Comparison of Decentralized and Centralized Update Paradigms for Remote Tracking of Distributed Dynamic Sources

***Abstract*—In this work, we perform a comparative study of *centralized* and *decentralized* update strategies for the basic remote tracking problem of many distributed users/devices with randomly evolving states. Our goal is to reveal the impact of the fundamentally different tradeoffs that exist between information accuracy and communication cost under these two update paradigms. In one extreme, decentralized updates are triggered by distributed users/transmitters based on *exact local state-information*, but also at a *higher cost* due to the need for uncoordinated multi-user communication. In the other extreme, centralized updates are triggered by the common tracker/receiver based on *estimated global state-information*, but also at a *lower cost* due to the capability of coordinated multi-user communication. We use a generic superlinear function to model the communication cost with respect to the number of simultaneous updates for multiple sources. We characterize the conditions under which transmitter-driven decentralized update policies outperform their receiver-driven centralized counterparts for symmetric sources, and vice versa. Further, we extend the results to a scenario where system parameters are unknown and develop learning-based update policies that asymptotically achieve the minimum cost levels attained by the optimal policies.**

I. INTRODUCTION

In recent years, there has been a growing number of applications requiring real-time updates of system status, especially in cyber-physical systems such as smart homes and buildings or health-care monitoring systems [1]. In such systems, each sensor samples and transmits time-varying information to a controller (or a monitor), and the controller makes system decisions based on the collected information from multiple sources. Although it is ideal to maintain the controller up-to-date all the way, this is often impractical due to limited resources of communication networks.

To address this challenge, Age of Information (AoI) has been introduced and utilized as a performance metric measuring the freshness of information [2], defined as the time elapsed since a new packet is generated at a source. In the single-source single-destination scenario, the optimal update rate to minimize AoI with random transmission time has been studied in [3]. In the multiple-source single-destination scenario, a scheduling problem under communication constraints has been studied in [4], [5]. In [4], at most one source can transmit a packet via a channel, where a packet is dropped with some probability. In [5], a channel is modeled as a FIFO queue with random service time.

Related to the AoI optimization problem, there have been studies on remote estimation problem of the system status [6]–[17], where the freshness of information is measured not in the age metric but directly in the estimation error, i.e., the absolute difference between the state information at the source and the destination. It has been observed that a sampling strategy that minimizes the AoI does not always minimize the estimation error [6], [7], where sampling problems of Wiener process and Ornstein-Uhoenbeck process with a channel modeled as FIFO queue are studied, respectively. In [8], remote estimation problems with a packet-drop channel for both finite state Markov source and First-order autoregressive source are investigated, where a channel state changes over time horizon following finite-state Markov chain, and the packet-drop probability depends on a channel state and the power level for transmission. In [9], the Automatic Repeat reQuest (ARQ)-based remote estimation framework are studied for the linear time-invariant (LTI) system, where a sensor's observation is noisy and a channel's gain changes over time following finite-state Markov chain. In this domain, several works have tackled the scheduling problem under communication constraints [10]–[13]. In [10], [11], the minimization problem of Mean Squared Error (MSE) of an estimator (or monitor) is considered when the number of transmissions over finite-time horizon is constrained. The scheduling problem with per-transmission communication cost has been studied in [11]–[13]. Each transmission is accompanied with communication cost and the objective is to schedule transmissions to optimize the MSE of an estimator and the communication cost of a transmitter over finite-time horizon [11], [12] or over infinite-time horizon [13].

Extending a single-source single-destination scenario investigated in [6]–[13], a multiple-source single-destination scenario has been investigated in [14]–[17]. In [14], [15], each source is modeled as a LTI system and, at each time slot, at most m out of n transmitters can update the remote monitor. The scheduling decision is made by a centralized controller (or the receiver). In [16], the scheduling decision is made by each distributed sensor/transmitter sensing a LTI system, and at most one transmitter can update the monitor. In contrast to [14], [15] where the objective is to minimize the estimation error covariance, the objective of [16] is to minimize the transmission power consumption that can stabilize the system.

In contrast to the aforementioned works , we investigate a remote estimation problem with communication cost that couples the simultaneous updates, where multiple sources update an estimator. We consider sources with simple random walk and assume a perfect channel, i.e, noiseless and no packet drop, as in [13], [17]. Different from [17], we consider shared communication channels (i.e., communication cost is correlated between sources).

Our contributions can be summarized as follows.

- We formulate the remote estimation problem in shared

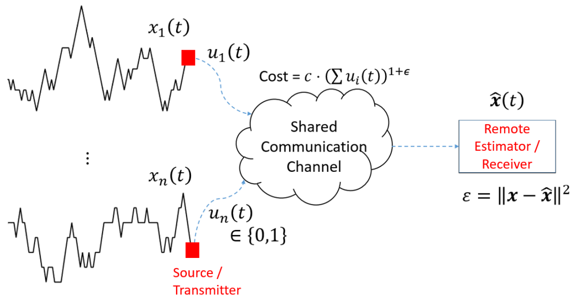


Fig. 1: System model.

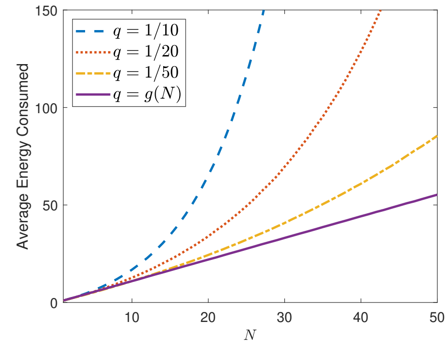


Fig. 2: Average energy consumed for successful update with respect to the number N of simultaneous transmissions, when each transmitter attempts with probability q.

communication channels, where the estimator remotely tracks the time-varying state of multiple sources. We demonstrate, with an example, that the (communication) cost associated with coordination between distributed transmitters increases super-linearly with respect to the number of transmitters.

- We study the information update policies that make decisions of *when* and *which* source information should be transmitted to the estimator. The update decisions can be triggered either by the distributed transmitters based on exact local state-information or by the receiver based on estimated global state-information assuming that system parameters are known a priori.
- We extend the results to a scenario where system parameters are unknown, and develop learning-based update policies employing the Upper Confidence Bound (UCB) technique from the (stochastic) Multi-Armed Bandit (MAB) literature [18]. Through numerical simulations, we show that our learning-based update policies asymptotically achieve the minimum cost levels.

The rest of paper is organized as follows. In Section II, we describe the system model and formulate the problem. In Sections III and IV, we study information update policies triggered by the distributed transmitters and by the receiver, respectively, when system parameters are known. In Section V, we compare the performance of the proposed update policies and extend them to the scenarios where the system parameters are unknown. In Section VI, we verify our analysis results through numerical simulations. In Section VII, we conclude our work.

II. System Model

We present our system model with n information sources (e.g., sensors) and one remote estimator (e.g., sink or collector), where the estimator remotely tracks the time-varying state of the sources through shared wireless channels, as shown in Fig. 1. We describe the cost models of information mismatch and update communication, and then formulate our problem. We use the terms of sensor and transmitter, interchangeably, and similarly for the terms of estimator and receiver.

A. Value of Information

We consider a time-slotted system. At each time t, the state of each source changes following a random walk process. Specifically, let $w_i(t)$ be an independent and identically distributed (i.i.d) random process with distribution as

$$w_i(t) = \begin{cases} 1, & \text{with probability } p_i, \\ 0, & \text{with probability } 1 - 2p_i, \\ -1, & \text{with probability } p_i, \end{cases} \tag{1}$$

for some $p_i \in [0, 0.5]$. The parameter p_i is known to the receiver[1]. This simple non-biased, scalar-valued model not only captures the essential aspect of the problem, but also can be converted to a biased case by adding a constant drift. Let $x_i(t)$ denote the state of source i at the beginning of time t, which is a random walk process associated with $w_i(t)$ as

$$x_i(t+1) = x_i(t) + w_i(t), \text{ for } t \geq 0,$$

with initial state $x_i(0)$.

Let $u_i(t) \in \{0, 1\}$ denote an update decision of transmitter i in time t, where $u_i(t) = 1$ implies that transmitter i updates the receiver at time t. At the end of time slot $t-1$, the update decision $u_i(t)$ can be made either in a decentralized manner by each transmitter or in a centralized manner by the receiver, based on their own observations up to time $t-1$. The detailed explanation will be made in Section II-D. Then the estimated state of source i at the receiver at time t, denoted by $\hat{x}_i(t)$, evolves as

$$\hat{x}_i(t) = \begin{cases} x_i(t), & \text{when } u_i(t) = 1, \\ \hat{x}_i(t-1), & \text{when } u_i(t) = 0. \end{cases}$$

Let $\varepsilon_i(t)$ denote the information mismatch (or error) between $x_i(t)$ and $\hat{x}_i(t)$, i.e., $\varepsilon_i(t) = |x_i(t) - \hat{x}_i(t)|$. Let $f(\varepsilon)$ be a penalty function, which increases with respect to the error ε. In this paper, we consider the mean squared error (MSE):

$$f(\varepsilon_i(t)) = |x_i(t) - \hat{x}_i(t)|^2.$$

[1] Later in section V-B, we will address the case when the parameter is unknown and has to be learned.

B. Cost of Communication

Let N_t denote the number of transmitters that take the update action simultaneously during time slot t, i.e., $N_t = \sum_{i=1}^{n} u_i(t)$. We assume that the communication cost c_t is a function of N_t, which may represent energy consumption, protocol overhead, delay, etc. In particular, we pay attention to the cost associated with coordination between the transmitters, since multiple distributed transmitters should communicate over shared channels.

In order to understand the property of the cost function, suppose that N transmitters access the shared channels as Slotted ALOHA: each time slot t is divided into mini-slots, in which N transmitters independently transmit an update packet with an identical probability q. At a mini-slot, if a transmission is successful (i.e., there is only one transmitter who attempts in the mini-slot), the corresponding transmitter receives an ACK by the end of the mini-slot, and stops transmitting in the subsequent mini-slots (by the end of the time slot). With a larger number of transmitters, the level of contention increases and there is a higher chance of collision, which occurs when multiple transmitters attempt their transmission at the same mini-slot. Once a collision occurs, no transmission is successful in the mini-slot. This implies that, on average, a transmitter makes more attempts in a time slot for a successful transmission as N increases. Suppose that a transmission consumes a unit energy (or power). Then average energy cost for a successful transmission, that can be considered as the communication cost per an update of a source, will increase with N. As an example, we simulate 50 transmitters with the Slotted ALOHA protocol, and measure average update cost of N sources when a transmission consumes 1 unit of energy. The results with different values of N and q are shown in Fig. 2, where $q = g(N)$ implies that we use the best q given N that is empirically found. We can observe that the setting of fixed q results in an exponential increase of average cost with N. The minimum cost can be achieved when q is properly set according to N.

Based on the observation, we model the total update cost as a function of N_t, the number of simultaneous transmission at time slot t, in the following exponential form:

$$c_t = cN_t^{1+\epsilon}, \tag{2}$$

for some constant $c > 0$ and nonlinearity coefficient $\epsilon \geq 0$. We will discuss later about the nonlinearity coefficient of the policies with centralized or decentralized update paradigm.

C. Problem formulation

Considering the aforementioned costs, the *per-source* cost associated with source i at each time t, under policy π, can be written as

$$C^{\pi}_{i,\epsilon}(c,t) = u_i(t)cN_t^{\epsilon} + (1 - u_i(t))f(\varepsilon_i(t)).$$

Suppose that $\mathbf{x}(0) = \hat{\mathbf{x}}(0)$. Our objective is to find an update policy π that minimizes the expected average cost over infinite time horizon:

$$\text{minimize } g_{\pi}(n, c, \epsilon),$$

where

$$g_{\pi}(n,c,\epsilon) = \mathbb{E}\left[\lim_{s\to\infty} \frac{1}{sn} \sum_{t=1}^{s} \sum_{i=1}^{n} C^{\pi}_{i,\epsilon}(c,t)\right].$$

In this work, we focus on the case of symmetric transmitters with $p_i = p$ for all i.

D. Decentralized and Centralized Update Paradigms

We organize our investigation under two fundamentally different paradigms: that of decentralized and centralized update policies. These can also be named transmitter-driven (TD) and receiver-driven (RD) paradigms, respectively, since the update decisions are triggered by each transmitter under the former one, while the update decisions are triggered by the receiver under the latter one. Also, they can be named event-driven and time-driven paradigms, respectively, since the former one triggers an update when the *error* since the last update becomes high enough, while the latter one triggers an update when the *time* since the last update is long enough.

Under a TD policy, each transmitter independently makes individual decision with the knowledge of its error $\varepsilon_i(t)$, but without the knowledge of the other's actions, e.g., the number N_t of transmitters in time slot t. On the other hand, under a RD policy, the receiver can decide on the update actions collectively (thus the set of transmitters at time slot t is under control), but it does not know current errors $\varepsilon_i(t)$. Intuitively when the update cost c_t is relatively small (i.e., N_t and ϵ are relatively small), the error cost $f(\varepsilon_i(t))$ dominates the update cost cN_t^{ϵ} and thus a TD policy may outperform a RD policy. However, when the update cost c_t is sufficiently large, the update cost starts dominating the error cost, and thus a RD policy will outperform a TD policy.

We consider two different types of TD policies based on their level of coordination: one type is for transmitters with only local information of p and c (called as TD-L policy), and the other type is for transmitters with global information of n and ϵ as well as the local information of p and c (called as TD-G policy). The information and control available for each policy is summarized in Table I.

Let ϵ_s and ϵ_r denote the nonlinearity coefficient for the TD update policy and the RD update policy, respectively. Note that the TD policy often aims for distributed operations while the RD policy is by nature centralized. Since, in general, the coordination in a distributed system is often more costly than a centralized counterpart, it is reasonable to assume $\epsilon_s \geq \epsilon_r \geq 0$. The objective of this work is to study TD and RD policies to minimize the expected average cost over infinite horizon and to compare their performance given the number n of transmitters and the nonliearity parameters ϵ_s and ϵ_r.

III. Decentralized Update Paradigm

We first investigate the behavior of TD policy, under which each transmitter decides distributedly at each time slot whether it updates the receiver or not.

TABLE I: Information and control available to the policies.

Policy	TD-L	TD-G	RD
local parms.	p, c	p, c	p, c
global parms.	$-$	n, ϵ_s	n, ϵ_r
error	$\varepsilon_i(t)$	$\varepsilon_i(t)$	$-$
controller	transmitter i	transmitter i	receiver
control var.	$u_i(t)$	$u_i(t)$	$u_1(t), \dots, u_n(t)$

In a single source system with constant per-transmission cost, it has been shown in [13] that an optimal update policy is *of threshold type* in the forms of

$$u^*(t) = \begin{cases} 0, & \text{if } \varepsilon(t) + w(t) < \gamma, \\ 1, & \text{if } \varepsilon(t) + w(t) \geq \gamma, \end{cases} \tag{3}$$

with some threshold $\gamma > 0$, where we omit subscript i for notational convenience. Then, given constant per-transmission update cost $\bar{c}$, it was shown that the expected average cost $h(\bar{c})$ over infinite time horizon can be obtained as

$$h(\bar{c}) = \frac{2}{\gamma^2}\left(\bar{c}p + \sum_{i=1}^{\gamma-1} f(i)(\gamma - i)\right), \tag{4}$$

and, for the MSE $f(\varepsilon) = \varepsilon^2$, the optimal threshold that minimizes (4) is $\gamma^* = \lfloor \sqrt[4]{12p\bar{c}} \rfloor$ or $\lceil \sqrt[4]{12p\bar{c}} \rceil$.

Inspired by these results, we consider a TD policy where each transmitter updates the receiver with a threshold γ_i. Then the number N_t of simultaneous transmissions at time slot t is a random variable, and thus the per-transmission cost is also a random variable from (2). To characterize the performance of the TD policy, we study asymptotic behavior of N_t, which will lead to our understanding of the expected average cost $g_{TD}(n, c, \epsilon_s)$.

Note that since each transmitter independently updates the receiver, we can consider the error $\varepsilon_i(t)$ for transmitter i as an independent renewal process because it is reset to 0 upon every update. An inter-renewal distribution is called *arithmetic* if inter-renewal intervals are integer multiples of some real number, and the span of an arithmetic distribution is defined as the largest number ρ such that this property holds. Then, the following theorem provides an asymptotic behavior of renewal probabilities.

Theorem 3.1 (Theorem 4.6.2 in [19]): If an inter-renewal distribution is arithmetic with span ρ, then

$$\lim_{t\to\infty} \mathbb{P}(\text{Renewal at } t\rho) = \frac{\rho}{\mathbb{E}[T]},$$

where T denote the inter-renewal interval.

The renewal process $\varepsilon_i(t)$ is arithmetic with span $\rho = 1$ since inter-renewal intervals can be $\gamma, \gamma+1, \gamma+2, \dots$. Also, the expectation of the inter-renewal interval under the threshold-type update policy with a threshold γ is known as $\mathbb{E}[T] = \frac{\gamma^2}{2p}$ [13]. Hence, we can obtain that $\lim_{t\to\infty} \mathbb{P}(\text{Renewal at } t) = \lim_{t\to\infty} \mathbb{P}(u(t) = 1) = \frac{1}{\mathbb{E}[T]} = \frac{2p}{\gamma^2}$.

Combined with the independence of the renewal processes, Theorem 3.1 can be used to characterize the asymptotic behavior of N_t.

Lemma 3.1: When n independent (symmetric) transmitters update the receiver with the same threshold γ, the number N_t of simultaneous transmissions at time slot t converges in distribution to a Binomial distribution with parameters n and $\frac{2p}{\gamma^2}$, i.e.,

$$\lim_{t\to\infty} N_t \sim B\left(n, \frac{2p}{\gamma^2}\right),$$

where $B(\cdot, \cdot)$ denotes the Binomial distribution.

Lemma 3.1 can be shown using the independence of the transmitters' decision $u_i(t)$ and Theorem 3.1. We refer to Appendix A for the proof.

Let $\tilde{g}_{\text{TD}}(\gamma, i)$ denote the expected average cost of a TD policy of source i with threshold γ, and let $N = \lim_{t\to\infty} N_t$. Replacing $\bar{c}$ in (4) with the per-transmission cost, and from Lemma 3.1, it can be obtained as

$$\begin{aligned} \tilde{g}_{\text{TD}}(\gamma, i) &= \sum_{k=1}^{n} \mathbb{P}(N = k \mid u_i(t) = 1)\, h(ck^{\epsilon_s}) \\ &= \sum_{k=0}^{n-1} \mathbb{P}\left(k; n-1, \frac{2p}{\gamma^2}\right) h(c(k+1)^{\epsilon_s}), \\ &= \mathbb{E}[h(c(K+1)^{\epsilon_s})], \end{aligned} \tag{5}$$

where $\mathbb{P}(k; n, q)$ is the probability that $N = k$ when $N \sim B(n, q)$, and K is a random variable that follows $B\left(n-1, \frac{2p}{\gamma^2}\right)$. Due to the symmetry, this holds for all i, and we can write $\tilde{g}_{\text{TD}}(\gamma, i) = \tilde{g}_{\text{TD}}(\gamma)$ for all i.

For the TD policies with local information (TD-L), each transmitter i optimizes its threshold γ agnostic about other transmitters, which results in $\gamma_L^* = \lfloor \sqrt[4]{12pc} \rfloor$ or $\lceil \sqrt[4]{12pc} \rceil$ that leads to the expected average cost

$$g_{\text{TD-L}}(n, c, \epsilon_s) = \tilde{g}_{\text{TD}}(\gamma_L^*).$$

On the other hand, for the TD policies with global information (TD-G), the transmitters can minimize (5) by further optimizing their threshold γ with respect to n and ϵ_s, which results in the expected average cost

$$g_{\text{TD-G}}(n, c, \epsilon_s) = \min_{\gamma \geq 0} \tilde{g}_{\text{TD}}(\gamma).$$

IV. Centralized Update Paradigm

In contrast to the TD policy where each transmitter can trace the error, the RD policy cannot access the error information. Thus, the receiver should make a decision based on estimation of the current error for each source. Due to the renewal property of the error, this results in time-periodic updates.

A. Expected Error

We first study how the expected error between a source and the estimator evolves over time. We omit subscript i for notational convenience. Let s denote the time elapsed since a transmitter updates the receiver. Then, there are $2s+1$ possible error states of the source (i.e., $x - \hat{x} \in [-s, s]$). Let $\mathbf{e}_s = [e_s(-s), ..., e_s(0), ..., e_s(s)]$ denote the expected error vector when the receiver is not updated by the transmitter for s consecutive time slots, where $e_s(k)$ is the probability that the error between the receiver and the source is k (i.e., $x - \hat{x} = k$). According to (1), the expected error evolves by Bayes rule as

$$e_s(k) = e_{s-1}(k)(1-2p) + (e_{s-1}(k-1) + e_{s-1}(k+1))p, \tag{6}$$

for $k \in \{-s, ..., s\}$, where $e_{s-1}(-s-1) = e_{s-1}(-s) = e_{s-1}(s) = e_{s-1}(s+1) = 0$.

Let $\xi(s)$ denote the expected error cost when the receiver has not been updated from the source for s consecutive time slots, i.e.,

$$\xi(s) = \sum_{k=-s}^{s} e_s(k) f(|k|). \tag{7}$$

In order to obtain the expected error cost $\xi(s)$, the receiver needs to update the expected error $e_s(k)$ for $k \in \{-s, .., s\}$ every time slot, and the memory for storing the expected error increases linearly with respect to s. However, in the special case of the mean squared error penalty function $f(\varepsilon) = \varepsilon^2$, the expected error cost $\xi(s)$ can be obtained without computing $\mathbf{e}_s$ as the following.

Lemma 4.1: If $f(\varepsilon) = \varepsilon^2$, the expected error cost after s consecutive time slots since the last update is

$$\xi(s) = 2ps, \text{ for } s \geq 1.$$

Lemma 4.1 can be shown by induction. We refer to Appendix B for the proof.

B. Single-transmitter Scenario

As a first step, we obtain an optimal solution to single-transmitter problem with the MSE penalty function. We first describe a discrete-time Markov Decision Process (MDP), whose state at time slot t is represented by s. For each state, the receiver has two possible actions: update $u = 1$ or not update $u = 0$. If $u = 0$, then the state evolves to $s + 1$. If $u = 1$, the state evolves to 1. If the per-transmission cost is $\bar{c}$, the expected cost under state s and action u is given as $u\bar{c} + (1 - u)\xi(s)$. Then, we have the following Bellman equation:

$$\phi(s) = \min\{\xi(s) + \phi(s+1) - \lambda, \bar{c} + \phi(1) - \lambda\}, \tag{8}$$

where $\xi(s) = 2ps$, $\phi(\cdot)$ denotes the cost-to-go function, and λ denotes the minimum expected average cost over infinite time horizon [20].

We show in Lemma 4.2 that an optimal policy that solves (8) is also of threshold type.

Lemma 4.2: There exists a threshold policy that solves the Bellman optimal equations (8). Specifically, given constant update cost $\bar{c}$, the policy has a real-valued time threshold $\tau^*(\bar{c})$ such that

$$u^*(s) = \begin{cases} 0, & \text{if } s < \tau^*(\bar{c}), \\ 1, & \text{if } s \geq \tau^*(\bar{c}). \end{cases} \tag{9}$$

Lemma 4.2 can be shown [13] using the fact that $\xi(s)$ is increasing in s when $f(\varepsilon) = \varepsilon^2$. We refer to Appendix C. Note that unlike the optimal TD policy (3), the optimal RD policy has a *time threshold* with periodic updates.

Given time threshold τ, the expected average cost $g_{\text{RD}}(\tau)$ under the RD update policy is given by

$$g_{\text{RD}}(\tau) = \frac{1}{\tau}\left(\bar{c} + \sum_{s=1}^{\tau-1} \xi(s)\right). \tag{10}$$

For $f(\varepsilon) = \varepsilon^2$, we have $\xi(s) = 2ps$ and thus $g_{\text{RD}}(\tau) = \bar{c}/\tau + p(\tau-1)$, which is convex in $\tau > 0$. Thus, by solving $\frac{dg_{\text{RD}}}{d\tau} = 0$, we can obtain a closed-form expression of an optimal time threshold that solves (8): $\tau^*(\bar{c}) = \lfloor\sqrt{\bar{c}/p}\rfloor$ or $\lceil\sqrt{\bar{c}/p}\rceil$.

Note that, under the TD policy with a single transmitter, we have the expected update interval $\mathbb{E}[T] = \sqrt{3\bar{c}/p}$. That is, the RD policy updates the receiver more frequently than the TD policy on average. This is because the controller does not use the error $\varepsilon(t)$ and thus it compensates for the lack of information by updating more frequently.

C. Multi-transmitter Scenario

Now, suppose that n (symmetric) transmitters update the receiver. Unlike the TD policy where each transmitter independently updates the receiver and thus the number of simultaneous transmissions at a given time is random, an RD policy can control the number of simultaneous transmissions so that the communication cost is not too high. Since we are considering symmetric transmitters with $p_i = p$ for all i, the update periods (i.e, the time thresholds) are the same for all transmitters.

Let τ_{n,ϵ_r} denote the update period. The receiver can optimize τ_{n,ϵ_r} by taking into account n and ϵ_r, and control the transmissions by assigning a time slot to each transmitter.

- When $n \leq \tau_{n,\epsilon_r}$, an optimal policy lets each transmitter i update at time slot t such that $(t \mod \tau_{n,\epsilon_r}) = i$ and there is at most one transmission at each time slot[2]. There are $\tau_0 = \tau_{n,\epsilon_r} - n$ idle (i.e., no-update) time slots within the update period τ_{n,ϵ_r}.
- When $n > \tau_{n,\epsilon_r}$, an optimal policy lets each transmitter i update at time slot t such that $(t \mod \tau_{n,\epsilon_r}) = (i \mod \tau_{n,\epsilon_r})$. Then, at each time slot, there are $\lceil\frac{n}{\tau_{n,\epsilon_r}}\rceil$ transmissions or $\lceil\frac{n}{\tau_{n,\epsilon_r}}\rceil - 1$ transmissions. Let $k_{n,\epsilon_r} = \lceil\frac{n}{\tau_{n,\epsilon_r}}\rceil$, and let τ_0 denote the number of time slots where $\lceil\frac{n}{\tau_{n,\epsilon_r}}\rceil - 1$ transmitters update the receiver within an update period τ_{n,ϵ_r}. The structure of an optimal RD policy given n and τ_{n,ϵ_r} is shown by Fig.3. Each slot on the x-axis represents one time slot, and each bin on the y-axis represents one transmission opportunity. The number in each bin is the index of transmitters of {1,2,…,n}. Note that each of the first $(\tau_{n,\epsilon_r} - \tau_0)$ time slots on a period has k_{n,ϵ_r} simultaneous transmissions, and yields the total cost of $c(k_{n,\epsilon_r})^{\epsilon_r}$. Each of the rest τ_0 time slots has $k_{n,\epsilon_r} - 1$ simultaneous transmissions, and the cost of $c(k_{n,\epsilon_r} - 1)^{\epsilon_r}$.

Let $\tilde{g}_{\text{RD}}(\tau_{n,\epsilon_r})$ denote the expected average cost given τ_{n,ϵ_r}, which is given by

$$\begin{aligned} \tilde{g}_{\text{RD}}(\tau) &= \tfrac{k(\tau-\tau_0)}{n\tau}(ck^{\epsilon_r} + p(\tau-1)\tau) \\ &\quad + \tfrac{(k-1)\tau_0}{n\tau+}(c(k-1)^{\epsilon_r} + p(\tau-1)\tau) \\ &= \tfrac{c}{n\tau}((\tau-\tau_0)k^{1+\epsilon_r} + \tau_0(k-1)^{1+\epsilon_r}) + p(\tau-1), \end{aligned} \tag{11}$$

where we omit subscripts n and ϵ_r for notational convenience (i.e., $k = k_{n,\epsilon_r}$. Note that $\tau = \tau_{n,\epsilon_r}$) and $k = \lceil n/\tau \rceil$ and

[2]This policy is optimal in the sense that there is no other policy that can make communication cost smaller, given n and τ_{n,ϵ_r}. Note that the expected cost associated with error is determined by a threshold τ_{n,ϵ_r}.

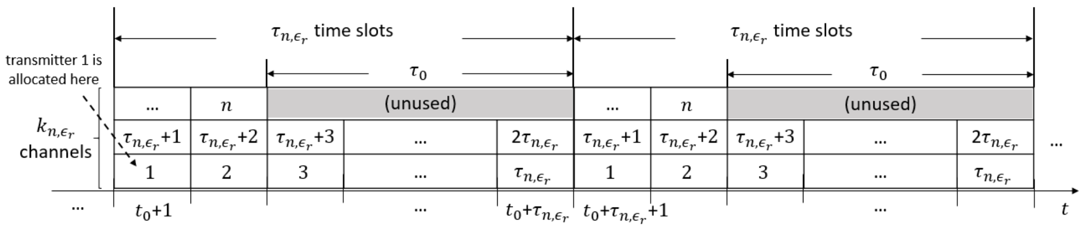


Fig. 3: Time slot and channel allocation of the RD policy.

$\tau_0 = k\tau - n$. The expected average cost, $g_{\text{RD}}(n, c, \epsilon_r)$, under the RD policy is given by

$$g_{\text{RD}}(n, c, \epsilon_r) = \min_{\tau \geq 1} \tilde{g}_{\text{RD}}(\tau). \tag{12}$$

V. Performance Comparison of Decentralized and Centralized Update Paradigms

In this section, we compare the performance of the two TD policies and the RD policy. Furthermore, we will extend our design to a scenario where system parameters are unknown.

A. TD-L Policy vs. TD-G Policy vs. RD Policy

We first consider the single-transmitter case of $n = 1$, in which TD-L is equivalent to TD-G. If $c < 2p$, then the optimal policy is to update at every time slot and we have $g_{\text{TD}}(1, c, \epsilon_s) = g_{\text{RD}}(1, c, \epsilon_r) = c$. Suppose that $c \geq 2p$. Then, from (10) with $\tau = \sqrt{c/p}$ and (4) with $\gamma = \sqrt[4]{12pc}$, we have

$$g_{\text{RD}}(1, c, \epsilon_r) = 2\sqrt{pc} - p \geq \frac{2}{\sqrt{3}}\sqrt{pc} - \frac{1}{6} = g_{\text{TD}}(1, c, \epsilon_s),$$

where $g_{\text{TD}} = g_{\text{TD-L}} = g_{\text{TD-G}}$ and the inequality comes from that $c \geq 2p$.

Not only this confirms the expected superiority of TD updates to RD updates for the single-transmitter case, but also reveals that the performance improvement is a function of system parameters p and c. Note that when $\epsilon_s = \epsilon_r = 0$, each transmitter pays the same per-transmitter cost c regardless of the number of simultaneous transmissions. Hence, we have $g_{\text{RD}}(n, c, 0) = g_{\text{RD}}(1, c, \epsilon_r) \geq g_{\text{TD}}(1, c, \epsilon_s) = g_{\text{TD}}(n, c, 0)$, i.e., TD policies always outperforms RD policy.

Now, we consider when $\epsilon_r > 0$, $\epsilon_s > 0$ and $n \gg 1$. Theorem 5.1 shows the asymptotic behavior of $g_{\text{TD-L}}$, $g_{\text{TD-G}}$ and g_{RD}, under the assumption that $c \geq 2p$.

Theorem 5.1: Under TD-L and TD-G policies, we have the asymptotic lower bounds such that

$$g_{\text{TD-L}}(n, c, \epsilon_s) = \Omega(n^{\epsilon_s}),$$

and

$$g_{\text{TD-G}}(n, c, \epsilon_s) = \Omega\left(n^{\frac{\epsilon_s}{\epsilon_s+2}}\right),$$

respectively, for $\epsilon_s > 0$. Under RD policy, we have an asymptotic upper bound such that

$$g_{\text{RD}}(n, c, \epsilon_r) = O\left(n^{\frac{\epsilon_r}{\epsilon_r+2}}\right)$$

for $\epsilon_r > 0$.

We refer to Appendix D for the detailed proof.

Since $g_{\text{RD}}(1, c, \epsilon_r) \geq g_{\text{TD-L}}(1, c, \epsilon_s) = g_{\text{TD-G}}(1, c, \epsilon_s)$ for any ϵ_s and ϵ_r, Theorem 5.1 implies that there is a crossing point where RD policy starts to outperform TD-L policy for $\epsilon_s \geq \epsilon_r > 0$ and TD-G policy for $\epsilon_s > \epsilon_r > 0$. In other words, changing the strategy depending on parameters n, ϵ_s and ϵ_r for some given p and c improves the system performance. In particular, when the value of information dominates the cost of communication, i.e., when n and $\epsilon_s - \epsilon_r$ are relatively small, it is better to use TD policies. On the other hands, when the cost of communication dominates the value of information, it is better to use RD policy.

B. Learning-based update policy

In this subsection, we consider scenarios where system parameters p, c, and ϵ_s (or ϵ_r) are *unknown*. We assume that the upper bounds of thresholds $\overline{\gamma}$ and $\overline{\tau}$ are given for TD and RD policies, respectively. In the following, we develop learning-based TD and RD policies employing the Multi-Armed Bandit (MAB) technique by considering each possible threshold as an arm[3].

Learning-based TD policy[4]: Let s_j denote the time at which the j^{th} update occurs with $s_0 = 0$. Let $\Delta_j := s_{j+1} - s_j$ denote the j^{th} update interval. In the learning-based algorithm, an update interval is a round for learning, where we apply the Upper Confidence Bound (UCB) technique [18]. At the beginning of the j^{th} interval, we select threshold γ_j and terminate the round when the error gets larger than the threshold according to (3). At the end of the interval, the average cost $\hat{r}_j$ during interval j can be written as

$$\hat{r}_j = \tfrac{1}{\Delta_j}\left(\textstyle\sum_{t=s_j+1}^{s_{j+1}} \varepsilon^2(t) + c\left(\sum_{i=1}^{n} u_i(s_{j+1})\right)^{\epsilon_s}\right).$$

For each threshold $\gamma \in [0, \overline{\gamma}]$, we store the empirical average cost and the number of selections up to now as $\hat{r}(\gamma)$ and $\eta(\gamma)$, respectively.

We run the following procedure independently for each transmitter i (subscript i is omitted for brevity). For the first $\overline{\gamma}+1$ update intervals, the transmitter selects $\gamma \in [0, \overline{\gamma}]$ exactly once. For each interval $j > \hat{\gamma} + 1$, it decides an action according to the following procedure.

- At the beginning of the j^{th} update interval:
  1) For each γ, $I(\gamma) \leftarrow \frac{\hat{r}(\gamma)}{\max_{\gamma'} \hat{r}(\gamma')} - \sqrt{\frac{2\log(j)}{\eta(\gamma)}}$.

[3]Since time and state space are discrete, we can employ the finite-armed Multi-Armed Bandits.

[4]Each transmitter follows the proposed procedure independently, thus we omit subscripts indicating the indices of transmitters.

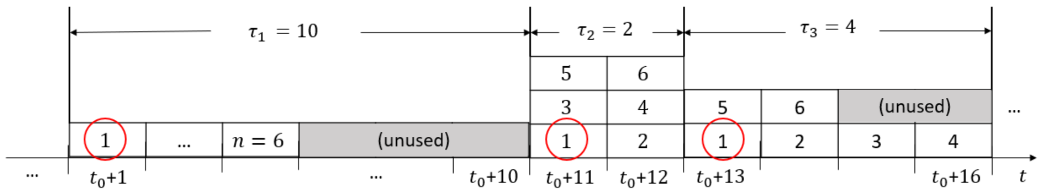


Fig. 4: Example of varying thresholds: Transmitter 1 marked with the circle has the same update intervals with the given thresholds.

2) $\gamma_j \leftarrow \arg\min_\gamma I(\gamma)$.

- When an update occurs and the interval ends:

1) $\eta(\gamma_j) \leftarrow \eta(\gamma_j) + 1$.

2) $\hat{r}(\gamma_j) \leftarrow \hat{r}(\gamma_j)\left(1 - \frac{1}{\eta(\gamma_j)}\right) + \frac{\hat{r}_j}{\eta(\gamma_j)}$.

Note that, for each possible threshold γ, the empirical average cost $\hat{r}(\gamma)$ can be greater than 1. Thus, when the UCB index $I(\gamma)$ is calculated, we normalize the empirical costs with the maximum value among them so that the values lie between 0 and 1.

Learning-based RD policy: Now, we develop the learning-based RD policy by employing the UCB technique. The receiver learns an optimal threshold τ^* among the possible threshold $\tau \in [1, \overline{\tau}]$. Let τ_j denote the threshold of the j^{th} interval. Since it is a time threshold, the j^{th} interval has exactly τ_j time slots. At the beginning of the j^{th} interval, the receiver collectively decides when source i will be updated in the interval as shown in Fig. 3. There is a problem though. Since the threshold changes across two consecutive intervals, in the perspective of individual source i, the update interval becomes somewhat arbitrary. For example, in Fig. 4, the update interval of source 4 is $(t_0 + 12) - (t_0 + 4) = 8$ and $(t_0 + 16) - (t_0 + 12) = 5$ when τ_j changes from 10 to 2 and 4. Thus, for the purpose of learning an optimal threshold, the receiver traces the empirical average cost of transmitter 1 only, since transmitter 1 has consistent update interval with τ_j.

As in the learning-based TD policy, the average cost during interval j is written as (13) replacing ϵ_s with ϵ_r. In the RD policy, the update interval Δ_j equals τ_j. Let $\hat{r}(\tau)$ and $\eta(\tau)$ denote the empirical average cost (of transmitter 1) for τ and the number of selections for τ, respectively. The learning-based RD policy is operated as in the learning-based TD policy by replacing γ with τ.

We verify the performance of learning-based TD and RD policies through simulations in Section VI.

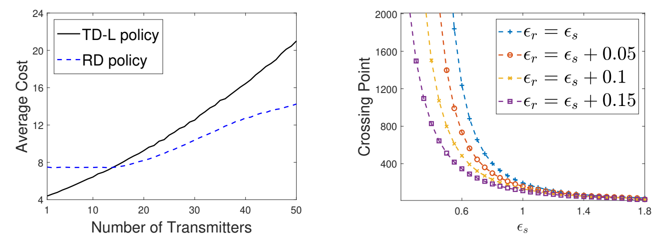
VI. Simulation Results

In this section, we compare the performance of TD-L, TD-G and RD policies through simulations. Throughout the simulations, we use $p = 0.3$ and $c = 50$.

It is obvious that TD-G policy outperforms TD-L policy for all $\epsilon_s \geq 0$ since TD-G policy uses more information than TD-L policy, and thus we do not compare between TD-L and TD-G policies. For numerical simulations, we use thresholds $\gamma_L^* = \lfloor \sqrt[4]{12pc} \rfloor$ or $\lceil \sqrt[4]{12pc} \rceil$ for TD-L policy, γ_G^* that minimizes (5) for TD-G policy, and τ^* that minimizes (12) for RD policy[5]. Based on the given threshold, each transmitter either updates the receiver ($u_i(t) = 1$) or not ($u_i(t) = 0$) at every time slot t. Then, the average cost $C(t)$ at time slot t is

$$C(t) := \frac{1}{tn}\sum_{s=1}^{t}\sum_{i=1}^{n}\left(\varepsilon_i^2(t) + u_i(t)\cdot c\left(\sum_{i=1}^{n} u_i(t)\right)^{\epsilon}\right),$$

where $\epsilon = \epsilon_s$ for TD-L and TD-G policies and $\epsilon = \epsilon_r$ for RD policy.



(a) Average cost when $\epsilon_s = \epsilon_r = 2$.

(b) Crossing point with respect to ϵ_s and ϵ_r.

Fig. 5: Performance comparison between TD-L and RD policies.

We first compare TD-L and RD policies. We run simulations for $T = 10^4$ time slots, and the results are averaged over 50 repetitions. Fig. 5(a) shows the average cost $C(T)$ at time T when $\epsilon_s = \epsilon_r = 2$. We observe that for a relatively small n, TD-L policy outperforms RD policy. However, as n increases, the gap becomes close to zero and eventually RD policy outperforms TD-L policy. We call the point (the number of transmitters) where RD policy starts to outperform TD-L policy as a *crossing point*. In Fig. 5(a), the crossing point is at 14. Fig. 5(b) shows the crossing point with respect to ϵ_s and ϵ_r. As expected, for relatively large n and ϵ_s, the communication cost of distributed updates dominates the value of (state) information, and the value of information dominates the update cost for small n and ϵ_s.

Now, we compare TD-G and RD policies. Note that, according to Theorem 5.1, the existence of a crossing point between TD-G and RD policies can be guaranteed only for $\epsilon_s > \epsilon_r > 0$ Fig. 6 shows the ratio of the average cost of RD policy to that of TD-G policy when $\epsilon_s = \epsilon_r = \epsilon$. The ratio greater than 1 implies TD-G policy outperforms

[5]Thresholds γ_G^* and τ^* can be found using numerical search methods.

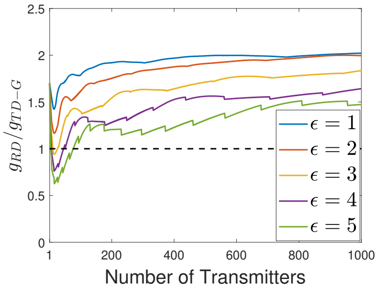
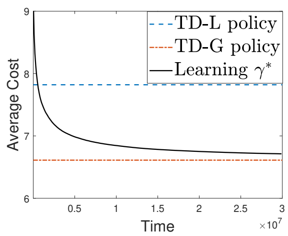
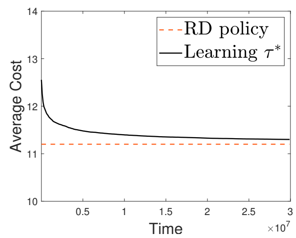


Fig. 6: Performance comparison between TD-G and RD policies when $\epsilon = \epsilon_s = \epsilon_r$.



(a) The learning-based TD policy.



(b) The learning-based RD policy.

Fig. 7: Performance of the learning-based policies.

RD policy. As a specific example, when $\epsilon = 1$, from (15) and (17), we can analytically show that $\lim_{n\to\infty} \frac{g_{\text{RD}}}{g_{\text{TD-G}}} \geq \lim_{n\to\infty} \frac{(4/48^{2/3}+48^{1/3}/6)(p^2cn)^{1/3}-1/6}{(p(2c/p)^{1/3}+c(p/2c)^{2/3})n^{1/3}} \approx 2.08$, which agrees with the simulation result. This implies that when transmitters have global information, i.e., n and ϵ_s, they can adjust their threshold reflecting the distribution of N_t and this leads to significant improvement of the performance of TD-L policy.

Now, we evaluate the learning-based TD and RD policies, where system parameters p, c, ϵ_s and ϵ_r are unknown to both transmitters and the receiver, and n is known to the receiver but not to the transmitters. Only the range of possible values of each parameter is known, and thus each transmitter and the receiver have the set of possible thresholds $\gamma \in [0, \overline{\gamma}]$ and $\tau \in [1, \overline{\tau}]$, respectively. We set $p = 0.3$, $c = 50$, $\epsilon_s = 2$, $\epsilon_r = 1$ and $n = 50$, and assume that $\overline{\gamma} = 10$ and $\overline{\tau} = 30$, respectively. We run simulations for $T = 3 \times 10^7$ time slots.

Fig. 7(a) shows the performance of the learning-based TD policy, which is compared to TD-L and TD-G policies that operate with known system parameters. As shown in Fig. 7(a), the average cost of the learning-based TD policy rapidly approaches that of TD-G policy, which implies that the learning-based TD policies find the global optimal threshold γ_G^*. Fig. 7(b) shows the performance of the learning-based RD policy, which is also compared to RD policy with known parameters. It verifies that the learning-based RD policy finds the optimal threshold τ^* of RD policy. These findings confirm that the findings of our work can be effectively translated into the learning environment where system parameters as well as value and cost functions are unknown.

VII. Conclusion

We investigated decentralized (transmitter-driven) and centralized (receiver-driven) update paradigms, where a receiver is updated from multiple sources of which states evolve according to a simple random walk process. In particular, we considered a scenario where each update is accompanied by communication cost, and we modeled communication cost as a superlinear function of the number of simultaneous transmissions at a given time since the transmitters communicate over shared channels. When the cost associated with the information mismatch (error) is the mean squared error, we obtained the expected average cost for the transmitter-driven and receiver-driven policies, and compared them for different number of transmitters. From the comparison, we provided insights into the tradeoff between the value of fresh information and the cost of distributed communication in the remote tracking of large-scale distributed systems. We also developed learning-based policies that asymptotically achieve the minimum costs attained by the optimal policies when the system parameters are unknown. Finally, through numerical simulations, we verified the performance of the proposed policies. Theoretical analysis of the performance of learning-based update policies is an interesting future work in consideration that each transmitter has different update periods. Another interesting future work is to study the case when each transmitter has different dynamics.

Appendix

A. Proof of Lemma 3.1

By the independence of the transmitters' decision $u_i(t)$ and Theorem 3.1, we have

$$\begin{aligned}\lim_{t\to\infty} \mathbb{P}(N_t = k) &= \lim_{t\to\infty} \mathbb{P}\left(\textstyle\sum_{i=1}^n u_i(t) = k\right) \\ &= \lim_{t\to\infty} \binom{n}{k} \mathbb{P}(u(t) = 1)^k \mathbb{P}(u(t) = 0)^{n-k} \\ &= \binom{n}{k} \left(\frac{1}{\mathbb{E}[T]}\right)^k \left(1 - \frac{1}{\mathbb{E}[T]}\right)^{n-k},\end{aligned}$$

where $\mathbb{E}[T]$ is the expectation of the inter-renewal interval under the threshold-type update policy with a threshold γ, which is $\frac{2p}{\gamma^2}$ [13].

B. Proof of Lemma 4.1

We prove Lemma 4.1 by induction. For $\tau = 0$, we have $e_0(0) = 1$ and $\xi(0) = 0$. For $\tau = 1$, we have $\mathbf{e}_1 = [e_1(-1), e_1(0), e_1(1)] = [p, 1-2p, p]$ and $\xi(1) = 2p$. Now assume the induction hypothesis that

$$\xi(\tau - 1) = \sum_{k=-\tau+1}^{\tau-1} k^2 e_{\tau-1}(k) = 2p(\tau - 1) \text{ for } \tau \geq 2. \quad (13)$$

From (6) and (7), we have

$$\begin{aligned}\xi(\tau) &= \textstyle\sum_{k=-\tau}^{\tau} k^2 e_\tau(k) \\ &= (1-2p)\textstyle\sum_{k=-\tau+1}^{\tau-1} k^2 e_{\tau-1}(k) \\ &\quad + p\textstyle\sum_{k=-\tau}^{\tau-2} k^2 e_{\tau-1}(k+1) + p\textstyle\sum_{k=-\tau+2}^{\tau} k^2 e_{\tau-1}(k-1) \\ &= \xi(\tau-1) \\ &\quad + e_{\tau-1}(-\tau+1)(p\tau^2 + p(\tau-2)^2 - 2p(\tau-1)^2) \\ &\quad + \cdots + e_{\tau-1}(0)2p \\ &\quad + e_{\tau-1}(\tau-1)(p(\tau-2)^2 + p\tau^2 - 2p(\tau-1)^2).\end{aligned}$$

Since $e_{\tau-1}(0) = 1 - \sum_{k=1}^{\tau-1} e_{\tau-1}(-k) - \sum_{k=1}^{\tau-1} e_{\tau-1}(k)$, we have $\xi(\tau) = 2p\tau$.

C. Proof of Lemma 4.2

From (8), let $A(s) = \xi(s) + A(s+1) - \lambda$ and $B(s) = c + \phi(1) - \lambda$. Then, an optimal action is to update if $B(s) < A(s)$ and not to update if $B(s) \geq A(s)$. Note that $B(0) = c + \phi(1) - \lambda > \xi(0) + \phi(1) - \lambda = A(0)$ since $\xi(0) = 0$ and $c > 0$. Thus, $u = 0$ is an optimal action for $s = 0$. Note that $B(s) = c + \phi(1) - \lambda$ is a constant, and $A(s)$ is an increasing-then-decreasing function or a decreasing function since $A(s+1) - A(s) = \lambda - 2ps$.

We show that there exists τ such that $A(s) \leq B(s)$ for $s \leq \tau$ and $A(\tau+1) > B(\tau+1)$ by contradiction. Suppose that such τ does not exist, which implies that $A(s) \leq B(s)$ for all s and thus $u = 0$ for all s. Then, the expected cost goes infinity since $\xi(s) = 2ps$. If we take an update policy such that $u = 1$ for all s, then the expected cost c, which leads to a contradiction. Hence, there exist s such that $A(s) > B(s)$ and, for $\tau = \min\{s : A(s) > B(s)\}$, we have the claim since $A(s)$ is increasing-then-decreasing.

D. Proof of Theorem 5.1

We first show the asymptotic lower bound for TD policies. The expected average cost $\tilde{g}_{\text{TD}}(\gamma)$ of TD policy given threshold γ is, from (5), given by

$$\tilde{g}_{\text{TD}}(\gamma) = \frac{2}{\gamma^2}\left(pc\mathbb{E}\left[(K+1)^{\epsilon_s}\right] + \frac{\gamma^2(\gamma^2-1)}{12}\right). \tag{14}$$

TD-L policy: Let $f(x) = (x+1)^{\epsilon_s}$ and $\mu = \mathbb{E}[K] = \frac{2p(n-1)}{\gamma^2}$. By expanding the Taylor series of $f(K)$ around μ by the second-order term, we have

$$f(K) = f(\mu) + f'(\mu)(K-\mu) + \frac{f''(\alpha)(K-\mu)^2}{2}$$

for some $\alpha \in [0, n-1]$. By taking the expectation on both sides, we have

$$\mathbb{E}[(K+1)^{\epsilon_s}] = (\mu+1)^{\epsilon_s} + \frac{\mathbb{E}[f''(\alpha)(K-\mu)^2]}{2}.$$

If $\epsilon_s \geq 1$, then $f(x)$ is convex and thus $f''(x) \geq 0$ for all $x \in [0, n-1]$. Then, we have

$$\mathbb{E}[(K+1)^{\epsilon_s}] \geq \left(\frac{2p(n-1)}{\gamma^2} + 1\right)^{\epsilon_s}, \tag{15}$$

and thus $\mathbb{E}[(K+1)^{\epsilon_s}] = \Omega(n^{\epsilon_s})$ with $\gamma_L^* = \lfloor \sqrt[4]{12pc} \rfloor$ or $\lceil \sqrt[4]{12pc} \rceil$.

Now, suppose that $0 < \epsilon_s < 1$. By expanding the Taylor series of $f(K)$ around μ by the third-order term, we have

$$\begin{aligned}f(K) = f(\mu) &+ f'(\mu)(K-\mu) \\ &+ \frac{f''(\mu)(K-\mu)^2}{2} + \frac{f^{(3)}(\alpha)(K-\mu)^3}{6}\end{aligned}$$

for some $\alpha \in [0, n-1]$. By taking the expectation on both sides, we have $\mathbb{E}[(K+1)^{\epsilon_s}] =$

$$(\mu+1)^{\epsilon_s} + \frac{f''(\mu)\text{Var}(K)}{2} + \frac{\mathbb{E}[f^{(3)}(\alpha)(K-\mu)^3]}{6}.$$

Note that $f^{(3)}(x) \geq 0$ for all $x \in [0, n-1]$ since $f'(x) = \epsilon_s(x+1)^{\epsilon_s - 1}$ is convex for $\epsilon_s \in (0,1)$, and $\mathbb{E}[(K-\mu)^3] \geq 0$ since, for $X \sim B(m,q)$, $\mathbb{E}[(X - \mathbb{E}[X])^3] = mq(2q-1)(q-1)$ and in our case $q = \frac{2p}{\gamma^2} < 0.5$ since $c \geq 2p^6$. Thus, we have

$$\mathbb{E}[(K+1)^{\epsilon_s}] \geq \left(\frac{2p(n-1)}{\gamma^2} + 1\right)^{\epsilon_s} + \frac{f''(\mu)\text{Var}(K)}{2}, \tag{16}$$

where $f''(x) = \epsilon_s(\epsilon_s - 1)(x+1)^{\epsilon_s - 2}$, $\mu = \frac{2p(n-1)}{\gamma^2}$ and $\text{Var}(K) = (n-1)\left(\frac{2p}{\gamma^2}\right)\left(1 - \frac{2p}{\gamma^2}\right)$, and thus $\mathbb{E}[(K+1)^{\epsilon_s}] = \Omega(n^{\epsilon_s})$.

TD-G policy: If $\epsilon_s \geq 1$, by (14) and (15), we have

$$\begin{aligned}\tilde{g}_{\text{TD}}(\gamma) &\geq \frac{2}{\gamma^2}\left(pc\left(\frac{2p(n-1)}{\gamma^2} + 1\right)^{\epsilon_s} + \frac{\gamma^2(\gamma^2-1)}{12}\right) \\ &\geq \frac{2pc(2p(n-1))^{\epsilon_s}}{\gamma^{2\epsilon_s + 2}} + \frac{\gamma^2 - 1}{6} = \overline{g}_{\text{TD}}(\gamma).\end{aligned}$$

Since $\overline{g}_{\text{TD}}(\gamma)$ is convex in γ, by solving $\frac{d\overline{g}_{\text{TD-G}}}{d\gamma} = 0$, we have $\gamma^* = \sqrt[2\epsilon_s+4]{6pc(2\epsilon_s+2)(2p(n-1))^{\epsilon_s}}$, with which we have $g_{\text{TD-G}}(n, c, \epsilon_s) = \Omega(n^{\frac{\epsilon_s}{\epsilon_s+2}})$ since $g_{\text{TD}}(n, c, \epsilon_s) = \min_{\gamma > 0} \tilde{g}_{\text{TD}}(\gamma) \geq \min_{\gamma > 0} \overline{g}_{\text{TD}}(\gamma)$.

If $0 < \epsilon_s < 1$, by (14) and (16), we have

$$\begin{aligned}\tilde{g}_{\text{TD}}(\gamma) &\geq \frac{\gamma^2-1}{6} + \frac{2pc(2p(n-1))^{\epsilon_s}}{\gamma^{2\epsilon_s+2}}\left(1 - \frac{(n-1)}{32(\mu+1)^2}\right) \\ &\geq \frac{\gamma^2-1}{6} + \frac{2pc(2p(n-1))^{\epsilon_s}}{\gamma^{2\epsilon_s+2}}\left(1 - o(n)\right),\end{aligned}$$

where $o(n) = \frac{\gamma^4}{128p^2(n-1)}$. Suppose that, for some $\delta \in (0,1)$, there exists an N_δ such that $o(n) \leq \delta$ for all $n \geq N_\delta$. Then, for $n \geq N_\delta$, we have

$$\tilde{g}_{\text{TD}}(\gamma) \geq \frac{\gamma^2-1}{6} + \frac{2pc(2p(n-1))^{\epsilon_s}}{\gamma^{2\epsilon_s+2}}(1-\delta) = \overline{g}_{\text{TD}}(\gamma).$$

Then, by minimizing $\overline{g}_{\text{TD}}(\gamma)$, we have an optimal threshold $\gamma^* = \sqrt[2\epsilon_s+4]{(1-\delta)6pc(2\epsilon_s+2)(2p(n-1))^{\epsilon_s}}$, with which we have $o(n) = O(n^{\frac{\epsilon_s - 2}{\epsilon_s+2}})$. Since $\epsilon_s \in (0,1)$, the result accords with the assumption on $o(n)$. Hence, we have $g_{\text{TD-G}}(n, c, \epsilon_s) = \Omega(n^{\frac{\epsilon_s}{\epsilon_s+2}})$.

RD policy: Under the RD policy, the expected average cost $\tilde{g}_{\text{RD}}(\tau)$ given by τ can be bounded, from (11), as

$$\tilde{g}_{\text{RD}}(\tau) \leq \frac{ck^{1+\epsilon_r}}{n} + p(\tau - 1) = \overline{g}_{\text{RD}}(\tau), \tag{17}$$

where $k = \lceil n/\tau \rceil$. Since $g_{\text{RD}}(n, c, \epsilon_r) = \min_{\tau \geq 1} \tilde{g}_{\text{RD}}(\tau) \leq \min_{\tau \geq 1} \overline{g}_{\text{RD}}(\tau) \leq \overline{g}_{\text{RD}}(\tau')$ for any $\tau' \geq 1$, by letting $\tau' = \sqrt[\epsilon_r+2]{(1+\epsilon_r)cn^{\epsilon_r}/p}$, we have $g_{\text{RD}}(n, c, \epsilon_r) = O(n^{\frac{\epsilon_r}{\epsilon_r}+2})$.

[6]Note that, for $X \sim \mathcal{N}(\mu, \sigma^2)$, the equality holds in (16) since $\mathbb{E}[(X - \mathbb{E}[X])^3] = 0$. Since a Binomial distribution, $B(m,q)$, can be approximated by a Gaussian distribution, $\mathcal{N}(mq, mq(1-q))$, for a sufficiently large m, the gap in inequality (16) vanishes as the number n of transmitters increases.

REFERENCES

[1] S. K. Khaitan and J. D. McCalley, "Design techniques and applications of cyberphysical systems: A survey," *IEEE Systems Journal*, vol. 9, no. 2, pp. 350–365, 2015.

[2] A. Kosta, N. Pappas, V. Angelakis, *et al.*, "Age of information: A new concept, metric, and tool," *Foundations and Trends® in Networking*, vol. 12, no. 3, pp. 162–259, 2017.

[3] 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.

[4] I. Kadota, A. Sinha, and E. Modiano, "Optimizing age of information in wireless networks with throughput constraints," in *Proc. INFOCOM*, 2018.

[5] A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, "Age-optimal sampling and transmission scheduling in multi-source systems," in *MobiHoc*, 2019, pp. 121–130.

[6] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the wiener process over a channel with random delay," in *IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 321–325.

[7] T. Z. Ornee and Y. Sun, "Sampling for remote estimation through queues: Age of information and beyond," in *IEEE WiOpt*, 2019.

[8] J. Chakravorty and A. Mahajan, "Remote estimation over a packet-drop channel with markovian state," *IEEE Transactions on Automatic Control*, 2019.

[9] K. Huang, W. Liu, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "Real-time remote estimation with hybrid arq in wireless networked control," *IEEE Transactions on Wireless Communications*, 2020.

[10] X. Gao, E. Akyol, and T. Başar, "Optimal estimation with limited measurements and noisy communication," in *IEEE Conference on Decision and Control (CDC)*, 2015, pp. 1775–1780.

[11] ——, "Optimal sensor scheduling and remote estimation over an additive noise channel," in *American Control Conference (ACC), 2015*, 2015, pp. 2723–2728.

[12] G. M. Lipsa and N. C. Martins, "Remote state estimation with communication costs for first-order lti systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 9, pp. 2013–2025, 2011.

[13] J. Yun, C. Joo, and A. Eryilmaz, "Optimal real-time monitoring of an information source under communication costs," in *IEEE Conference on Decision and Control (CDC)*, 2018, pp. 4767–4772.

[14] D. Han, J. Wu, H. Zhang, and L. Shi, "Optimal sensor scheduling for multiple linear dynamical systems," *Automatica*, vol. 75, pp. 260–270, 2017.

[15] S. Wu, X. Ren, S. Dey, and L. Shi, "Optimal scheduling of multiple sensors over shared channels with packet transmission constraint," *Automatica*, vol. 96, pp. 22–31, 2018.

[16] K. Gatsis, A. Ribeiro, and G. J. Pappas, "Random access design for wireless control systems," *Automatica*, vol. 91, pp. 1–9, 2018.

[17] J. Yun and C. Joo, "Information update for multi-source remote estimation in real-time monitoring networks," in *International Conference on Information and Communication Technology Convergence (ICTC)*, 2019, pp. 633–635.

[18] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine learning*, vol. 47, no. 2-3, pp. 235–256, 2002.

[19] R. G. Gallager, *Discrete stochastic processes*. Springer Science & Business Media, 2012, vol. 321.

[20] D. P. Bertsekas, D. P. Bertsekas, D. P. Bertsekas, and D. P. Bertsekas, *Dynamic programming and optimal control*. Athena scientific Belmont, MA, 1995, vol. 1, no. 2.