Heterogeneous Multimodal Remote Inference under Bandwidth Constraints: Two-Modality Scheduling

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**

I. Introduction

The advent of sixth-generation (6G) technology opens up vast opportunities for intelligent systems—including autonomous transportation, unmanned mobility, industry automation, and extended reality—by enabling reliable connectivity between diverse sensing nodes and remote processing nodes. Many complex tasks in these systems often involve multiple *data modalities*, such as visual, auditory, spatial (e.g., LiDAR), and motion (e.g., inertial sensor) 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 [1]. Similarly, for teleoperated robotics, tactile sensors provide additional information about the surface properties of objects [2].

Moreover, for many complex real-time tasks that utilize data-driven machine learning (ML) models, timely input data delivery is crucial for enabling accurate estimations and informed decisions in dynamic environments. However, limited bandwidth and other realistic network constraints often prevent the ML model from consistently accessing fresh data across all available modalities simultaneously. Traditionally, data freshness has not been a primary concern when evaluating the model performance; however, recent work [3] demonstrates both theoretically and empirically that the freshness of multiple correlated features can significantly impact the accuracy of ML models under various time-series forecasting scenarios. This insight leads to a natural research question: *How can we schedule modalities transmission to control their freshness and eventually optimize ML model performance, given limited network resources?*

To that end, we study a time-slotted *multimodal remote inference* problem: at each time slot, the scheduler must decide which modality to transmit from multiple sensors to a remote receiver. 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 affected by data freshness, we adopt the system model in recent work [4]–[7]. Specifically, we first use the Age of Information (AoI) [8] to measure data freshness: for each modality $m = 1, 2, \ldots, M$, the AoI of modality m at time slot t, denoted as $\Delta^{(m)}(t)$, is given by $\Delta^{(m)}(t) := t - U(t)$, where $U(t)$ is the generation time of the most recently delivered data packet. Then, the inference error at time slot t is characterized as a function of the AoI vector of all modalities $\mathbf{\Delta}(t) = (\Delta^{(1)}(t), \Delta^{(2)}(t), \ldots, \Delta^{(M)}(t))$. Moreover, we assume that not all modalities can be transmitted simultaneously due to the bandwidth constraint. Our goal is to find an optimal transmission scheduling policy to minimize the time-averaged expected inference error over an infinite horizon. We summarize the contribution of this paper as follows:

- We formulate a two-modality selection 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 problem poses two main challenges: (i) 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. (ii) We consider heterogeneous transmission times across modalities, which naturally arise from differences in data sizes or sampling rates. This introduces additional complexity and requires additional technical effort.

- We solve the formulated problem and theoretically derive the optimal policy. To tackle the challenge, we first cast the problem into an infinite-horizon average-cost Semi-Markov Decision Problem (SMDP), which is typically solved using dynamic programming [9]. Yet, dynamic programming methods are usually computationally expensive and lack insight. We show that the optimal policy can be expressed as an index-based threshold policy. Similar policies have proven effective for handling non-monotonic inference error in single-source scenarios, but it remained unclear whether a similarly structured policy could be extended to the scheduling of two modalities.
- We conduct numerical experiments to evaluate our proposed algorithm. Results show that it reduces inference error by up to 55% compared to baseline methods.

A. Related Work

AoI penalty function. The concept of the Age of Information (AoI) has been extensively studied in a large body of work (see eg., [3]–[6], [8], [10]–[25] and a recent survey [26]). Over time, its concept has progressively become more general, encompassing a wider range of applications. Initially, research efforts primarily focused on analyzing and optimizing linear AoI, such as average AoI and peak AoI [8], [10], [13], [15], [22]. Later, as surveyed in [11], several non-decreasing non-linear utility functions of AoI were suggested to characterize data freshness. Meanwhile, many studies also considered various AoI-related metrics, such as Age of Incorrect Information (AoII) [27], Age of Outdated Information (Ao^{2}I) [28], and Uncertainty of Information [29]. More recently, it was shown in [3] that the impact of AoI from multiple correlated features on inference error can be non-monotonic. However, how to design a scheduling policy when the inference error is a non-monotonic, non-additive function of AoI vector remains underexplored. This paper partially bridges this gap by showing an optimal policy for the two-modality case.

Remote inference. Recently, a line of research [5]–[7] has focused on feature transmission scheduling for the remote inference problem. In [7], an optimal policy was proposed for the single-source case, incorporating the single-source settings studied in [5], [6] as special cases. However, these results do not directly extend to the two-modality case, as the modalities are coupled due to the non-additive objective function and bandwidth constraints. On the other hand, the multi-source case has been explored in [5], [6], where the objective function is formulated as a weighted sum of inference errors from each source-predictor pair. This work considers a setting where multiple data modalities contribute to the same learning task, resulting in a non-additive objective function. Without additivity, previous methods cannot be applied to our setting.

Remote state estimation. This work also relates to the field of remote state estimation [14], [18], [30]–[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], [30]–[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, 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 slot index $t \in \{0, \dots, T\}$. 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 a communication channel that allows at most one modality to be transmitted per time slot due to the bandwidth constraint. 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 modality m. We assume that $T^{(m)}$ is a known constant for each modality, 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

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

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 [8], 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 receive side at time slot t. For $D_n \le t < D_{n+1}$, given the scheduler's decisions a_{n+1}, the AoI $\Delta^{(m)}(t)$ is represented by

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

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 $m' = 3 - m$ as the other modality. Then, for each n, we have

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

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 evaluate 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 $\ell : \mathcal{P} \times \mathcal{Y} \mapsto \mathbb{R}$, where $\ell(y, p)$ denotes the value of inference error at time slot t when the prediction is $P_t = p$ and the target is $Y_t = y$. 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 $\mathbf{\Delta}(t)$ [4]. 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 $\mathbf{X}_{t-\boldsymbol{\delta}}$.

C. Structure of the Scheduling Policy

The scheduling policy π can be represented as a sequence of modality choices, $\pi := \{a_n, n = 1, 2, \dots\}$. 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 weighted 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 Eq. (2).

III. Optimal Scheduling Policy Design

In this section, we present the optimal policy for **OPT**. We first reformulate the problem and cast it as a Semi-Markov Decision Process (SMDP). Then, we derive 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 selection decisions $\{a_n\}$. 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 decisions $\{a_n\}$ can be represented by another sequence

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

where $\tau_i^{(m)}$ denotes the number of consecutive tranmission of modality m before switching to the other modality. Instead of selecting a modality for each transmission, the transmitter now decides how many times to transmit one modality consecutively; once these transmissions are completed, it selects the

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

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

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 transmission from modality 1, and finally one transmission from modality 2. We refer to each pair $(\tau_i^{(1)},\tau_i^{(2)})$ as the i-th round, and to $\tau_i^{(m)}$ as the continuation times for modality m during the i-th round. Note that although the two policies (i.e., $\{a_n\}$ and $\{(\tau_i^{(1)},\tau_i^{(2)})\}$) have different decision times, they are actually *equivalent* for the following two reasons: (i) any sequence $\{a_n\}$ can be derived from $\{(\tau_i^{(1)},\tau_i^{(2)})\}$ and vice versa; (ii) the decisions are made based on the same available information because the transmission times $T^{(1)}$ and $T^{(2)}$ are determinisitc. Thus, with a slight abuse of notation, we denote the new policy as $\pi=\{(\tau_i^{(1)},\tau_i^{(2)}),i=1,2,\dots\}$.

Now, with the new policy, we cast **OPT** as a Semi-Markov Decision Process (SMDP) [9]. In the SMDP, the action is the continuation time $\tau_i^{(m)}$, and the system state at time slot t is represented by the AoI vector $\mathbf{\Delta}(t)$. 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 $h(\Delta^{(1)},\Delta^{(2)})$ denote the relative value function at the state $(\Delta^{(1)},\Delta^{(2)})$; the function $\mathcal{L}^{(m)}(\cdot)$ is defined as

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

where $\mathcal{L}^{(m)}(x,y)$ represents the total expected inference error incurred when transmitting modality m from state (x,y) 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

Finding the optimal policy involves addressing two key questions:

(i) Does there exist an optimal cyclic scheduler; that is, are there constants $\tau_{\text{opt}}^{(1)}$ and $\tau_{\text{opt}}^{(2)}$ such that $\tau_i^{(1)}=\tau_{\text{opt}}^{(1)}$ and $\tau_i^{(2)}=\tau_{\text{opt}}^{(2)}$ for all i, and the resulting policy is optimal?

(ii) If (i) is true, how to determine $\tau_{\text{opt}}^{(1)}$ and $\tau_{\text{opt}}^{(2)}$?

Next, we answer these two questions by solving the Bellman optimality equations. Note that Eq. (5) and Eq. (6) are still coupled due to the common variable $\bar{L}_{\text{opt}}$. To address this, we first try to solve a simpler problem where one of the continuation times is fixed. Specifically, since two modalities are symmetric, we assume that $T^{(1)}\geq T^{(2)}$ and fix the continuation time of modality 2 without loss of generality. Let Π_τ represent the set of all causal scheduling policies with a fixed continuation time sequence for modality 2:

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

Given τ, the simplified problem is then expressed as

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

Solving **OPT-S** is crucial to deriving the optimal solution to the original problem **OPT**. We now present the solution to **OPT-S** in the following proposition.

Define an index function of modality 1 as

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

Proposition 1: When $\tau_i^{(2)}=\tau$ for any i, and $|L(\mathbf{\Delta})|\leq M$ is uniformly bounded by some finite M for all AoI $\mathbf{\Delta}$, the optimal policy to **OPT-S** is $\tau_i^{(1)}=\tau_{\text{opt}}^{(1)}(\beta(\tau))$, 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 to **OPT-S**, i.e., $\bar{L}_{\text{opt}}(\tau)=\beta(\tau)$.

Proof sketch: With $\tau_i^{(2)}=\tau$ fixed, the Bellman optimality equations reduce to a one-variable equation with respect to $\tau_i^{(1)}$. The idea is to show that Eq. (11) can solve the reduced Bellman optimality equation. First, we show that **OPT-S** boils down to a single-state SMDP, which implies that the subscript i can 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 unqiue root, we establish the existence of $\bar{L}_{\text{opt}}(\tau)$ and $\tau_{\text{opt}}^{(1)}(\bar{L}_{\text{opt}}(\tau))$. 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 continuation time of modality 2 fixed as τ, it is optimal to select modality 1 consecutively for θ times until $\gamma^{(1)}(\theta+1)$ exceeds a $\beta(\tau)$. 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 known parameters. There exist low-complexity algorithms, such as bisection search, that can efficiently solve this equation [14, Algorithm 1-3].

Remark 1: Although similar index-based policies have been recently explored for remote inference systems [5]–[7], our setting involves two coupled sources with heterogeneous transmission times, which is reflected in the structure of the index function. Deriving the index-based threshold rule in this case is nontrivial and requires additional technical efforts. Particularly, when $T^{(1)}=T^{(2)}$, the index function in Eq. (10) reduces to

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

Now we are ready to solve our original problem **OPT**.

Theorem 1: If $|L(\boldsymbol{\Delta})|\le M$ is uniformly bounded by some finite M for all AoI vectors $\boldsymbol{\Delta}$, then an optimal policy $\{\tau_i^{(1)}\}$ and $\{\tau_i^{(2)}\}$ for **OPT** is determined as follows:

(a) $\tau_1^{(2)}=\tau_2^{(2)}=\cdots=\tau_{\text{opt}}^{(2)}$, where $\tau_{\text{opt}}^{(2)}$ is given by

$$\tau_{\text{opt}}^{(2)}=\arg\inf_{\tau\in\{1,2,\dots\}}\bar{L}_{\text{opt}}(\tau), \tag{13}$$

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

(b) $\tau_1^{(1)}=\tau_2^{(1)}=\cdots=\tau_{\text{opt}}^{(1)}$, where $\tau_{\text{opt}}^{(1)}$ is given by

$$\tau_{\text{opt}}^{(1)}=\inf\{\theta\in\mathbb{Z}_+:\gamma^{(1)}(\theta+1)\ge\bar{L}_{\text{opt}}\}. \tag{14}$$

The index function $\gamma^{(1)}(\theta)$ is defined in (10), and the threshold $\bar{L}_{\text{opt}}$ is the optimal objective value of **OPT**, given by

$$\bar{L}_{\text{opt}}=\inf_{\tau\in\{1,2,\dots\}}\bar{L}_{\text{opt}}(\tau). \tag{15}$$

Proof sketch: The key point in this proof has two steps. Step 1: show that $\tau_{\text{opt}}^{(2)}$ is independent of i. Step 2: show that $\tau_{\text{opt}}^{(2)}$ solves the Bellman optimality equations if Eq. (13) holds. Then, by applying Proposition 1 with $\tau_i^{(2)}=\tau_{\text{opt}}^{(2)}$, we immediately obtain the result for $\tau_i^{(1)}$ as stated in Eq. (14). For Step 1, by substitute Eq. (5) into Eq. (6) and with Proposition 1, we can show that the solution to the Bellman optimality equation is independent of i. For Step 2, the intuition is that the policy solving the Bellman optimality equation must also achieve the optimal objective value. Thus, we use proof by contradiction and show that Eq. (13) is a sufficient condition for the optimal policy. The detailed proof is provided in Appendix C. ■

Theorem 1 answers the two questions posed above. First, it establishes the existence of an optimal cyclic scheduling policy. Second, it offers a simple way to determine the two continuation times: after obtaining $\bar{L}_{\text{opt}}(\tau)$ and $\tau_{\text{opt}}^{(1)}(\bar{L}_{\text{opt}}(\tau))$ by solving **OPT-S** using Proposition 1, the optimal solution to **OPT** can be found by searching the value of τ that minimizes the objective value $\bar{L}_{\text{opt}}(\tau)$.

Remark 2: Theorem 1 can be viewed as an extension result to the two-source scheduling with a monotonic increasing loss function (with respect to AoI) and uniform transmission time. Specifically, when $L(\cdot)$ is monotonic increasing and $T^{(1)}=T^{(2)}$, 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 [33]. 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 1.

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,

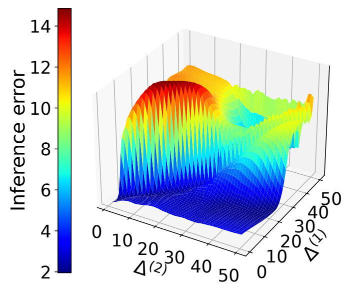
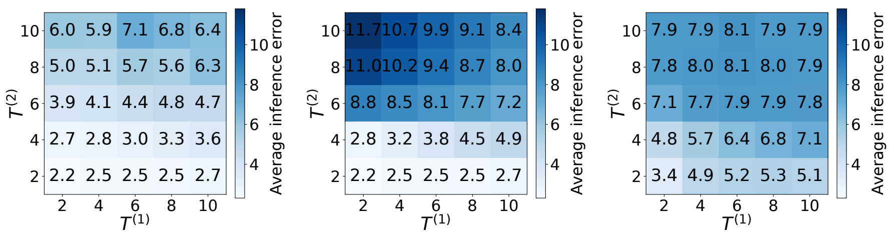


Fig. 1: Inference error vs. AoI for two modalities.



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

Fig. 2: Performance comparison under varying transmission times.

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 across stumps, pitfalls, and other obstacles (see [34] for details). We use the TD3-FORK reinforcement learning algorithm [35] 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 a 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 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 incoporate 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. 1 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 1.

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

(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 $2, 4, \ldots, 10$, respectively. Fig. 2 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 a coupled two-modality case, extending it to handle a general multimodal setting remains an open problem. A natural question arises: can we still obtain a decomposable index function for scheduling when more than two modalities are involved? 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 the analysis. 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 machine learning inference models. Evaluation in such settings would further validate the practical effectiveness of the proposed scheduling policy.

VI. CONCLUSION

In this paper, we considered 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] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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*, 2023.

[6] 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.

[7] 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.

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

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

[10] 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.

[11] 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.

[12] 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.

[13] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *2015 IEEE International Symposium on Information Theory (ISIT)*, 2015, pp. 3008–3012.

[14] 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.

[15] 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.

[16] M. Klügel, M. H. Mamduhi, S. Hirche, and W. Kellerer, "Aoi-penalty minimization for networked control systems with packet loss," in *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2019, pp. 189–196.

[17] J. Sun, Z. Jiang, B. Krishnamachari, S. Zhou, and Z. Niu, "Closed-form whittle's index-enabled random access for timely status update," *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1538–1551, 2020.

[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 *Proceedings of the Twenty-Fourth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*, ser. MobiHoc '23. New York, NY, USA: Association for Computing Machinery, 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, R. Singh, and N. B. Shroff, "Optimizing information freshness using low-power status updates via sleep-wake scheduling," in *Proceedings of the Twenty-First International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*, ser. Mobihoc '20, New York, NY, USA, 2020, p. 51–60.

[21] 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.

[22] 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.

[23] Z. Liu and B. Ji, "Towards the tradeoff between service performance and information freshness," in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*. IEEE, 2019, pp. 1–6.

[24] 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.

[25] 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.

[26] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE*

Journal on Selected Areas in Communications, vol. 39, no. 5, pp. 1183–1210, 2021.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] T. Z. Ornee and Y. Sun, “Performance bounds for sampling and remote estimation of gauss-markov processes over a noisy channel with random delay,” in *2021 IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2021, pp. 1–5.

[32] R. V. Ramakanth, V. Tripathi, and E. Modiano, “Monitoring correlated sources: Aoi-based scheduling is nearly optimal,” *IEEE Transactions on Mobile Computing*, 2024.

[33] 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.

[34] 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.

[35] 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.

APPENDIX A
THE SMDP FORMULATION OF PROBLEM **OPT**

Here, we provide a detailed description of the components of the SMDP [9] for Problem **OPT**, along with the derivation of the Bellman optimality equations.

(i) **Action:** For each round i, the transmitter determines the i-th continuation time $\tau_i^{(m)}$ 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., $\mathbf{\Delta}(S_i^{(m)})$, 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

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

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

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

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

(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

$$\begin{aligned}
&\sum_{j=1}^{\tau_i^{(1)}-1}\sum_{t=0}^{T^{(1)}-1} L(T^{(1)}+t, jT^{(1)}+T^{(2)}+t) \\
&+\sum_{t=0}^{T^{(1)}-1} L(T^{(1)}+\tau_{i-1}^{(2)}T^{(2)}+t, T^{(2)}+t) \\
&=\sum_{j=1}^{\tau_i^{(1)}-1} \mathcal{L}^{(1)}(T^{(1)}, jT^{(1)}+T^{(2)}) \\
&\quad+\mathcal{L}^{(1)}(T^{(1)}+\tau_{i-1}^{(2)}T^{(2)}, T^{(2)}).
\end{aligned} \tag{16}$$

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

$$\begin{aligned}
&\sum_{j=1}^{\tau_i^{(2)}-1}\sum_{t=0}^{T^{(2)}-1} L(T^{(1)}+jT^{(2)}+t, T^{(2)}+t) \\
&+\sum_{t=0}^{T^{(2)}-1} L(T^{(1)}+t, \tau_i^{(1)}T^{(1)}+T^{(2)}+t) \\
&=\sum_{j=1}^{\tau_i^{(2)}-1} \mathcal{L}^{(2)}(T^{(1)}+jT^{(2)}, T^{(2)}) \\
&\quad+\mathcal{L}^{(2)}(T^{(1)}, \tau_i^{(1)}T^{(1)}+T^{(2)}).
\end{aligned} \tag{17}$$

Given the state, action, transmission time, and transmission 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. (18). Our goal is to obtain the optimal $\tau_i^{(1)}$ that minimizes the relative value function $h(T^{(1)}+\tau^{(2)}T^{(2)}, T^{(2)})$. First, from the last equation of (18), we observe that the SMDP always remains in the same state $(T^{(1)}+\tau^{(2)}T^{(2)}, T^{(2)})$, regardless of the choice of $\tau_i^{(1)}$. Hence, we can drop the subscript i and focus on solving the following equation:

$$\begin{aligned}
\inf_{\tau^{(1)}\in\{1,2,\dots\}} \Big[&\sum_{j=1}^{\tau^{(1)}-1} \mathcal{L}^{(1)}(T^{(1)}, jT^{(1)}+T^{(2)}) \\
&- \tau^{(1)}T^{(1)}\bar{L}_{\mathrm{opt}}(\tau) + \mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}T^{(1)}+T^{(2)})\Big].
\end{aligned} \tag{19}$$

$$
\begin{aligned}
&h(T^{(1)}+\tau T^{(2)},T^{(2)})\\
&=\inf_{\tau_i^{(1)}\in\{1,2,\dots\}}\Big[\mathcal{L}^{(1)}(T^{(1)}+\tau T^{(2)},T^{(2)})+\sum_{j=1}^{\tau_i^{(1)}-1}\mathcal{L}^{(1)}(T^{(1)},jT^{(1)}+T^{(2)})-\tau_i^{(1)}\bar{L}_{\rm opt}(\tau)+h(T^{(1)},T^{(2)}+\tau_i^{(1)}T^{(1)})\Big]\\
&=\inf_{\tau_i^{(1)}\in\{1,2,\dots\}}\Big[\mathcal{L}^{(1)}(T^{(1)}+\tau T^{(2)},T^{(2)})+\sum_{j=1}^{\tau_i^{(1)}-1}\mathcal{L}^{(1)}(T^{(1)},jT^{(1)}+T^{(2)})-\tau_i^{(1)}T^{(1)}\bar{L}_{\rm opt}(\tau)\\
&\qquad+\mathcal{L}^{(2)}(T^{(1)},\tau_i^{(1)}T^{(1)}+T^{(2)})+\sum_{j=1}^{\tau-1}\mathcal{L}^{(2)}(T^{(1)}+jT^{(2)},T^{(2)})-\tau T^{(2)}\bar{L}_{\rm opt}(\tau)+h(T^{(1)}+\tau T^{(2)},T^{(2)})\Big]\\
&=\inf_{\tau_i^{(1)}\in\{1,2,\dots\}}\Big[\sum_{j=1}^{\tau_i^{(1)}-1}\mathcal{L}^{(1)}(T^{(1)},jT^{(1)}+T^{(2)})-\tau_i^{(1)}T^{(1)}\bar{L}_{\rm opt}(\tau)+\mathcal{L}^{(2)}(T^{(1)},\tau_i^{(1)}T^{(1)}+T^{(2)})\Big]\\
&\qquad+\mathcal{L}^{(1)}(T^{(1)}+\tau T^{(2)},T^{(2)})+\sum_{j=1}^{\tau-1}\mathcal{L}^{(2)}(T^{(1)}+jT^{(2)},T^{(2)})-\tau T^{(2)}\bar{L}_{\rm opt}(\tau)+h(T^{(1)}+\tau T^{(2)},T^{(2)}).
\end{aligned}\tag{18}
$$

Let $\tau_{\rm opt}^{(1)}(\bar{L}_{\rm opt}(\tau))$ be an optimal solution to Eq. (19). From Eq. (19), one equivalent condition for $\tau_{\rm opt}^{(1)}(\bar{L}_{\rm opt}(\tau))=1$ is

$$
\begin{aligned}
&\inf_{\tau^{(1)}\in\{2,3,\dots\}}\Big[\sum_{j=1}^{\tau^{(1)}-1}\mathcal{L}^{(1)}(T^{(1)},jT^{(1)}+T^{(2)})\\
&\quad-\tau^{(1)}T^{(1)}\bar{L}_{\rm opt}(\tau)+\mathcal{L}^{(2)}(T^{(1)},\tau^{(1)}T^{(1)}+T^{(2)})\Big]\\
&\geq\mathcal{L}^{(2)}(T^{(1)},T^{(1)}+T^{(2)})-T^{(1)}\bar{L}_{\rm opt}(\tau),
\end{aligned}\tag{20}
$$

where the right-hand side comes from substituting $\tau^{(1)}>1$ into Eq. (19). Rearrange the term we obtain,

$$
\begin{aligned}
&\mathcal{L}^{(1)}(T^{(1)},T^{(1)}+T^{(2)})-\mathcal{L}^{(2)}(T^{(1)},T^{(1)}+T^{(2)})\\
&+\inf_{\tau^{(1)}\in\{2,3,\dots\}}\Big[\sum_{j=1}^{\tau^{(1)}-2}\mathcal{L}^{(1)}(T^{(1)},T^{(1)}+jT^{(1)}+T^{(2)})\\
&\quad-(\tau^{(1)}-1)T^{(1)}\bar{L}_{\rm opt}(\tau)+\mathcal{L}^{(2)}(T^{(1)},\tau^{(1)}T^{(1)}+T^{(2)})\Big]\\
&\geq 0.
\end{aligned}\tag{21}
$$

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

$$
\begin{aligned}
&\mathcal{L}^{(1)}(T^{(1)},T^{(1)}+T^{(2)})-\mathcal{L}^{(2)}(T^{(1)},T^{(1)}+T^{(2)})\\
&+\inf_{k\in\{1,2,\dots\}}\Big[\sum_{j=1}^{k-1}\mathcal{L}^{(1)}(T^{(1)},T^{(1)}+jT^{(1)}+T^{(2)})\\
&\qquad+\mathcal{L}^{(2)}(T^{(1)},T^{(1)}+kT^{(1)}+T^{(2)})\\
&\qquad-kT^{(1)}\bar{L}_{\rm opt}(\tau)\Big]\geq 0.
\end{aligned}\tag{22}
$$

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

$$
\begin{aligned}
&\mathcal{L}^{(1)}(T^{(1)},T^{(1)}+T^{(2)})-\mathcal{L}^{(2)}(T^{(1)},T^{(1)}+T^{(2)})\\
&=\sum_{t=T^{(2)}}^{T^{(1)}-1}L(T^{(1)}+t,T^{(1)}+T^{(2)}+t).
\end{aligned}\tag{23}
$$

Combining Eq. (23) and Eq. (22), we now have

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

Next, similar to [30, Lemma 2], Eq. (24) is equivalent to the following inequality:

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

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

$$
\begin{aligned}
&\inf_{k\in\{1,2,\dots\}}\frac{1}{kT^{(1)}}\Big[\sum_{j=1}^{k-1}\mathcal{L}^{(1)}(T^{(1)},(\theta+j)T^{(1)}+T^{(2)})\\
&+\mathcal{L}^{(2)}(T^{(1)},(\theta+k)T^{(1)}+T^{(2)})\\
&+\sum_{t=T^{(2)}}^{T^{(1)}-1}\mathcal{L}^{(1)}(T^{(1)}+t,\theta T^{(1)}+T^{(2)}+t)\Big]\geq\bar{L}_{\rm opt}(\tau).
\end{aligned}\tag{26}
$$

The condition $\tau_{\rm opt}^{(1)}(\bar{L}_{\rm opt}(\tau))\neq 1,2,\dots,\theta-1$ implies that $\gamma^{(1)}(k)<\bar{L}_{\rm opt}(\tau)$ for $k=2,\dots,\theta$; and Eq. (26) derives

$\gamma^{(1)}(\theta+1) \geq \bar{L}_{\rm opt}(\tau)$. Therefore, we conclude that the optimal continuation time $\tau^{(1)} = \tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau))$ is determined by

$$\tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau)) = \inf\{\theta : \gamma^{(1)}(\theta+1) \geq \bar{L}_{\rm opt}(\tau)\}. \tag{27}$$

Now, we are ready to compute the optimal objective value $\bar{L}_{\rm opt}(\tau)$. From Eq. (18), cancel the term $h(T^{(1)}+\tau T^{(2)}, T^{(2)})$ from both sides and substitute $\tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau))$, we obtain Eq. (12) in Proposition 1, i.e.,

$$\begin{aligned}
&\sum_{j=1}^{\tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau))-1} \mathcal{L}^{(1)}(T^{(1)}, jT^{(1)}+T^{(2)}) \\
&+\mathcal{L}^{(1)}(T^{(1)}+\tau T^{(2)}, T^{(2)}) - \tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau))T^{(1)}\bar{L}_{\rm opt}(\tau) \\
&+\mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau))T^{(1)}+T^{(2)}) \\
&+\sum_{j=1}^{\tau-1}\mathcal{L}^{(2)}(T^{(1)}+jT^{(2)}, T^{(2)}) - \tau T^{(2)}\bar{L}_{\rm opt}(\tau) = 0.
\end{aligned} \tag{28}$$

Finally, the following lemma shows that there exists a unique, finite solution to Eq. (28). Define

$$\begin{aligned}
J(\beta) \coloneqq &\sum_{j=1}^{\tau^{(1)}_{\rm opt}(\beta)-1} \mathcal{L}^{(1)}(T^{(1)}, jT^{(1)}+T^{(2)}) \\
&+\mathcal{L}^{(1)}(T^{(1)}+\tau T^{(2)}, T^{(2)}) - \tau^{(1)}_{\rm opt}(\beta)T^{(1)}\beta \\
&+\mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}_{\rm opt}(\beta)T^{(1)}+T^{(2)}) \\
&+\sum_{j=1}^{\tau-1}\mathcal{L}^{(2)}(T^{(1)}+jT^{(2)}, T^{(2)}) - \tau T^{(2)}\beta.
\end{aligned} \tag{29}$$

Lemma 2: [6, 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 1

Proof: The idea is to show the optimality of the solution in Theorem 1 by solving the Bellman optimality equations (5), (6). For convenience, we restate the solution again:

(a) $\tau^{(2)}_1 = \tau^{(2)}_2 = \cdots = \tau^{(2)}_{\rm opt}$, where $\tau^{(2)}_{\rm opt}$ is given by

$$\tau^{(2)}_{\rm opt} = \arg\inf_{\tau\in\{1,2,\dots\}} \bar{L}_{\rm opt}(\tau). \tag{31}$$

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

(b) $\tau^{(1)}_1 = \tau^{(1)}_2 = \cdots = \tau^{(1)}_{\rm opt}$, where $\tau^{(1)}_{\rm opt}$ is given by

$$\tau^{(1)}_{\rm opt} = \inf\{\theta\in\mathbb{Z}_+ : \gamma^{(1)}(\theta+1) \geq \bar{L}_{\rm opt}\}. \tag{32}$$

The index function $\gamma^{(1)}(\theta)$ is defined in (10), and the threshold $\bar{L}_{\rm opt}$ is the optimal objective value of **OPT**, given by

$$\bar{L}_{\rm opt} = \inf_{\tau\in\{1,2,\dots\}} \bar{L}_{\rm opt}(\tau). \tag{33}$$

First, we show that the existence of an optimal periodic scheduling policy, i.e., there exists $\tau^{(1)}$ and $\tau^{(2)}$ such that $\tau^{(1)}_i = \tau^{(1)}$ and $\tau^{(2)}_i = \tau^{(2)}$ for each i that solves Eq. (5), (6). For any $\tau^{(2)}_i$, Proposition 1 implies that the solution to Eq. (5) should be $\tau^{(1)}_i = \tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau^{(2)}_i))$. Then, substitute $h(T^{(1)}+\tau^{(2)}_i T^{(2)}, T^{(2)})$ into Eq. (6) we obtain Eq. (30). In the right-hand side of Eq. (30), the only term related to the state $(T^{(1)}, \tau^{(1)}_{i-1}T^{(1)}+T^{(2)})$ is $\mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}_{i-1}T^{(1)}+T^{(2)})$, which is independent of $\tau^{(2)}_i$. Therefore, the solution of Eq. (30) can be obtained by solving the following equation:

$$\begin{aligned}
\inf_{\tau^{(2)}\in\{1,2,\dots\}} \Big[&\sum_{j=1}^{\tau^{(2)}-1} \mathcal{L}^{(2)}(T^{(1)}+jT^{(2)}, T^{(2)}) - \tau^{(2)}T^{(2)}\bar{L}_{\rm opt} \\
&+\mathcal{L}^{(1)}(T^{(1)}+\tau^{(2)}T^{(2)}, T^{(2)}) - \tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau^{(2)}))T^{(1)}\bar{L}_{\rm opt} \\
&+\sum_{j=1}^{\tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau^{(2)}))-1} \mathcal{L}^{(1)}(T^{(1)}, jT^{(1)}+T^{(2)}) \\
&+\mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau^{(2)}))T^{(1)}+T^{(2)})\Big].
\end{aligned} \tag{34}$$

Note that we can simply change $\tau^{(2)}_i$ to $\tau^{(2)}$ and drop the subscript i because all terms are independent of $\tau^{(1)}_{i-1}$. Also, the term $\mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau^{(2)}))T^{(1)}+T^{(2)})$ is the only term related to $\tau^{(2)}$ in $h(T^{(1)}, \tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau^{(2)}))T^{(1)}+T^{(2)})$. So we can use it to replace $h(T^{(1)}, \tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau^{(2)}))T^{(1)}+T^{(2)})$ in Eq. (34) without changing the solution. Let $\tau^{(2)}_{\rm opt}$ be the optimal solution to Eq. (34). Now, we use proof by contradiction to show that Eq. (31) holds, i.e., $\bar{L}_{\rm opt}(\tau^{(2)}_{\rm opt}) = \inf_{\tau\in\{1,2,\dots\}} \bar{L}_{\rm opt}(\tau)$. Assume $\tau\in\mathbb{Z}_+$ is the optimal solution to Eq. (34) while there exists $\tau'\in\mathbb{Z}_+$ such that $\bar{L}_{\rm opt}(\tau') \leq \bar{L}_{\rm opt}(\tau)$. Since $\tau^{(2)}_{\rm opt} = \tau$, according to Proposition 1, we must have $\bar{L}_{\rm opt} = \bar{L}_{\rm opt}(\tau)$, otherwise the bellman optimality equation Eq. (5) does not hold. From Eq. (12), we also have the following equation

$$\begin{aligned}
&\sum_{j=1}^{\tau-1}\mathcal{L}^{(2)}(T^{(1)}+jT^{(2)}, T^{(2)}) - \tau T^{(2)}\bar{L}_{\rm opt}(\tau) \\
&+\mathcal{L}^{(1)}(T^{(1)}+\tau T^{(2)}, T^{(2)}) - \tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau))T^{(1)}\bar{L}_{\rm opt}(\tau) \\
&+\sum_{j=1}^{\tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau))-1} \mathcal{L}^{(1)}(T^{(1)}, jT^{(1)}+T^{(2)}) \\
&+\mathcal{L}^{(2)}(T^{(1)}, \tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau))T^{(1)}+T^{(2)}) = 0.
\end{aligned} \tag{35}$$

This equation holds true if we replace τ, $\tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau))$, and $\bar{L}_{\rm opt}(\tau)$ by τ', $\tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau'), \bar{L}_{\rm opt}(\tau'))$, respectively. If we fix τ and $\tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau))$ in Eq. (35), then clearly the left-hand side is decreasing with respect to $\bar{L}_{\rm opt}(\tau)$. With the assumption that $\bar{L}_{\rm opt}(\tau') \leq \bar{L}_{\rm opt}(\tau)$, we replace τ, $\tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau))$ by τ', $\tau^{(1)}_{\rm opt}(\bar{L}_{\rm opt}(\tau'))$ and left-hand side of Eq. (35) will below 0,

$$
\begin{aligned}
&h(T^{(1)}, \tau_{i-1}^{(1)} T^{(1)} + T^{(2)}) \\
&= \mathcal{L}^{(2)}(T^{(1)}, \tau_{i-1}^{(1)} T^{(1)} + T^{(2)}) + \inf_{\tau_i^{(2)} \in \{1,2,\dots\}} \Big[\sum_{j=1}^{\tau_i^{(2)}-1} \mathcal{L}^{(2)}(T^{(1)} + jT^{(2)}, T^{(2)}) - \tau_i^{(2)} T^{(2)} \bar{L}_{\text{opt}} + \mathcal{L}^{(1)}(T^{(1)} + \tau_i^{(2)} T^{(2)}, T^{(2)}) \\
&\quad + \sum_{j=1}^{\tau_{\text{opt}}^{(1)}(\bar{L}_{\text{opt}}(\tau_i^{(2)}))-1} \mathcal{L}^{(1)}(T^{(1)}, jT^{(1)} + T^{(2)}) - \tau_{\text{opt}}^{(1)}(\bar{L}_{\text{opt}}(\tau_i^{(2)})) T^{(1)} \bar{L}_{\text{opt}} + h(T^{(1)}, \tau_{\text{opt}}^{(1)}(\bar{L}_{\text{opt}}(\tau_i^{(2)})) T^{(1)} + T^{(2)}) \Big].
\end{aligned} \tag{30}
$$

i.e.,

$$
\begin{aligned}
&\sum_{j=1}^{\tau'-1} \mathcal{L}^{(2)}(T^{(1)} + jT^{(2)}, T^{(2)}) - \tau' T^{(2)} \bar{L}_{\text{opt}}(\tau) \\
&+ \mathcal{L}^{(1)}(T^{(1)} + \tau' T^{(2)}, T^{(2)}) - \tau_{\text{opt}}^{(1)}(\bar{L}_{\text{opt}}(\tau')) T^{(1)} \bar{L}_{\text{opt}}(\tau) \\
&+ \sum_{j=1}^{\tau_{\text{opt}}^{(1)}(\bar{L}_{\text{opt}}(\tau'))-1} \mathcal{L}^{(1)}(T^{(1)}, jT^{(1)} + T^{(2)}) \\
&+ \mathcal{L}^{(2)}(T^{(1)}, \tau_{\text{opt}}^{(1)}(\bar{L}_{\text{opt}}(\tau')) T^{(1)} + T^{(2)}) < 0,
\end{aligned} \tag{36}
$$

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

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

and

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

■