})\Big]\\
&\qquad+\mathcal{L}^{(1)}(T^{(1)}+\tau T^{(2)},T^{(2)})+\sum_{j=1}^{\tau-1}\mathcal{L}^{(2)}(T^{(1)}+jT^{(2)},T^{(2)})-\tau T^{(2)}\bar{L}_{\text{opt}}(\tau)+h(T^{(1)}+\tau T^{(2)},T^{(2)}).
\end{aligned}
\tag{21}
$$

and our target result Eq.(11) remains unclear. To see this, we first rearrange the term in Eq. (23),

$$
\begin{aligned}
&\mathcal{L}^{(1)}(T^{(1)},T^{(1)}+T^{(2)})-\mathcal{L}^{(2)}(T^{(1)},T^{(1)}+T^{(2)})\\
&+\inf_{\tau^{(1)}\in\{2,3,\dots\}}\Big[\sum_{j=1}^{\tau^{(1)}-2}\mathcal{L}^{(1)}(T^{(1)},T^{(1)}+jT^{(1)}+T^{(2)})\\
&\quad-(\tau^{(1)}-1)T^{(1)}\bar{L}_{\text{opt}}(\tau)+\mathcal{L}^{(2)}(T^{(1)},\tau^{(1)}T^{(1)}+T^{(2)})\Big]\\
&\geq 0.
\end{aligned}
\tag{24}
$$

Replace $\tau^{(1)}$ with $k=\tau^{(1)}-1$, we obtain

$$
\begin{aligned}
&\mathcal{L}^{(1)}(T^{(1)},T^{(1)}+T^{(2)})-\mathcal{L}^{(2)}(T^{(1)},T^{(1)}+T^{(2)})\\
&+\inf_{k\in\{1,2,\dots\}}\Big[\sum_{j=1}^{k-1}\mathcal{L}^{(1)}(T^{(1)},T^{(1)}+jT^{(1)}+T^{(2)})\\
&\qquad+\mathcal{L}^{(2)}(T^{(1)},T^{(1)}+kT^{(1)}+T^{(2)})\\
&\qquad-kT^{(1)}\bar{L}_{\text{opt}}(\tau)\Big]\geq 0.
\end{aligned}
\tag{25}
$$

Next, expand $\mathcal{L}^{(1)}$ and $\mathcal{L}^{(2)}$ by their definition, we have

$$
\begin{aligned}
&\mathcal{L}^{(1)}(T^{(1)},T^{(1)}+T^{(2)})-\mathcal{L}^{(2)}(T^{(1)},T^{(1)}+T^{(2)})\\
&=\sum_{t=T^{(2)}}^{T^{(1)}-1}L(T^{(1)}+t,T^{(1)}+T^{(2)}+t).
\end{aligned}
\tag{26}
$$

Combining Eq. (26) and Eq. (25), we now have

$$
\begin{aligned}
&\inf_{k\in\{1,2,\dots\}}\Big[\sum_{t=T^{(2)}}^{T^{(1)}-1}L(T^{(1)}+t,T^{(1)}+T^{(2)}+t)\\
&+\sum_{j=1}^{k-1}\mathcal{L}^{(1)}(T^{(1)},T^{(1)}+jT^{(1)}+T^{(2)})\\
&+\mathcal{L}^{(2)}(T^{(1)},T^{(1)}+kT^{(1)}+T^{(2)})-kT^{(1)}\bar{L}_{\text{opt}}(\tau)\Big]\geq 0.
\end{aligned}
\tag{27}
$$

Next, similar to [37, Lemma 2], Eq. (27) is equivalent to the following inequality:

$$
\begin{aligned}
&\inf_{k\in\{1,2,\dots\}}\frac{1}{kT^{(1)}}\Big[\sum_{t=T^{(2)}}^{T^{(1)}-1}L(T^{(1)}+t,T^{(1)}+T^{(2)}+t)\\
&\quad+\sum_{j=1}^{k-1}\mathcal{L}^{(1)}(T^{(1)},T^{(1)}+jT^{(1)}+T^{(2)})\\
&\quad+\mathcal{L}^{(2)}(T^{(1)},T^{(1)}+kT^{(1)}+T^{(2)})\Big]\geq\bar{L}_{\text{opt}}(\tau).
\end{aligned}
\tag{28}
$$

The left-hand side is exactly $\gamma^{(1)}(2)$ according to the definition of the index function (10). Therefore, we have shown that $\gamma^{(1)}(2)\geq\bar{L}_{\text{opt}}(\tau)$ is the condition for $\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))=1$. By repeating this process, for any θ, we can derive $\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))=\theta$ if $\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))\neq 1,2,\dots,\theta-1$ and

$$
\begin{aligned}
&\inf_{k\in\{1,2,\dots\}}\frac{1}{kT^{(1)}}\Big[\sum_{j=1}^{k-1}\mathcal{L}^{(1)}(T^{(1)},(\theta+j)T^{(1)}+T^{(2)})\\
&+\mathcal{L}^{(2)}(T^{(1)},(\theta+k)T^{(1)}+T^{(2)})\\
&+\sum_{t=T^{(2)}}^{T^{(1)}-1}\mathcal{L}^{(1)}(T^{(1)}+t,\theta T^{(1)}+T^{(2)}+t)\Big]\geq\bar{L}_{\text{opt}}(\tau).
\end{aligned}
\tag{29}
$$

The condition $\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))\neq 1,2,\dots,\theta-1$ implies that $\gamma^{(1)}(k)<\bar{L}_{\text{opt}}(\tau)$ for $k=2,\dots,\theta$; and Eq. (29) derives $\gamma^{(1)}(\theta+1)\geq\bar{L}_{\text{opt}}(\tau)$. Therefore, we conclude that the optimal continuation time $\tau^{(1)}=\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))$ is determined by

$$
\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))=\inf\{\theta:\gamma^{(1)}(\theta+1)\geq\bar{L}_{\text{opt}}(\tau)\}.
\tag{30}
$$

Now, we are ready to compute the optimal objective value $\bar{L}_{\text{opt}}(\tau)$. From Eq. (21), cancel the term $h(T^{(1)}+\tau T^{(2)},T^{(2)})$

from both sides and substitute $\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))$, we obtain Eq. (12) in Proposition 1, i.e.,

$$\begin{aligned}
&\sum_{j=1}^{\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))-1} \mathcal{L}^{(1)}(T^{(1)}, jT^{(1)}+T^{(2)}) \\
&+\mathcal{L}^{(1)}(T^{(1)}+\tau T^{(2)}, T^{(2)}) - \tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))T^{(1)}\bar{L}_{\text{opt}}(\tau) \\
&+\mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))T^{(1)}+T^{(2)}) \\
&+\sum_{j=1}^{\tau-1}\mathcal{L}^{(2)}(T^{(1)}+jT^{(2)}, T^{(2)}) - \tau T^{(2)}\bar{L}_{\text{opt}}(\tau) = 0.
\end{aligned} \tag{31}$$

Finally, the following lemma shows that there exists a unique, finite solution to Eq. (31). Define

$$\begin{aligned}
J(\beta) \coloneqq &\sum_{j=1}^{\tau^{(1)}_{\text{opt}}(\beta)-1} \mathcal{L}^{(1)}(T^{(1)}, jT^{(1)}+T^{(2)}) \\
&+\mathcal{L}^{(1)}(T^{(1)}+\tau T^{(2)}, T^{(2)}) - \tau^{(1)}_{\text{opt}}(\beta)T^{(1)}\beta \\
&+\mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}_{\text{opt}}(\beta)T^{(1)}+T^{(2)}) \\
&+\sum_{j=1}^{\tau-1}\mathcal{L}^{(2)}(T^{(1)}+jT^{(2)}, T^{(2)}) - \tau T^{(2)}\beta.
\end{aligned} \tag{32}$$

Lemma 2. *[7, Lemma 9] The function $J(\beta)$ has the following properties:*

1) *The function $J(\beta)$ is concave, continuous, and strictly decreasing in β.*
2) $\lim_{\beta\to\infty} J(\beta) = -\infty$ *and* $\lim_{\beta\to-\infty} J(\beta) = \infty$.

■

Appendix C
Proof for Theorem 2

Proof: The idea is to show the optimality of the solution in Theorem 2 by solving the Bellman optimality equations (5), (6). For convenience, we restate the solution again:

$$\tau_1^{(m)} = \tau_2^{(m)} = \cdots = \tau^{(m)}_{\text{opt}}(\beta), \tag{34}$$

where $\tau^{(m)}_{\text{opt}}(\beta)$ is given by

$$\tau^{(m)}_{\text{opt}}(\beta) = \inf\{\theta\in\mathbb{Z}_+ : \gamma^{(m)}(\theta+1)\geq\beta\},\ m\in\{1,2\}. \tag{35}$$

First, we show that the existence of an optimal periodic scheduling policy, i.e., there exists $\tau^{(1)}$ and $\tau^{(2)}$ such that $\tau_i^{(1)} = \tau^{(1)}$ and $\tau_i^{(2)} = \tau^{(2)}$ for each i that solves Eq. (5), (6). For any $\tau_i^{(2)}$, Proposition 1 implies that the solution to Eq. (5) should be $\tau_i^{(1)} = \tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau_i^{(2)}))$. Then, substitute $h(T^{(1)}+\tau_i^{(2)}T^{(2)}, T^{(2)})$ into Eq. (6) we obtain Eq. (33). In the right-hand side of Eq. (33), the only term related to the state $(T^{(1)}, \tau_{i-1}^{(1)}T^{(1)}+T^{(2)})$ is $\mathcal{L}^{(2)}(T^{(1)}, \tau_{i-1}^{(1)}T^{(1)}+T^{(2)})$, which is independent of $\tau_i^{(2)}$. Therefore, the solution of Eq. (33) can be obtained by solving the following equation:

$$\begin{aligned}
&\inf_{\tau^{(2)}\in\{1,2,\dots\}}\Big[\sum_{j=1}^{\tau^{(2)}-1}\mathcal{L}^{(2)}(T^{(1)}+jT^{(2)}, T^{(2)}) - \tau^{(2)}T^{(2)}\bar{L}_{\text{opt}} \\
&+\mathcal{L}^{(1)}(T^{(1)}+\tau^{(2)}T^{(2)}, T^{(2)}) - \tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau^{(2)}))T^{(1)}\bar{L}_{\text{opt}} \\
&+\sum_{j=1}^{\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau^{(2)}))-1}\mathcal{L}^{(1)}(T^{(1)}, jT^{(1)}+T^{(2)}) \\
&+\mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau^{(2)}))T^{(1)}+T^{(2)})\Big].
\end{aligned} \tag{36}$$

Note that we can simply change $\tau_i^{(2)}$ to $\tau^{(2)}$ and drop the subscript i because all terms are independent of $\tau_{i-1}^{(1)}$. Also, the term $\mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau^{(2)}))T^{(1)}+T^{(2)})$ is the only term related to $\tau^{(2)}$ in $h(T^{(1)}, \tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau^{(2)}))T^{(1)}+T^{(2)})$. So we can use it to replace $h(T^{(1)}, \tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau^{(2)}))T^{(1)}+T^{(2)})$ in Eq. (36) without changing the solution. Let $\tau^{(2)}_{\text{opt}}$ be the optimal solution to Eq. (36). Now, we use proof by contradiction to show that Eq. (16) holds, i.e., $\bar{L}_{\text{opt}}(\tau^{(2)}_{\text{opt}}) = \inf_{\tau\in\{1,2,\dots\}}\bar{L}_{\text{opt}}(\tau)$. Assume $\tau\in\mathbb{Z}_+$ is the optimal solution to Eq. (36) while there exists $\tau'\in\mathbb{Z}_+$ such that $\bar{L}_{\text{opt}}(\tau')\leq\bar{L}_{\text{opt}}(\tau)$. Since $\tau^{(2)}_{\text{opt}} = \tau$, according to Proposition 1, we must have $\bar{L}_{\text{opt}} = \bar{L}_{\text{opt}}(\tau)$, otherwise the bellman optimality equation Eq. (5) does not hold. From Eq. (12), we also have the following equation

$$\begin{aligned}
&\sum_{j=1}^{\tau-1}\mathcal{L}^{(2)}(T^{(1)}+jT^{(2)}, T^{(2)}) - \tau T^{(2)}\bar{L}_{\text{opt}}(\tau) \\
&+\mathcal{L}^{(1)}(T^{(1)}+\tau T^{(2)}, T^{(2)}) - \tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))T^{(1)}\bar{L}_{\text{opt}}(\tau) \\
&+\sum_{j=1}^{\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))-1}\mathcal{L}^{(1)}(T^{(1)}, jT^{(1)}+T^{(2)}) \\
&+\mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))T^{(1)}+T^{(2)}) = 0.
\end{aligned} \tag{37}$$

This equation holds true if we replace τ, $\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))$, and $\bar{L}_{\text{opt}}(\tau)$ by τ', $\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau'), \bar{L}_{\text{opt}}(\tau'))$, respectively. If we fix τ and $\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))$ in Eq. (37), then clearly the left-hand side is decreasing with respect to $\bar{L}_{\text{opt}}(\tau)$. With the assumption that $\bar{L}_{\text{opt}}(\tau')\leq\bar{L}_{\text{opt}}(\tau)$, we replace τ, $\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau))$ by τ', $\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau'))$ and left-hand side of Eq. (37) will below 0, i.e.,

$$\begin{aligned}
&\sum_{j=1}^{\tau'-1}\mathcal{L}^{(2)}(T^{(1)}+jT^{(2)}, T^{(2)}) - \tau' T^{(2)}\bar{L}_{\text{opt}}(\tau) \\
&+\mathcal{L}^{(1)}(T^{(1)}+\tau' T^{(2)}, T^{(2)}) - \tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau'))T^{(1)}\bar{L}_{\text{opt}}(\tau) \\
&+\sum_{j=1}^{\tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau'))-1}\mathcal{L}^{(1)}(T^{(1)}, jT^{(1)}+T^{(2)}) \\
&+\mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}_{\text{opt}}(\bar{L}_{\text{opt}}(\tau'))T^{(1)}+T^{(2)}) < 0,
\end{aligned} \tag{38}$$

$$
\begin{aligned}
&h(T^{(1)}, \tau_{i-1}^{(1)} T^{(1)} + T^{(2)}) \\
&= \mathcal{L}^{(2)}(T^{(1)}, \tau_{i-1}^{(1)} T^{(1)} + T^{(2)}) + \inf_{\tau_i^{(2)} \in \{1,2,\dots\}} \Big[\sum_{j=1}^{\tau_i^{(2)}-1} \mathcal{L}^{(2)}(T^{(1)} + jT^{(2)}, T^{(2)}) - \tau_i^{(2)} T^{(2)} \bar{L}_{\mathrm{opt}} + \mathcal{L}^{(1)}(T^{(1)} + \tau_i^{(2)} T^{(2)}, T^{(2)}) \\
&\quad + \sum_{j=1}^{\tau_{\mathrm{opt}}^{(1)}(\bar{L}_{\mathrm{opt}}(\tau_i^{(2)}))-1} \mathcal{L}^{(1)}(T^{(1)}, jT^{(1)} + T^{(2)}) - \tau_{\mathrm{opt}}^{(1)}(\bar{L}_{\mathrm{opt}}(\tau_i^{(2)})) T^{(1)} \bar{L}_{\mathrm{opt}} + h(T^{(1)}, \tau_{\mathrm{opt}}^{(1)}(\bar{L}_{\mathrm{opt}}(\tau_i^{(2)})) T^{(1)} + T^{(2)}) \Big].
\end{aligned} \tag{33}
$$

which leads to the contradiction that τ achieves the infimum.
Now, by applying Proposition 1 with $\tau^{(2)} = \tau_{\mathrm{opt}}^{(2)}$, we have

$$
\bar{L}_{\mathrm{opt}} = \bar{L}_{\mathrm{opt}}(\tau_{\mathrm{opt}}^{(2)}) = \inf_{\tau \in \{1,2,\dots\}} \bar{L}_{\mathrm{opt}}(\tau),
$$

and

$$
\begin{aligned}
\tau_{\mathrm{opt}}^{(1)} &= \tau_{\mathrm{opt}}^{(1)}(\bar{L}_{\mathrm{opt}}(\tau_{\mathrm{opt}}^{(2)})) \\
&= \inf\{\theta \in \mathbb{Z}_+ : \gamma^{(1)}(\theta + 1) \geq \bar{L}_{\mathrm{opt}}(\tau_{\mathrm{opt}}^{(2)})\} \\
&= \inf\{\theta \in \mathbb{Z}_+ : \gamma^{(1)}(\theta + 1) \geq \bar{L}_{\mathrm{opt}}\}.
\end{aligned}
$$

■