IEEE INFOCOM 2021 - IEEE Conference on Computer Communications | 978-1-6654-0325-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/INFOCOM42981.2021.9488702

Real-time Sampling and Estimation on Random Access Channels: Age of Information and Beyond

1st Xingran Chen
Electrical and System Engineering
University of Pennsylvania, USA
xingranc@seas.upenn.edu

2nd Xinyu Liao
Applied Mathematics and Computational Science
University of Pennsylvania, USA
xinyul@sas.upenn.edu

3rd Shirin Saeedi Bidokhti
Electrical and System Engineering
University of Pennsylvania, USA
saeedi@seas.upenn.edu

***Abstract*—Real-time sampling and estimation of autoregressive Markov processes is considered in random access channels. Two classes of policies are studied: (i) oblivious policies in which decision making is independent of the source realizations, and (ii) non-oblivious policies in which sources are observed causally for decision making. In the first class, minimizing the expected time-average estimation error is equivalent to minimizing the expected age of information (AoI). Lower and upper bounds are provided for the achievable estimation error in this class and age-based threshold policies are shown to provide a two-fold improvement compared to the state-of-the-art. In the second class, an error-based threshold policy is proposed: a transmitter becomes active when its error exceeds a threshold in which case it transmits probabilistically following slotted ALOHA. A closed-form expression is derived for the estimation error as a function of the peak age, the transmission delay, a term which we call the silence delay, as well as the source realization. It is analyzed approximately by considering the underlying source as a discretized Wiener process. The proposed threshold policy provides a three-fold improvement compared to oblivious policies and its performance is close to that of centralized greedy scheduling.**

***Index Terms*—Remote Estimation, Age of Information, Sampling, Decentralized Systems, Random Access, Collision Channel, Slotted ALOHA.**

I. Introduction

A. Motivation

The Internet of Things (IoT) paradigm is changing our conception of communications: It is no longer realistic to assume that information is known and stored at a source, waiting to be transmitted and replicated at the destination. Oftentimes, information is to be collected and communicated real-time within a decentralized network. For example, in applications of remote estimation and control, physical processes are observed at decentralized sensors that communicate wirelessly with a fusion center. In such applications, it is not realistic to assume a central scheduler that monitors all the sensors for decision making. In this paper, we study the problem of decentralized sampling and remote estimation of autoregressive Markov processes over a wireless collision channel.

B. Related Work

Below, we discuss three major facets of the problem.

Sampling: Remote estimation of physical processes requires efficient sampling and communication strategies that minimize not only the estimation error cost but also the sampling and transmission costs. With this viewpoint, prior works have studied optimal sampling strategies and their structural properties for various point-to-point scenarios. [1] designs optimal sampling strategies with limited measurements. [2] studies the problem for continuous sources. [3] proves the joint optimality of symmetric thresholding policies and Kalman-like estimators for autoregressive Markov processes. [4] formulates a two-player team problem and designs efficient iterative algorithms. Systems with energy harvesting sensors are considered in [5]. Noisy channels and packet drop channels are considered in [6], [7]. The above-mentioned works have all considered single-user channels and the developed methodologies do not generalize to random access networks with multiple sensors.

Reliable v.s. Timely Communication: In estimation and control applications, timeliness of communication is key and that is why traditional rate-distortion frameworks and channel coding paradigms that propose asymptotic block coding solutions are not applicable. More importantly, it is oftentimes observed that as the rate and/or reliability of compression/communication schemes improve, their timeliness decrease. This aspect of sampling and remote estimation is barely studied in the estimation literature. One of the few existing works in this direction is [8] which proposes and optimizes a hybrid automatic repeat request (HARQ)-based remote estimation protocol and improves the performance of the remote estimation systems compared to conventional non-HARQ policies. Recently, tradeoffs between reliability/rate timeliness of communication have been looked at in the context of age of information (AoI) – a metric of timeliness defined in [9]. In channels with queue constraints, [10] establishes a tradeoff between AoI and rate. [11] finds the optimal blocklength of channel coding for minimizing AoI. [12] provides a centralized scheduling framework to attain tradeoffs between rate and AoI in broadcast channels. [13] proposes decentralized transmission strategies for random access channels that benefit from the availability of fresh packets and improve both communication rate and AoI. It is known that AoI is closely related to the expected estimation error of schemes that are oblivious to the processes they monitor [14]. Non-oblivious sampling schemes are, however, signal-dependent and known to outperform oblivious schemes. In [14], threshold policies are shown to be optimal for point-to-

978-1-6654-0325-2/21/$31.00 ©2021 IEEE

point channels with a random delay and closed form solutions are found for the optimal threshold value. It is further shown that the oblivious policies can be far from optimal. We build on our prior work in [13] that concerned AoI minimization and propose decentralized threshold policies for minimizing estimation error in random access channels with many users.

Distributed decision making: In random access networks, a large number of sensors communicate with a single fusion center over a wireless channel. To avoid collision, most works in this direction have considered centralized oblivious policies that do not observe the process realizations for decision making (see, e.g., [15]–[21] and the references therein). In the IoT applications, however, it is not realistic to assume a central scheduler that monitors all the sensors for decision making. We seek decentralized solutions in which each sensor decides when to sample and transmit information based only on its local observations. In decentralized setups (and in the context of control, rather than estimation) [22], [23] consider wireless control architectures with multiple control loops over a random access channel and optimize the access rate of the sensors who randomly communicate. Policies that adapt to the state of the systems are proposed in [24]. The work [25] (which was carried out concurrently and independently) designs decentralized policies for the remote estimation of i.i.d processes over a collision channel. Decision making in both [24] and [25] is thresholding and based on the realization of the process (or a function of that). But since neither of the two works exploits channel collision feedback, adaptations of them (or other policies that impose a fixed rate on the channel) are far from optimal in our setup.

C. Contributions

In this paper, we study sampling and remote estimation of M independent random walk processes over a wireless collision channel. As opposed to all prior works, we seek decentralized solutions in which decision at each node is based solely on its local observations and channel collision feedback. Our goal is to minimize the estimation error, and specifically a normalized metric that we call the normalized expected weighted sum of estimation errors (NEWSEE). This metric looks at the expected time-average estimation error, normalized by the number of source nodes M. We are interested in the asymptotic regime where $M \to \infty$.

Two general classes of policies are considered, namely oblivious policies and non-oblivious policies. In the former class, decision making is independent of the processes that are monitored and we prove that minimizing the expected time-average estimation error, in the class of oblivious policies, is equivalent to minimizing the age of information. This leads to lower and upper bounds on the minimum achievable estimation error in this class along with efficient oblivious policies that are age-based. In particular, the NEWSEE under age-based policies is lower bounded by $.88\sigma^2$ and upper bounded by $\frac{e}{2}\sigma^2$.

We next ask if non-oblivious policies can provide a significant gain by observing the processes as they progress. Since all source nodes are provided with the channel collision feedback, they can compute their age-function and reproduce their respective estimated processes (at the destination) in each time slot. Furthermore, using the collision feedback, the nodes can implicitly coordinate for communication. We define the notion of *error process* at each node which is a function of the sample values and age. We then propose a threshold policy, called error-based thinning, in which source nodes become active only when their corresponding error process is beyond a given threshold. Once a node becomes active, it transmits stochastically following a slotted ALOHA policy.

To find an optimal threshold and find a closed-form solution for the resulting NEWSEE, we first provide a closed-form expression for the NEWSEE that is a function of the peak age, the transmission delay, a term which we call the silence delay, as well as the process realization. We approximately find the NEWSEE under an optimal threshold policy by considering the underlying autoregressive Markov process as a discretized Wiener process. An optimal threshold is then shown to be approximately $\sigma\sqrt{eM}$ and the resulting NWESEE to be $\frac{e}{6}\sigma^2$. The approximation error increases linearly as a function of the variance of the innovation process and decreases as M gets large.

Simulation results show that the proposed decentralied threshold policy outperforms oblivious policies. Moreover, oblivious policies are shown to outperform all state-of-the-art policies (both oblivious and non-oblivious) that impose a fixed rate (without using the collision feedback). Finally, it is numerically shown that the performance of the optimal threshold policy is very close to that of centralized greedy policies that schedule transmissions according to the instantaneous error reduction or age reduction.

The paper is organized as follows. In Section II, we introduce the system model. Oblivious policies are studied in Section III and non-oblivious policies are discussed in Section IV. Simulation results are presented for various policies in Section V and our assumptions and derivations are verified numerically. Finally, we conclude in Section VI.

D. Notation

We use the notations $\mathbb{E}[\cdot]$ and $\Pr(\cdot)$ for expectation and probability, respectively. Scalars are denoted by lower case letters, e.g. s, and random variables are denoted by capital letters, e.g. S. The notation $A \sim B$ implies that A has the same distribution as B and $\mathcal{N}(0, \sigma^2)$ stands for the Gaussian distribution with mean 0 and variance σ^2. The notations $O(\cdot)$ and $o(\cdot)$ represent the Big O and little o notations according to Bachmann-Landau notation, respectively.

II. System Model

Consider a system with M statistically identical sensors and a fusion center. We often refer to the sensor nodes as nodes or transmitters and the fusion center as the receiver/destination. Let time be slotted. Each node i, $i = 1, 2, \cdots, M$, observes a process $\{X_i(k)\}_{k\geq 0}$ which is a random walk process as follows

$$X_i(k+1) = X_i(k) + W_i(k) \tag{1}$$

where $W_i(k) \sim \mathcal{N}(0, \sigma^2)$. The processes $\{X_i(k)\}_{k=0}^{\infty}$ are assumed to be mutually independent across i and for simplicity we let $X_i(0) = 0$.

At the beginning of each time slot, the nodes have the capability to sample the underlying process and decide whether or not to communicate the sample with the receiver. The communication medium is modeled by a collision channel: If two or more nodes transmit in the same time slot, then the packets interfere with each other (collide) and do not get delivered at the receiver. We use the binary variable $d_i(k)$ to indicate whether a packet is transmitted from node i and delivered at the receiver in time slot k. Specifically, $d_i(k) = 0$ if node i does not transmit or if collision occurs; $d_i(k) = 1$ otherwise. We assume a delay of one time unit in delivery for packets. At the end of time slot k, all transmitters are informed (through a low-rate feedback link) whether or not collision occurred, which is indicated by an indicator $c(k)$. If collisions happen in time slot k, then $c(k) = 1$; if a packet is delivered successfully at the receiver or no packet is transmitted, then $c(k) = 0$.

We assume that the buffer size of every transmitter is one packet and that new packets replace older undelivered packets at the transmitter. This assumption relies on the fact that the underlying processes that are monitored are Markovian.

The receiver estimates the process in every time slot based on the collection of the received samples. Denote by $\hat{X}_i(k)$ the estimate of $X_i(k)$ in time slot k. We define the following normalized expected weighted sum of estimation errors (NEWSEE) as our performance metric:

$$\begin{aligned} L^{\pi}(M) &= \lim_{K\to\infty} \mathbb{E}[L_K^{\pi}] \\ L_K^{\pi}(M) &= \frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=1}^{K}\left(X_i(k) - \hat{X}_i(k)\right)^2 \end{aligned} \tag{2}$$

where M is the number of sources, $\pi \in \Pi$ refers to the sampling and transmission policy in place, and Π is the set of all decentralized sampling and transmission policies. Note that the metric (2) is normalized by M. This allows us to study the asymptotic performance in the regime of large M. The minimum attainable NEWSEE is then denoted by $L(M)$:

$$L(M) = \min_{\pi\in\Pi} L^{\pi}(M). \tag{3}$$

Our objective is to design *decentralized* sampling and transmission mechanisms to attain $L(M)$.

Consider the i^{th} node. Let $\{k_{\ell}^{(i)}\}_{\ell\geq 0}$ be the sequence of time slots at the end of which packets are received at the destination from node i. In any time slot k, $k_{\ell-1}^{(i)} < k \leq k_{\ell}^{(i)}$, the latest sample from node i is received at $k_{\ell-1}^{(i)}$ and since the delay is one time unit, it is time stamped at the beginning of $k_{\ell-1}^{(i)}$. So the age of information (AoI) [13] with respect to node i, denoted by $h_i(k)$, is

$$h_i(k) = k - k_{\ell-1}^{(i)}. \tag{4}$$

Without loss of generality, assume $k_0^{(i)} = 0$. At the beginning of time slot k, the receiver knows the information of all packets delivered before time k, i.e., $\{X_i(j)\}_{j=0}^{k-1}$ and reconstructs $\hat{X}_i(k)$ by the minimum mean square error (MMSE) estimator:

$$\hat{X}_i(k) = \mathbb{E}\left[X_i(k) | \{X_i(k_t^{(i)})\}_{t=0}^{\ell-1}\right].$$

For the class of policies that we consider in this paper (oblivious policies and symmetric thresholding policies), the MMSE estimator reduces to a Kalman-like estimator:

$$\hat{X}_i(k) = \mathbb{E}[X_i(k) | X_i(k_{\ell-1}^{(i)})] = X_i(k_{\ell-1}^{(i)}). \tag{5}$$

One of the major challenges in this problem arises from the decentralized nature of decision making. A decentralized policy is one in which the action of each node is only a function of its own local observations and actions. In this setup, the action of node i at time k depends on the history of feedback and actions as well as casual observations of the process $\{X_i(j)\}_{j\leq k}$.

We also consider a simpler class of policies Π', called *oblivious* policies, in which the action of each node depends only on the history of feedback and actions at that node. In particular, oblivious policies do not take into account the realization (value) of the samples, but only the time they were sampled, transmitted, and received (if successfully received). We denote the minimum attainable NEWSEE in the class of oblivious policies by

$$\bar{L}(M) = \min_{\pi\in\Pi'} L^{\pi}(M). \tag{6}$$

We argue in section III that this simplification equivalently transforms the estimation problem into the problem of timely communication of packets for age minimization. By additionally exploiting the value of the samples, in Section IV, we design and analyze decentralized mechanisms that outperform oblivious schemes in minimizing the expected average estimation error.

III. Oblivious Policies and Age of Information

Oblivious policies are independent of the processes they observe and they are therefore less costly to implement. Moreover, they can still benefit from the channel collision feedback to (i) quantify how stale the information at the receiver has become (in order to decide when to sample and communicate) and (ii) adapt to the channel state (for communication purposes). In this section, we show that minimizing NEWSEE in the class of oblivious policies is equivalent to minimizing the normalized expected weighted sum of AoI (NEWSAoI) as we have previously defined in [13].

First, we establish the following relationship between the expected estimation error and the expected age.

Lemma 1. *In oblivious policies, the expected estimation error associated with process i has the following relationship with the expected age function:*

$$\mathbb{E}[\big(X_i(k) - \hat{X}_i(k)\big)^2] = \mathbb{E}[h_i(k)]\sigma^2. \tag{7}$$

Remark 1. *Lemma 1 does not hold for non-oblivious policies. As a matter of fact, finding $\mathbb{E}[\big(X_i(k)-\hat{X}_i(k)\big)^2]$ in closed-form is non-trivial and its numerical computation can be intractable when M is large. The reason is that even though the estimation error is the sum of $h_i(k)$ Gaussian noise variables, once we condition on $h_i(k)$, their distributions change because $h_i(k)$ can be dependent on the process that is being monitored.*

Proof. The proof of Lemma 1 is given in [26]. □

Lemma 1 is reminiscent of [27, Lemma 4]. Using Lemma 1, the metric NEWSEE in (2) can be re-written as follows:

$$L^\pi(M) = \lim_{K\to\infty} \sigma^2 J^\pi(M) \tag{8}$$

where

$$J^\pi(M) = \frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=1}^{K}\mathbb{E}[h_i^\pi(k)]. \tag{9}$$

Note that $J^\pi(M)$ is only a function of the age function $h_i^\pi(k)$. The metric in (9) is the NEWSAoI defined in [13] and, threfore, the decentralized threshold policies of [13] apply directly. In particular, [13, Algorithm 2] outlines a stationary age-based thinning (SAT) policy in which a source transmits only when the corresponding AoI is larger than a pre-determined threshold. Using this algorithm, we can achieve the following age performance in the limit of large M:

$$\lim_{M\to\infty} J^{\text{SAT}}(M) = \frac{e}{2} \tag{10}$$

$$\lim_{M\to\infty} L^{\text{SAT}}(M) = \frac{e}{2}\sigma^2. \tag{11}$$

Results from [13, Proposition 1] also lead to the following lower bound on NEWSAoI $J^\pi(M)$ for any decentralized policy π:

$$\lim_{M\to\infty} J^\pi(M) \geq .88. \tag{12}$$

Using (11) and (12), we arrive at the following proposition.

Proposition 1. *The minimum attainable NEWSEE in the class of oblivious policies has the following bounds*

$$.88\sigma^2 \leq \lim_{M\to\infty} \bar{L}(M) \leq \frac{e}{2}\sigma^2. \tag{13}$$

A. Comparison with Oblivious Centralized Policies

In this section, we compare the SAT policy in [13, Algorithm 2] with an oblivious centralized policy – the Max-Weight (MW) policy [12], [13], [28]–[30]. Denote $\underline{T}^{(i)}(k) = \{k_j^{(i)}\}_{j=0}^{\ell}$ with $k_\ell^{(i)} \leq k$. We devise the MW policy using techniques from Lyapunov Optimization. Define the Lyapunov function

$$L(k) = \frac{1}{M}\sum_{i=1}^{M}\big(X_i(k) - \hat{X}_i(k)\big)^2 \tag{14}$$

and the one-slot Lyapunov Drift

$$LD(k) = \mathbb{E}[L(k+1) - L(k)|\underline{T}^{(i)}(k)]. \tag{15}$$

We devise the MW policy such that it minimizes the one-slot Lyapunov Drift.

Definition 1. *At the beginning of each slot k, the MW policy chooses the action i^* such that*

$$h_{i^*}(k) = \max_i h_i(k). \tag{16}$$

Note that this policy is exactly the MW policy derived in [30] for age minimization. From Lemma 2 in [28, Section III], the policy defined in Definition 1 is optimal.

Proposition 2. *The MW policy in Definition 1 minimizes the one-slot Lyapunov Drift in each slot, and*

$$\lim_{M\to\infty} L^{MW}(M) = \frac{\sigma^2}{2}. \tag{17}$$

Proof. The proof of Proposition 2 is given in [26, Appendix A]. □

Comparing (11) with (17), we have

$$\lim_{M\to\infty}\frac{L^{SAT}(M)}{L^{MW}(M)} = e.$$

The NEWSEE of the decentralized SAT policy is e times that of the optimal centralized policy in the limit of large M. The conclusion coincides with one's intuition: the throughput of the decentralized SAT policy in [13] is e^{-1}, while the throughput of the centralized MW policy is 1, which implies the amount of delivered fresh packets in the centralized MW policy is e times that of the decentralized SAT policy. We illustrate their performances through simulations in Section V.

IV. Non-oblivious Policies

We now consider a more general class of policies in which nodes can observe their corresponding Markov processes for decision making. In other words, we seek to benefit from not only the AoI, but also the process realization (in a casual manner). Clearly, if all nodes try to transmit their samples at every time slot, no packet will go through due to collisions. The nodes, therefore, need to transmit packets with a lower rate. This means that they have to decide, in a decentralized manner, when to transmit. Motivated by the optimality of threshold policies in various point-to-point setups [1], [2], [5], [14], as well as their applications in age minimization over many-to-one random access channels [13], we propose threshold policies for decision making.

A. Error-based Thinning

Define the *error process* $\psi_i(k)$ at node i as follows:

$$\psi_i(k) = |X_i(k) - \hat{X}_i(k)|. \tag{18}$$

Since the transmitters have access to collision feedback, they can calculate $\hat{X}_i(k)$, and hence $\psi_i(k)$, in each time slot and use this information for decision making. One way to understand $\psi_i(k)$ is as follows. At time k, if the sample of node i is successfully delivered, the estimation error will reduce by $\psi_i(k)$. So $\psi_i(k)$ quantifies the amount of instantaneous estimation error reduction upon successful delivery from transmitter i. With this viewpoint, we devise a threshold policy in which transmitters prioritize packets that have large $\psi_i(k)$. In particular, we design a fixed threshold β to distinguish and prioritize nodes that offer high instantaneous gain.

The action of each node is thus as follows: node i becomes "active" if the error process $\psi_i(k)$ has crossed a pre-determined threshold β. Once a transmitter is active, it remains active until a packet is successfully delivered from that node. Active nodes transmit stochastically following Rivest's stabilized slotted ALOHA protocol [31, Chapter 4.2.3]. Denote the number of active nodes and an estimate of the number of active nodes in time slot k as $N(k)$, $\hat{N}(k)$, respectively. In particular, each active node transmits its sample with probability $p_b(k)$ which is calculated adaptively as follows based on an estimate of the number of active nodes[1]:

$$\begin{aligned} p_b(k) &= \min(1, \frac{1}{\hat{N}(k)}) \\ \hat{N}(k) &= \begin{cases} \min\Big(\hat{N}(k-1) + \hat{\lambda}(k) + (e-2)^{-1}, M\Big) & \text{if } c(k-1) = 1 \\ \min\Big(\hat{\lambda}(k) + \big(\hat{N}(k-1) - 1\big)^{+}, M\Big) & \text{if } c(k-1) = 0. \end{cases} \end{aligned} \tag{19}$$

Here, $\hat{\lambda}(k)$ is an estimate of $\lambda(k)$, and $\lambda(k)$ is the sum arrival rate in time slot k. It is well-known that the maximum sum throughput of the slotted ALOHA is e^{-1} [31, Chapter 4.2.3] and the regime of interest is $\lambda(k) < e^{-1}$ when k is sufficiently large. In our setup, $\lambda(k)$ corresponds to the *expected* number of nodes that become *active* in time slot k (see Definition 2 ahead). We refer to $\lambda(k)$ as the activation rate or the effective arrival rate in time slot k.

So far, we have outlined a threshold policy in which a node decides to become active if its local error process is larger than a pre-determined threshold value β. We call this procedure *Error-based Thinning* (EbT). The main underlying challenge is, however, in the design of an *optimal* β. In the rest of this section, we will find an (approximately) optimal choice for β and analyze the corresponding NEWSEE approximately. We start by some preliminaries.

B. Preliminaries

Consider node i and an inter-delivery interval $(k_{\ell-1}^{(i)}, k_\ell^{(i)}]$ (see Figure 1). The inter-delivery time $I_\ell^{(i)}$ is given by $I_\ell^{(i)} = k_\ell^{(i)} - k_{\ell-1}^{(i)}$. For any time slot k, $k_{\ell-1}^{(i)} < k \le k_\ell^{(i)}$, we can write the error process $\psi(k)$ as follows:

$$\psi_i(k) = |X_i(k) - \hat{X}_i(k)| = \Big|\sum_{j=k_{\ell-1}^{(i)}}^{k-1} W_i(j)\Big|. \tag{20}$$

The term on the right hand side of (20) is the sum of $h_i(k)$ independent Gaussian noise variables $\big($see (4)$\big)$. Indeed, (20) demonstrates that $\psi_i(k)$ contains both the information of sample values as well as the age with respect to source i.

We next define "active" nodes as follows.

Definition 2 (Active Nodes)**.** *If there exists a time slot $k_0 \in (k_{\ell-1}^{(i)}, k_\ell^{(i)}]$ such that (i) $\psi_i(j) < \beta$ for all $k_{\ell-1}^{(i)} < j < k_0$ and (ii) $\psi_i(k_0) \ge \beta$, then we say that node i is active in the entire interval $[k_0, k_\ell^{(i)}]$.*

Definition 3 (Silence Delay and Transmission Delay)**.** *Let k_0 be as defined in Definition 2. We define $J_\ell^{(i)} = k_0 - k_{\ell-1}^{(i)}$ as the silence delay, and $U_\ell^{(i)} = k_\ell^{(i)} - k_0 + 1$ as the transmission delay (see Figure 1).*

An active source becomes inactive immediately after a successful delivery. By the above two definitions, the inter-delivery time $I_\ell^{(i)}$ consists of two components – the silence delay $J_\ell^{(i)}$ and the transmission delay $U_\ell^{(i)}$:

$$I_\ell^{(i)} = J_\ell^{(i)} - 1 + U_\ell^{(i)}. \tag{21}$$

In this equation, $J_\ell^{(i)}$ is the first time slot after $k_{\ell-1}^{(i)}$ at which $\psi_i(k) > \beta$ (as defined in Definition 3). So $J_\ell^{(i)} - 1$ represents the number of time slots in which node i is not active, and $U_\ell^{(i)}$ represents the number of time slots in which node i is in active state. Recall that active nodes transmit with probability $p_b(k)$. So $U_\ell^{(i)}$ may be larger than 1 either because the node is active and it does not transmit or because the node transmits and experiences collision. By the stationarity of the transmission scheme, the processes $\{I_\ell^{(i)}\}_{i,\ell}$, $\{J_\ell^{(i)}\}_{i,\ell}$, and $\{U_\ell^{(i)}\}_{i,\ell}$ are statistically identical across i and ℓ. We define I_β, J_β, and U_β to have the same distributions as $\{I_\ell^{(i)}\}_{i,\ell}$, $\{J_\ell^{(i)}\}_{i,\ell}$, and $\{U_\ell^{(i)}\}_{i,\ell}$, respectively.

Let $\{W_j\}_j$ be an i.i.d sequence with the same distribution as $\{W_j(k)\}_j$. Define

$$S_n = \sum_{j=1}^{n} W_j.$$

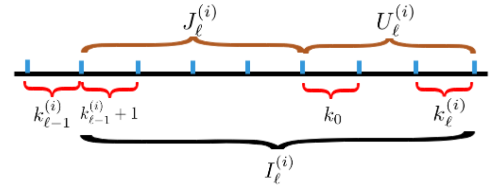

Fig. 1: an example of $J_\ell^{(i)}$, $U_\ell^{(i)}$, and $I_\ell^{(i)}$. Packets are generated at the beginning of every time slot, so $J_\ell^{(i)}$ arrivals/generations means $J_\ell^{(i)} - 1$ time slots.

[1]Since the sensors have unit buffer sizes, the number of "backlogged" nodes $N(k)$ in Rivest's algorithm is at most M. One notes that this has been incorporated in (19).

Using the definition of $h_i(k)$ in (4), and by the stationarity of $\{W_j\}_j$, we conclude that

$$\psi_i(k) \sim |S_{h_i(k)}|. \tag{22}$$

Recall that J_β has the same distribution as $J_\ell^{(i)}$. Then, J_β is the smallest time index at which $|S_n| \geq \beta$. J_β is a stopping time for S_n. From [32, Chapter 7.5.1, Lemma 7.5.1], it follows that J_β has finite moments of all orders. Moreover, using [32, Chapter 7.5.2], we have

$$\mathbb{E}[S_{J_\beta}^2] = \sigma^2 \mathbb{E}[J_\beta]. \tag{23}$$

Finding an optimal β is non-trivial because β impacts both J_β and U_β. In the remainder of this subsection, we establish some useful expressions for the expectations of I_β and U_β in an optimal design.

Let $a(k)$ denote the number of newly active nodes at time k. We have $\mathbb{E}[a(k)] = \lambda(k)$, where $\lambda(k)$ is the expected sum rate/throughput in time slot k (imposed by our sampling and transmission policy). Now recall that in a traditional slotted Aloha-based random access channel, the maximum sum throughput is asymptotically e^{-1}. This is true also for the case with buffer size 1 where only the latest packets are stored, as discussed in [13, Appendix E]) and which applies to our setting here. Define $c(M)$ as the sum throughput when the system contains M sources.

Definition 4. *The random access system is stabilized*[2] *if* $\lambda_m = \limsup_{k\to\infty} \lambda(k) < e^{-1}$.

We provide our analysis under the following two assumptions:

Assumption 1. *Under an optimal β, when M is sufficiently large, $\{a(k)\}_{k=1}^{\infty}$ are approximately independent.*

Assumption 2. *Under an optimal β, when M is sufficiently large, the random access system is stabilized, and $\lambda_m \approx e^{-1}$, $c(M) \approx e^{-1}$.*

Assumptions 1, 2 are given for analysis tractablility, and we will verify them for our proposed β later. In the rest of the paper, let M be sufficiently large. We analyze the optimal β under assumptions 1, 2. To transmit as many fresh samples as possible, β is designed such that $\lambda(k)$ is as large as possible. Recall that we focus on the regime where $\lambda(k)$ is close to e^{-1} when k is large. From Assumption 2, $\lambda_m \approx e^{-1}$. For tractability in analysis, we let the estimate $\hat{\lambda}(k) = e^{-1}$ for all k.

Note that the system is stationary, so $U_\ell^{(i)}$ (or U_β) is a random variable, hence $U_\ell^{(i)}$ (or U_β) is measurable. Recall that J_β has finite moments of all orders. Then, I_β is measurable. We remark that $\{I_\ell^{(i)}\}_\ell$ is not independent but rather weakly correlated across ℓ as we prove in [26, Appendix B]. We can thus conclude that the strong law of large numbers holds for $\{I_\ell^{(i)}\}_\ell$, see also [33].

[2] Here, contrary to traditional slotted ALOHA schemes, the term “stabilized” does not refer to ”stability of queues” in our problem setup. However, similar to traditional slotted ALOHA schemes, the term “stabilized” implies that the system is stationary, which has sum arrival rate less than e^{-1}

Recall that $N(k)$ is the number of active nodes at the *end* of time slot k. The fraction of active nodes at the *end* of time slot k is hence $N(k)/M$.

Definition 5. *Define $\alpha_\beta(k)$ as the expected fraction of active nodes:*

$$\alpha_\beta(k) = \frac{\mathbb{E}[N(k)]}{M}. \tag{24}$$

If $\beta = 0$, then all nodes are active and $\alpha_0(k) = 1$; if $\beta = +\infty$, then all nodes are inactive and $\alpha_{+\infty}(k) = 0$. In the limit of $k \to \infty$, we denote the expected fraction of active nodes by α_β:

$$\alpha_\beta = \lim_{k\to\infty} \frac{\mathbb{E}[N(k)]}{M}. \tag{25}$$

Lemma 2. *When the system is stabilized, α_β exists, and $\alpha_\beta = \frac{\mathbb{E}[U_\beta]}{\mathbb{E}[I_\beta]}$.*

Proof. The proof of Lemma 2 is given in [26, Appendix C]. □

Note that $\alpha_\beta = \frac{\mathbb{E}[U_\beta]}{\mathbb{E}[I_\beta]}$ represents the probability of each node being active when the system is steady. Since α_β exists, then, when $k \to \infty$, the expected number of nodes that become active in every time slot is $(1-\alpha_\beta)M\alpha_\beta$, and

$$(1-\alpha_\beta)M\alpha_\beta = \lim_{k\to\infty} \lambda(k) = \limsup_{k\to\infty} \lambda(k) = \lambda_m. \tag{26}$$

Lemma 3. *When the system is stabilized,*

$$\mathbb{E}[I_\beta] = \frac{M}{c(M)} \tag{27}$$

$$\mathbb{E}[U_\beta] = \frac{M}{c(M)}\alpha_\beta = o(M) \tag{28}$$

where α_β is the expected fraction of active nodes in the steady state as defined in (25)*.*

Remark 2. *Lemma 3 coincides with one's intuition. Recall that the throughput of the channel is $c(M)$, so the throughput for each node is $\frac{c(M)}{M}$ (due to the symmetry). From the perspective of expectation, every successful delivery takes $\frac{M}{c(M)}$ time slots, i.e., $\mathbb{E}[I_\beta] = \frac{M}{c(M)}$. In addition, note that the expected number of active node is $M\alpha_\beta$, so the throughput of every active node is $\frac{c(M)}{M\alpha_\beta}$. Again, from the perspective of expectation, every successful delivery from active nodes takes $\frac{M}{c(M)}\alpha_\beta$ time slots, i.e., $\mathbb{E}[U_\beta] = \frac{M}{c(M)}\alpha_\beta$.*

Proof. The proof of Lemma 3 is given in [26, Appendix D]. □

C. The closed form of NEWSEE

We next derive a closed form expression for the attained NEWSEE, $L^{EbT}(M)$. Using (22), we re-write (2) as follows.

$$L^{EbT}(M) = \lim_{K\to\infty} \mathbb{E}[\frac{1}{M^2K}\sum_{i=1}^{M}\sum_{k=1}^{K} S_{h_i(k)}^2]. \tag{29}$$

Define $\Delta_\ell^{(i)}$ as the sum of $S_{h_i(k)}^2$ in the interval $k \in (k_{\ell-1}^{(i)}, k_\ell^{(i)}]$:

$$\Delta_\ell^{(i)} = \sum_{k=k_{\ell-1}^{(i)}+1}^{k_\ell^{(i)}} S_{h_i(k)}^2. \tag{30}$$

The next lemma shows that the expected time average in (29) takes a closed form expression in terms of $\mathbb{E}[\Delta_\beta]$ and $\mathbb{E}[I_\beta]$.

Lemma 4. *The proposed EbT policy attains the following NEWSEE:*

$$L^{EbT}(M) = \frac{1}{M}\frac{\mathbb{E}[\Delta_\beta]}{\mathbb{E}[I_\beta]}. \tag{31}$$

Proof. The proof of Lemma 4 is given in [26, Appendix E]. □

The NEWSEE in (31) can now be re-written as follows

$$\begin{aligned} L^{EbT}(M) =& \frac{1}{M}\frac{\mathbb{E}\left[\sum_{j=1}^{I_\beta} S_j^2\right]}{\mathbb{E}[I_\beta]} \\ =& \frac{1}{M}\frac{\mathbb{E}\left[\sum_{j=1}^{J_\beta+U_\beta-1} S_j^2\right]}{\mathbb{E}[I_\beta]} \\ \triangleq& L_1^{EbT}(M) + L_2^{EbT}(M) \end{aligned} \tag{32}$$

where

$$L_1^{EbT}(M) = \frac{1}{M}\frac{\mathbb{E}\left[\sum_{j=1}^{J_\beta} S_j^2\right]}{\mathbb{E}[I_\beta]} \tag{33}$$

and (34)

$$L_2^{EbT}(M) \tag{35}$$

$$= \frac{1}{M}\frac{\mathbb{E}\left[\sum_{j=J_\beta+1}^{J_\beta+U_\beta-1} S_j^2\right]}{\mathbb{E}[I_\beta]} \tag{36}$$

$$= \frac{1}{M}\cdot\frac{2\mathbb{E}[J_\beta](\mathbb{E}[U_\beta]-1)+\mathbb{E}[U_\beta^2]-\mathbb{E}[U_\beta]}{2\mathbb{E}[I_\beta]}\sigma^2. \tag{37}$$

The equality in (37) is proved in [26, Appendix F]. Note that L^{EbT} is a function of the peak age I_β, the silience delay J_β, the transmission delay U_β, and the process realization through W_j.

D. Optimizing β Approximately

Finally, we find approximate closed form expressions for $L_1^{EbT}(M)$ and $L_2^{EbT}(M)$. Let M be sufficient large. From (27) and (28) in Lemma 3, $\mathbb{E}[I_\beta] = \frac{M}{c(M)}$ and $\mathbb{E}[U_\beta] = o(M)$, then (37) can be re-written as

$$L_2^{EbT}(M) = \frac{1}{M}\cdot\frac{\mathbb{E}[U_\beta^2]}{2\mathbb{E}[I_\beta]}\sigma^2. \tag{38}$$

The following lemma comes in handy in our approximations.

Lemma 5. *Consider a Brown motion B_t. Define $J = \inf\{t \geq 0, |B_t| \geq a\}$. The following holds:*

(1) [34, Chapter 7, Theorem 7.5.5, Theorem 7.5.9] $\mathbb{E}[J] = a^2$ and $\mathbb{E}[J^2] = \frac{5a^4}{3}$;
(2) $\mathbb{E}[\int_0^J B_t^2 dt] = \frac{1}{10}\mathbb{E}[J^2] = \frac{1}{6}a^4$.

Proof. The proof of Lemma 5 is given in [26, Appendix G]. □

For any j, $\frac{S_j}{\sigma}$ is Gaussian with mean zero and variance j. We propose to use B_j as an approximation of $\frac{S_j}{\sigma}$. Letting $a = \beta/\sigma$ in Lemma 5, we obtain

$$\mathbb{E}\big[J_\beta\big] \approx \frac{\beta^2}{\sigma^2},\ \mathbb{E}[J_\beta^2] \approx \frac{5\beta^4}{3\sigma^4} \tag{39}$$

$$\mathbb{E}\big[\sum_{j=1}^{J_\beta} S_j^2\big] \approx \frac{\beta^4}{6\sigma^2} \approx \frac{1}{10}\mathbb{E}[J_\beta^2]. \tag{40}$$

The approximation error analysis is provided in Section IV-E.

Substituting (40) into (32), we find the following approximation for L^{EbT}:

$$\hat{L}^{EbT}(M) = \frac{\frac{1}{5}\mathbb{E}[J_\beta^2] + \mathbb{E}[U_\beta^2]}{2M\mathbb{E}[I_\beta]}\sigma^2. \tag{41}$$

Theorem 1. *Let M be sufficient large. The optimal β^* is approximately given by*

$$\beta^* = \sigma\sqrt{eM},$$

and

$$\hat{L}^{EbT} = \frac{e}{6}\sigma^2. \tag{42}$$

Proof. The proof of Theorem 1 is given in [26, Appendix H]. □

Finally, assumptions 1, 2 are verified (approximately) for β^* when M is sufficiently large in [26, Appendix I].

It is interesting to compare the performance of the proposed EbT policy with the oblivious decentralized and centralized policies of Section III. From (8), (9), and (10),

$$\lim_{M\to\infty} L^{SAT}(M) = \frac{e}{2}\sigma^2.$$

Using (11) and (42), we obtain

$$\lim_{M\to\infty} \frac{L^{SAT}(M)}{\hat{L}^{EbT}(M)} \approx 3.$$

The NEWSEE of oblivious SAT policy is around three times that of the EbT policy. From (13), the NEWSEE of the oblivious MW policy of Section III is asymptotically $\frac{\sigma^2}{2}$ and comparing with $\frac{e}{6}\sigma^2 = 0.455\sigma^2$ one concludes that the NEWSEE of the EbT policy is close to that of the oblivious MW policy. We remark that since $\hat{L}^{EbT}(M)$ is an estimate of $L^{EbT}(M)$, these comparisons are not exact. We will also compare the numerical performance of Algorithm 1 with oblivious policies as well as other state-of-the-art algorithms in Section V.

Algorithm 1 below summarizes the proposed decentralized error-based transmission policy.

Algorithm 1 Error-based Thinning (EbT)

Set the time horizon K.
Set initial points: $k = 1$; $h_i(0) = 1$, $X_i(0) = \hat{X}_i(0) = 0$ for $i = 1, 2, \cdots, M$; $c(0) = 0$; $d_i(0) = 0$; $p_b(0) = 1$; $\hat{N}(0) = 0$.
Set $\beta^* = \sigma\sqrt{eM}$.
repeat
Step 1: For each node i, observe the collision feedback $c(k-1)$ and $d_i(k-1)$ at the end of time slot $k-1$, and update $k_\ell^{(i)}$'s and $\hat{X}_i(k)$, respectively.
Step 2: For each node i, observe $X_i(k)$ $\big($which evolves according to (1)$\big)$ and compute $\psi_i(k)$ by (18).
Step 3: If $\psi_i(k) < \beta^*$, then node i does not transmit packets; otherwise it transmits a packet with probability $p_b(k)$.
Step 4: Calculate $p_b(k)$ by (19) in which $\lambda(k) = e^{-1}$.
until $k = K$
Calculate

$$L_K^{EbT} = \frac{1}{M^2}\sum_{i=1}^{M}\frac{1}{K}\sum_{k=0}^{K}\psi_i^2(k).$$

E. Approximation Error Analysis

Now we analyze the approximation error. In particular, we discuss how the approximation error changes with σ. The approximation error of L^{EbT} consists of (i) the approximation error in (39) and (ii) the approximation error in (40), both of which incurred when approximating an autoregressive Markov process with a Wiener process. In other words, the approximation error is due to the discretization of the Wiener process. This discretization is analyzed by the Langevin dynamics in [35]. In particular, $\frac{S_n}{\sigma} = \sum_{i=1}^{n} W_i \approx B_n$ can be regarded as an overdamped Langevin dynamics with step size 1 to approximate the Brownian motion. The approximation error in each step remains constant due to the unit step size.

We first consider $\mathbb{E}[J_\beta]$. Substituting $\beta = \sigma\sqrt{eM}$ into $a = \beta/\sigma$ in Lemma 5, $a = \sqrt{eM}$ is constant. So the distribution of J in Lemma 5 does not change when σ changes. Thus, the approximation error in (39) keeps invariant when σ changes.

Then, we consider (40). J_β is an approximation of J, and

$$\sum_{j=1}^{J_\beta} S_j^2 = \sigma^2 \sum_{j=1}^{J_\beta} S_j^2/\sigma^2. \tag{43}$$

The distribution of J does not change with σ, nor does the distribution of J_β. The terms $\frac{S_j}{\sigma} \sim \mathcal{N}(0, j)$ inside the sum in (43) are independent of σ. The distribution of $\sum_{j=1}^{J_\beta} S_j^2/\sigma^2$ does not change with σ. Thus, the approximation error in (40) increases linearly with σ^2.

By Lemma 5 (2), we have $\mathbb{E}[J^2] = 10\mathbb{E}[\int_0^T B_t^2 dt]$. Recall that the approximation error in (40) increases linearly with σ^2, thus the approximation error in $\mathbb{E}[J^2]$ also increases linearly with σ^2. Using (41), we conclude that the approximation error in $L^{EbT}(M)$ increases linearly with σ^2.

V. Numerical Results

In this section, we verify our findings through simulations. Figure 2 compares the NEWSEE of our proposed policy with the state of the art for $M = 500$ under different σ^2. In this plot, the green (plus) curve corresponds to an optimal stationary randomized policy in which each node transmits with an optimal pre-determined probability. The performance of threshold policies like [24], [25] that impose the optimal (fixed) transmission rate for each sensor also coincides with this curve, i.e, the green (plus) one. These policies do not exploit the available feedback for decision making. The purple (diamond) curve shows the performance of a standard pseudo-Bayesian slotted ALOHA. Slotted ALOHA does use feedback, but treats all packets similarly, independent of their corresponding sample values. The red (circle) and blue (squared) curves correspond to oblivious (age-based) policies [13, Algorithm 1] and [13, Algorithm 2], respectively. The black (star) curve shows the performance of our proposed decentralized policy in Algorithm 1 and the red (x) curve shows the approximation we find in (42). The gap between the two is small but increases linearly in σ^2 as discussed in Section IV-E. On this plot, we have also included an oblivious and a non-oblivious centralized policy. The former (green dashed curve) schedules the transmitter with the largest age and the latter (yellow smooth curve) schedules the transmitter with the largest estimation error. Oblivious centralized policies are optimal (from [28, Section III]) while non-oblivious centralized policies are not necessarily optimal (as they only optimize one time step ahead), they are often observed to be numerically very close to the optimal.

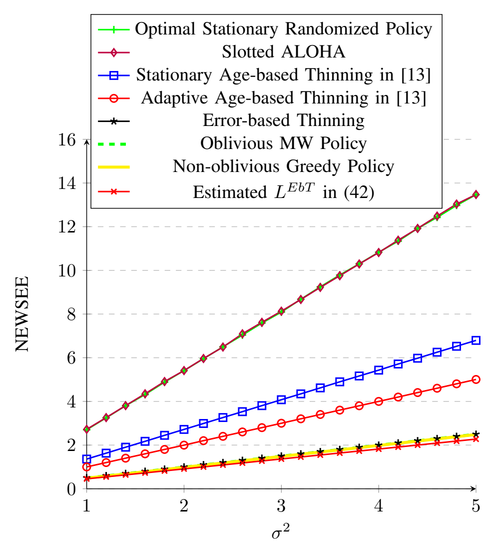

Fig. 2: NEWSEE as a function of σ^2 for various state-of-the-art scheme with $M = 500$.

The numerical calculation and analytical approximation of $\mathbb{E}[J_\beta]$, $\mathbb{E}[\sum_{j=1}^{J_\beta} S_j^2]$ and $\mathbb{E}[U_\beta]$ are given in Figure 3,

Figure 4 and Figure 5, respectively. Recall that $\mathbb{E}[J_\beta^2]$ is 10 times $\mathbb{E}\big[\sum_{j=1}^{J_\beta} S_j^2\big]$, so we only consider one of them. In order to offset the effect introduced by the number of nodes, we consider the normalized silence delay $\mathbb{E}[J_\beta]/M$, the normalized transmission delay $\mathbb{E}[U_\beta]/M$, and $\mathbb{E}[\sum_{j=1}^{J_\beta} S_j^2]/M$. The estimation error of the normalized silence delay $\mathbb{E}[J_\beta]/M$ is invariant of σ^2 (Figure 3), while the estimation error of $\mathbb{E}[\sum_{j=1}^{J_\beta} S_j^2]/M$ increases linearly with σ^2 (Figure 4). This coincides with the analysis in Section IV-E. In the simulation, we numerically find the expected fraction of active nodes to be $\alpha_\beta = 0.0173$. Substituting $\alpha_\beta = 0.0173$ into (28), we get $\mathbb{E}[U_\beta]$. From Figure 5, we can see that normalized transmission delay $\mathbb{E}[U_\beta]$ coincides with analytical results in (28).

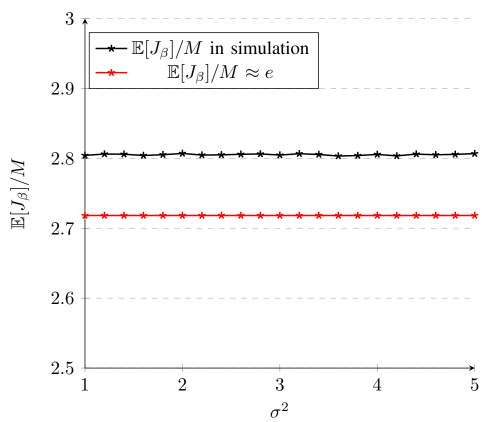

Fig. 3: $\mathbb{E}[J_\beta]/M$ as a function of σ^2 for $M = 500$.

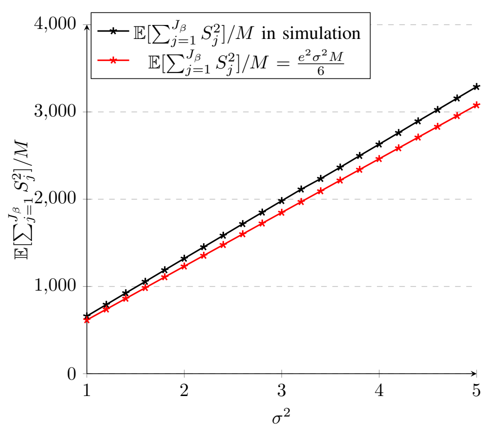

Fig. 4: $\mathbb{E}[\sum_{j=1}^{J_\beta} S_j^2]/M$ as a function of σ^2 for $M = 500$.

Finally, we show in Figure 6 that the gap between $L^{EbT}(M)$ and $\hat{L}^{EbT}(M)$ decreases as M gets large. In other words, the influence of approximation error caused by Langevin dynamics in Algorithm 1 weakens (but does not vanish) as M increases.

VI. Conclusion and Future Work

We considered the problem of decentralized sampling and remote estimation over wireless collision channels with

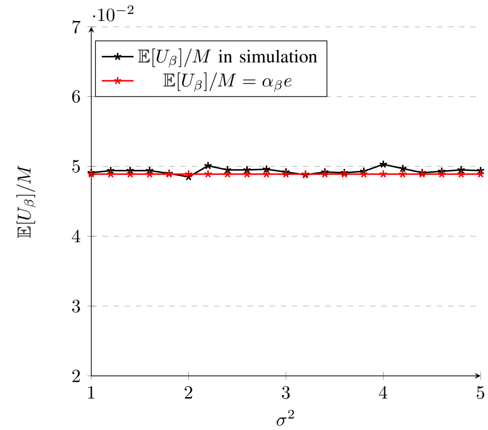

Fig. 5: $\mathbb{E}[U_\beta]/M$ as a function of σ^2 for $M = 500$.

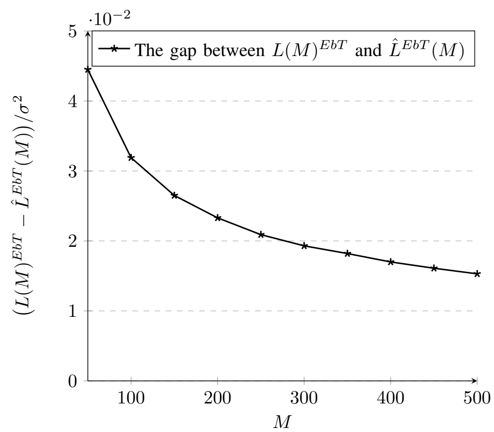

Fig. 6: The gap (normalized by σ^2) between $L^{EbT}(M)$ and $\hat{L}^{EbT}(M)$ as a function of M for $\sigma^2 = 3$.

M statistically identical source nodes, observing independent random walk processes. The goal is to minimize a normalized metric of estimation error, which we call the normalized expected weighted sum of estimation error (NEWSEE) in the regime of large M. We defined two general classes of policies: oblivious policies and non-oblivious policies. We showed in the former class that minimizing the expected estimation error is equivalent to minimizing the expected age and consequently proved lower and upper bounds on the optimal estimation error. We then proposed and analyzed a (non-oblivious) threshold policy in which (1) nodes become active if their estimation error has crossed a threshold and (2) active nodes transmit stochastically with probabilities that adapt to the state of the channel (exploiting the collision feedback). We showed that the NEWSEE performance of oblivious (age-based) policies is at least twice better than the state-of-the-art schemes (which impose a fixed rate of transmission at the nodes) such as standard slotted ALOHA and optimal stationary randomized policy. Moreover, our proposed threshold policy offers a multiplicative gain close to 3 compared to oblivious policies.

Future research includes generalizations to accommodate the following scenarios: 1) dynamic networks, i.e., the number of sensors changes with time; 2) asymmetric networks, i.e., the sensors are no longer statistically identical; 3) adaptive error-based thinning policies, i.e., the threshold $\beta(k)$ changes with time k; 4) correlated sources, i.e., sensors are no longer mutually independent. For the first scenario, we can simply replace M with $M(k)$ in every time slot. Subsequently, the error-based threshold is also a time-variant variable, $\beta(k)$. For the remaining three scenarios, the method we have proposed can not be applied directly. In particular, in the second scenario, we use the profile of all the sources to find an estimate on any individual source. In the third scenario, the nodes need statistical information about the sensors (and their underlying processes) to decide which ones are of priority. In the fourth scenario, the policies should change to account for the correlation between the observations.

Acknowledgment

This work was supported by Grants NSF 1910594 and NSF 1850356. The authors would like to thank Dr. Hamed Hassani and Dr. Konstantinos Gatsis for helpful discussions.

References

[1] O. Imer and T. Basar, "Optimal estimation with limited measurements," *Int. J. Sysmtes, Control and Communications*, vol. 2, pp. 5 – 29, 2010.
[2] G. V. M. M. Rabi and J. S. Baras, "Adaptive Sampling for Linear State Estimation," *SIAM Journal on Control and Optimization*, vol. 50, no. 2, pp. 672–702, 2012.
[3] G. Lipsa and N. 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.
[4] A. Molin and S. Hirche, "Event-Triggered State Estimation: An Iterative Algorithm and Optimality Properties," *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 5939 – 5946, 2017.
[5] A. Nayyar, T. Basar, D. Teneketzis, V. V. Veeravalli, "Communication scheduling and remote estimation with energy harvesting sensor," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, 2012.
[6] J. Chakravorty and A. Mahajan, "Remote Estimation Over a Packet-Drop Channel With Markovian State," *IEEE Transactions on Automatic Control*, vol. 65, no. 5, pp. 2016 – 2031, 2020.
[7] X. Gao, E. Akyol, and T. Basar, "Optimal communication scheduling and remote estimation over an additive noise channel," *Automatica*, vol. 88, pp. 57 – 69, 2018.
[8] K. Huang, W. Liu, Y. Li, and B. Vucetic, "To Retransmit or Not: Real-Time Remote Esimation in Wireless Networked Control," arXiv: 1902.07820, Feb 2019.
[9] S.Kaul, M.Gruteser, V. Rai, and J. Kenny, "Minimizing age of information in vehicular networks," in *2011 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, 2011, pp. 350–358.
[10] R. Talak and E. Modiano, "Age-Delay Tradeoffs in Queueing Systems," arXiv: 1911.05601, Nov 2019.
[11] H. Sac, T. Bacinoglu, E. Uysal-Biyikoglu, and G. Durisi, "Age-Optimal Channel Coding Blocklength for an M/G/1 Queue with HARQ," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2018.
[12] X. Chen and S. Saeedi-Bidokhti, "Benefits of Coding on Age of Information in Broadcast Networks," in *2019 IEEE Information Theory Workshop (ITW)*, 2019.
[13] X. Chen, K. Gatsis, H. Hassani and S. Saeedi-Bidokhti, "Age of information in random access channels," arXiv: 1912.01473, 2019.
[14] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote Estimation of the Wiener Process over a Channel with Random Delay," *IEEE Transactions on Information Theory*, vol. 66, no. 2, pp. 1118 – 1135, 2020.
[15] L. Meier, J. Peschon, and R.M. Dressler, "Optimal control of measurement subsystems," *IEEE Transactions on Automatic Control*, vol. 12, no. 5, pp. 528 – 536, 1967.
[16] Y. Oshman, "Optimal sensor selection strategy for discrete-time state estimators," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 2, pp. 307 – 314, 1994.
[17] A. Logothetis and A. Isaksson, "On sensor scheduling via information theoretic criteria," in *Proceedings of the 1999 American Control Conference*, 1999.
[18] V. Gupta, T. H. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage," *Automatica*, vol. 42, no. 2, pp. 251 – 260, 2006.
[19] M. P. Vitus, W. Zhang, A. Abate, J. Hu, and C. J. Tomlin, "On efficient sensor scheduling for linear dynamical systems," in *Proceedings of the 2010 American Control Conference*, 2010.
[20] L. Zhao, W. Zhang, J. Hu, A. Abate, and C. J. Tomlin, "On the Optimal Solutions of the Infinite-Horizon Linear Sensor Scheduling Problem," *IEEE Transactions on Automatic Control*, vol. 59, no. 10, pp. 2825 – 2830, 2014.
[21] S. T. Jawaid and S. L. Smith, "Submodularity and greedy algorithms in sensor scheduling for linear dynamical systems," *Automatica*, vol. 61, pp. 282 – 288, 2015.
[22] K. Gatsis, M. Pajic, A. Ribeiro and G. Pappas, "Opportunistic Control Over Shared Wireless Channels," *IEEE Transactions on Automatic Control*, vol. 60, no. 12, 2015.
[23] K. Gatsis, A. Ribeiro, and G. Pappas, "Control with Random Access Wireless Sensors," in *2015 IEEE 54th Annual Conference on Decision and Control (CDC)*, 2015.
[24] G. Taricco, "Joint Channel and Data Estimation for Wireless Sensor Networks," *IEEE Wireless Communications Letters*, vol. 1, no. 5, pp. 532 – 535, 2012.
[25] X. Zhang, M. M. Vasconcelos, W. Cui, and U. Mitra, "Distributed remote estimation over the collision channel with and without local communication," arXiv: 2005.11438, May 2020.
[26] X. Chen, X. Liao and S. Saeedi-Bidokhti, "Real-time sampling and estimation on random access channels: Age of information and beyond," arXiv: 2007.03652, Jul 2020.
[27] T. Z. Ornee and Y. Sun, "Sampling for Remote Estimation through Queues: Age of Information and Beyond," in *International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT)*, 2019.
[28] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling Policies for Minimizing Age of Information in Broadcast Wireless Networks," *IEEE/ACM Transactions on Networking*, vol. 26, no. 6, pp. 2637 – 2650, 2018.
[29] I. Kadota, A. Sinha and E. Modiano, "Scheduling Algorithms for Optimizing Age of Information in Wireless Networks With Throughput Constraints," *IEEE/ACM Transactions on Networking*, vol. 27, no. 4, pp. 1359 – 1372, 2019.
[30] I. Kadota and E. Modiano, "Minimizing the Age of Information in Wireless Networks with Stochastic Arrivals," *IEEE Transactions on Automatic Control*, Early Access, 2019.
[31] D. Bertsekas and R. Gallager, *Data Networks (2Nd Ed)*. Prentice-Hall, Inc., 1992.
[32] R. Gallager, *Discrete Stochastic Process*. Springer, Boston, MA, 1996.
[33] R. Lyons, "Strong laws of large numbers for weakly correslated random variables," *Michigan Math. J*, vol. 35, pp. 353 – 359, 1988.
[34] R. Durrett, *Probability: Theorey and Examples*. New York, NY: Cambridge University Press, 2019., 2019.
[35] A. M. Horowitz, "The second order Langevin equation and numerical simulations," *Nuclear Physics B*, vol. 280, no. 1, pp. 510 – 522, 1987.