Timely Multi-Process Estimation with Erasures

Karim Banawan[1], Ahmed Arafa[2], and Karim G. Seddik[3]

[1]Department of Electrical Engineering, Alexandria University, Egypt.
[2]Department of Electrical and Computer Engineering, University of North Carolina at Charlotte, USA
[3]Electronics and Communications Engineering Department, American University in Cairo, Egypt.

Abstract—We consider a multi-process remote estimation system observing K independent Ornstein-Uhlenbeck processes. In this system, a shared sensor samples the K processes in such a way that the long-term average sum mean square error (MSE) is minimized. The sensor operates under a total sampling frequency constraint $f_{\max}$ and samples the processes according to a Maximum-Age-First (MAF) schedule. The samples from all processes consume random processing delays, and then are transmitted over an erasure channel with probability ϵ. Aided by optimal structural results, we show that the optimal sampling policy, under some conditions, is a *threshold policy*. We characterize the optimal threshold and the corresponding optimal long-term average sum MSE as a function of K, $f_{\max}$, ϵ, and the statistical properties of the observed processes.

I. INTRODUCTION

We study the problem of timely tracking of multiple random processes using shared resources. This setting arises in many practical situations of remote estimation applications. Recent works have drawn connections between the quality of the estimates at the destination, measured through mean square error (MSE), and the *age of information* (AoI) metric that assesses timeliness and freshness of the received data, see, e.g., the survey in [1, Section VI]. We extend these results to *multi-process* estimation settings in this work.

AoI is defined as the time elapsed since the latest received message has been generated at its source. It has been studied extensively in the past few years in various contexts, see, e.g., [2]–[8]. Relevant to this work is the fact that AoI can be closely tied to MSE in random processes tracking applications. The works in [9]–[11] characterize implicit and explicit relationships between MSE and AoI under different estimation contexts. References [12], [13], however, consider the notion of the value of information (mainly through MSE) and show that optimizing it can be different from optimizing AoI. Lossy source coding and distorted updates for AoI minimization is considered in [14]–[16]. The notion of age of incorrect information (AoII) is introduced in in [17], adding more context to AoI by capturing erroneous updates. The works in [18], [19] consider sampling of Wiener and Ornstein-Uhlenbeck (OU) processes for the purpose of remote estimation, and draw connections between MSE and AoI. Our recent work in [20] also focuses on characterizing the relationship of MSE and AoI, yet with the additional presence of coding and
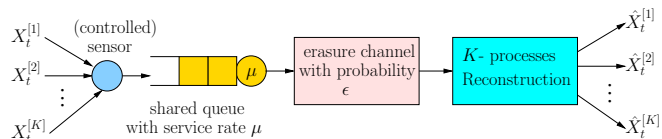

Fig. 1. System model.

quantization. Reference [21] shows the optimality of threshold policies for tracking OU processes under rate constraints.

Reference [19] is closely related to our setting, in which optimal sampling methods to minimize the long-term average MSE for an OU process is derived. It is shown that if sampling times are independent of the instantaneous values of the process (signal-independent sampling) the minimum MSE (MMSE) reduces to an increasing function of AoI (age penalty). Then, threshold policies are shown optimal in this case, in which a new sample is acquired only if the expected age-penalty surpasses a certain value. This paper extends [19] (and the related studies in [20], [21]) to multiple OU processes.

In this paper, we study a remote sensing problem consisting of a shared controlled sensor, a shared queue, and a receiver (see Fig. 1) to track K independent, but not necessarily identical, OU processes.[1] The sensor transmits the collected samples over an erasure channel with probability ϵ after being processed for a random delay with service rate μ. The sensor generates the samples *at will*, subject to a total sampling frequency constraint $f_{\max}$. The goal is to minimize the long-term average sum MSE of the K processes. We focus on *maximum-age-first* (MAF) scheduling, where the scheduler chooses the process with the largest AoI to be sampled. MAF scheduling results in obtaining a *fresh* sample from the *same* process until an unerased sample from that process is conveyed to the receiver. We show that the optimal stationary deterministic policy is a *threshold policy*. We characterize the optimal threshold $\tau^*(K, f_{\max}, \epsilon, \boldsymbol{\theta}, \boldsymbol{\sigma})$ and the corresponding long-term average sum MSE in terms of the processes statistical properties $(\boldsymbol{\theta}, \boldsymbol{\sigma})$, ϵ, and $f_{\max}$. The threshold is a maximum of two threshold values: one due to a nonbinding sampling frequency constraint scenario, and another due to a binding scenario. Our numerical results show that 1) the optimal threshold τ^* is an increasing function in the erasure probability ϵ, and 2) the optimal threshold is an increasing function in the number of the observed processes K.

---

[1]The OU process is the continuous-time analogue of the first-order autoregressive process [22], [23], and is used to model various physical phenomena, and has relevant applications in control and finance.

II. System Model

We consider a sensing system in which K independent, OU processes are remotely monitored using a *shared* sensor that transmits samples from the processes over an erasure channel to a receiver. Denote the kth process value at time s by $X_s^{[k]}$. Given $X_s^{[k]}$, the kth process evolves, for $t \geq s$, as [22], [23]

$$X_t^{[k]} = X_s^{[k]} e^{-\theta_k(t-s)} + \frac{\sigma_k}{\sqrt{2\theta_k}} e^{-\theta_k(t-s)} W_{e^{2\theta_k(t-s)}-1}, \quad (1)$$

where W_t denotes a Wiener process, while $\theta_k > 0$ and $\sigma_k > 0$ are fixed parameters that control how fast the process evolves. The processes are initiated as $X_0^{[k]} \sim \mathcal{N}\left(0, \sigma_k^2/2\theta_k\right)$.

To estimate the kth process at the receiver, the sensor observes the kth process at time instants $\left\{ S_i^{[k]} \right\}$ and sends the samples to the receiver. Samples are *generated-at-will*. We focus on *signal-independent* sampling policies, where optimal sampling instants depend on the processes' statistics and not on exact processes' values. The sensor must obey a *total sampling frequency constraint* $f_{\max}$. Let ℓ_i denote the ith sampling instance *regardless* of the identity of the process being sampled. We write the sampling constraint as,

$$\liminf_{n \to \infty} \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^{n} \ell_{i+1} - \ell_i \right] \geq \frac{1}{f_{\max}}, \quad (2)$$

Samples go through a shared processing queue, whose service model follows a Poisson process with service rate μ. Served samples are prune to erasures with probability ϵ independently across samples. Immediate erasure status feedback is available.

Samples are time-stamped prior to transmissions. The *age-of-information* (AoI) of the kth process, denoted $\texttt{AoI}^{[k]}(t)$, is defined as the time elapsed since the latest successfully received sample's time stamp. We focus on *Maximum-Age-First (MAF) scheduling*, in which the sampling priority given to the process with highest AoI. Hence, at time t, process

$$\kappa(t) \triangleq \arg\max_k \texttt{AoI}^{[k]}(t) \quad (3)$$

is sampled. The value of $\kappa(t)$ will *not* change unless a successful transmission occurs. Therefore, in case of erasure events, a *fresh* sample is generated from the *same* process being served. Under MAF scheduling, each process is sampled at a rate of $f_{\max}/K$, and the constraint in (2) becomes

$$\liminf_{n \to \infty} \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^{n} S_{i+1}^{[k]} - S_i^{[k]} \right] \geq \frac{K}{f_{\max}}, \quad \forall k. \quad (4)$$

Let $\tilde{S}_i^{[k]}$ denote the sampling instance of the ith *successfully* received sample from the kth process. Let $S_i^{[k]}(m)$ be the sampling instance of the mth *attempt* to convey the ith sample of the kth process, $m = 1, \ldots, M_i^{[k]}$, with $M_i^{[k]} \sim \text{geometric}(1-\epsilon)$, denoting the number of trials. Each attempt incurs an i.i.d. service time, $Y_i^{[k]}(m) \sim \exp(\mu)$. The successfully received sample arrives at the receiver at $D_i^{[k]}$,

$$D_i^{[k]} = \tilde{S}_i^{[k]} + Y_i^{[k]}\left(M_i^{[k]}\right). \quad (5)$$

The AoI of the kth OU process is as follows:

$$\texttt{AoI}^{[k]}(t) = t - \tilde{S}_i^{[k]}, \qquad D_i^{[k]} \leq t < D_{i+1}^{[k]}. \quad (6)$$

The receiver constructs minimum mean square error (MMSE) estimates using the collected samples. Since the processes are independent, and by the strong Markov property of the OU process, the MMSE estimate for the kth process, $\hat{X}_t^{[k]}$, is based solely on the latest received sample from that process. Thus, for $D_i^{[k]} \leq t < D_{i+1}^{[k]}$, we have [19], [20]

$$\hat{X}_t^{[k]} = \mathbb{E}\left[X_t^{[k]} \Big| \tilde{S}_i^{[k]}, X_{\tilde{S}_i^{[k]}} \right] \overset{(1)}{=} X_{\tilde{S}_i^{[k]}} e^{-\theta_k(t-\tilde{S}_i^{[k]})}. \quad (7)$$

Hence, the instantaneous mean square error (MSE) in estimating the kth process at time $t \in \left[D_i^{[k]}, D_{i+1}^{[k]} \right)$ is [19], [20]

$$\texttt{mse}^{[k]}\left(t, \tilde{S}_i^{[k]}\right) = \frac{\sigma_k^2}{2\theta_k}\left(1 - e^{-2\theta_k\left(t-\tilde{S}_i^{[k]}\right)} \right), \quad (8)$$

which is an increasing function of the AoI in (6). Next, we define the long-term time average MSE of the kth process as

$$\overline{\texttt{mse}^{[k]}} \triangleq \limsup_{T \to \infty} \frac{\sum_{i=1}^{T} \mathbb{E}\left[\int_{D_i^{[k]}}^{D_{i+1}^{[k]}} \texttt{mse}^{[k]}\left(t, \tilde{S}_i^{[k]}\right) dt \right]}{\sum_{i=1}^{T} \mathbb{E}\left[D_{i+1}^{[k]} - D_i^{[k]} \right]}. \quad (9)$$

Our goal is to choose the sampling instants to minimize a penalty function $g(\cdot)$ of $\left\{ \overline{\texttt{mse}^{[k]}} \right\}$. More specifically, to solve

$$\min_{\{S_i^{[k]}(m)\}} \quad g\left(\overline{\texttt{mse}^{[1]}}, \cdots, \overline{\texttt{mse}^{[K]}} \right)$$

$$\text{s.t.} \quad \liminf_{n \to \infty} \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^{n} S_{i+1}^{[k]} - S_i^{[k]} \right] \geq \frac{K}{f_{\max}}, \quad \forall k. \quad (10)$$

III. Stationary Policies: Problem Re-Formulation

We re-formulate problem (10) in terms of a *waiting policy*. We define $W_i^{[k]}(m)$ as the mth waiting time before taking the mth sample towards conveying the ith sample from the kth process, $1 \leq m \leq M_i^{[k]}$. Without loss of generality, let the MAF schedule be in the order $1, 2, \ldots, K$. Thus,

$$S_i^{[k]}(m) = D_i^{[k-1]} + \sum_{j=1}^{m-1} Y_i^{[k]}(j) + \sum_{j=1}^{m} W_i^{[k]}(j), \quad (11)$$

with $D_i^{[0]} \triangleq D_{i-1}^{[K]}$. We define the ith *epoch* of the kth process, $\Gamma_i^{[k]}$, as the inter-reception time in between its ith and $(i+1)$th *unerased* samples, i.e., $\Gamma_i^{[k]} = D_{i+1}^{[k]} - D_i^{[k]}$.

We focus on *stationary waiting policies* in which the waiting policy $\left\{ W_i^{[k]}(m) \right\}$ has the same distribution across all processes' epochs. Under MAF scheduling, each process epoch entails a successful transmission of every other process. This induces a stationary distribution across all processes' epochs given that the service times and erasures are i.i.d. Therefore, dropping the indices i and k, we have $\Gamma_i^{[k]} \sim \Gamma$, $\forall i, k$, where

$$\Gamma = \sum_{k=1}^{K} \sum_{m=1}^{M^{[k]}} W^{[k]}(m) + Y^{[k]}(m). \quad (12)$$

By stationarity, one can write (9) for a typical epoch as

$$\overline{\mathrm{mse}^{[k]}} = \frac{\mathbb{E}\left[\int_{D^{[k]}}^{D^{[k]}+\Gamma} \mathrm{mse}^{[k]}\left(t, \tilde{S}^{[k]}\right) dt\right]}{\mathbb{E}\left[\Gamma\right]}. \quad (13)$$

where $D_i^{[k]} \sim D^{[k]}$ and $\tilde{S}_i^{[k]} \sim \tilde{S}^{[k]}$, $\forall i$. In the sequel, *we treat the Kth (last) process's epoch as the typical epoch.*

In the next lemma, we show that one can achieve the same long-term average MSE penalty by *grouping all the waiting times at the beginning of the (typical) epoch.*

Lemma 1 *Under signal-independent sampling with MAF scheduling and stationary waiting policies, problem* (10) *is equivalent to the following optimization problem:*

$$\min_{W \geq 0} \quad g\left(\overline{mse^{[1]}}, \cdots, \overline{mse^{[K]}}\right)$$

$$s.t. \quad \mathbb{E}\left[(1-\epsilon)W + \sum_{k=1}^{K} Y^{[k]}\right] \geq \frac{K}{f_{\max}}, \quad (14)$$

where $W \triangleq \sum_{k=1}^{K} \sum_{m=1}^{M^{[k]}} W^{[k]}(m)$ and the waiting is only performed at the beginning of the epoch.

Proof: Observing the average MSE function in (13), the waiting times appear in the numerator and denominator as the sum $\sum_{k=1}^{K} \sum_{m=1}^{M^{[k]}} W^{[k]}(m)$. Thus, for the optimal waiting times $\left\{W^{[k]^*}(m)\right\}$ that solve (10), the waiting time $W^* = \sum_{k=1}^{K} \sum_{m=1}^{M^{[k]}} W^{[k]^*}(m)$ achieves the same $\overline{\mathrm{mse}^{[k]}}$. Conversely, starting with W^* in (14) and breaking it arbitrarily to any waiting times such that $W^* = \sum_{k=1}^{K} \sum_{m=1}^{M^{[k]}} W^{[k]^*}(m)$ gives the same objective function in (10).

For the sampling constraint, we have that for process k,

$$\liminf_{n \to \infty} \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^{n} S_{i+1}^{[k]} - S_i^{[k]}\right] = \liminf_{n \to \infty} \frac{1}{n}\mathbb{E}\left[S_{n+1}^{[k]}\right]. \quad (15)$$

Define $e(n)$ to be the index of the epoch corresponding to the nth sample. Hence, we can write the sampling constraint as

$$\liminf_{n \to \infty} \frac{e(n)}{n} \cdot \frac{\mathbb{E}[S_{n+1}^{[k]}]}{e(n)}$$

$$= \frac{1}{\mathbb{E}[M^{[k]}]} \cdot \liminf_{n \to \infty} \frac{1}{e(n)}\left(\sum_{i=1}^{e(n)-1} \mathbb{E}\left[\sum_{k=1}^{K}\sum_{m=1}^{M_i^{[k]}} W_i^{[k]}(m)\right.\right.$$

$$\left.\left. + Y_i^{[k]}(m)\right] + o\left(e(n)\right)\right) \quad (16)$$

$$= \frac{1}{\mathbb{E}[M^{[k]}]} \cdot \liminf_{n \to \infty} \frac{1}{e(n)}\sum_{i=1}^{e(n)-1}\left(\mathbb{E}[W]\right.$$

$$\left. + \mathbb{E}\left[M_i^{[k]}\right] \cdot \mathbb{E}\left[\sum_{k=1}^{K} Y_i^{[k]}\right]\right) \quad (17)$$

$$= \mathbb{E}\left[(1-\epsilon)W + \sum_{k=1}^{K} Y^{[k]}\right], \quad (18)$$

where (16) follows from the strong law of large numbers and the fact that the time spent in the $e(n)$th epoch, $\Delta = \sum_{\tilde{k}=1}^{k-1}\sum_{m=1}^{M^{[\tilde{k}]}} W_i^{[\tilde{k}]}(m) + Y_i^{[\tilde{k}]}(m) + \sum_{m=1}^{\tilde{m}} W_i^{[k]}(m) + Y_i^{[k]}(m)$, is $o(e(n))$ and hence $\liminf_{n\to\infty}\frac{\Delta}{e(n)} = 0$, and (17) follows from Wald's identity. ∎

Remark 1 *The sampling constraint in problem (14) is not active if $f_{\max} > \mu$. This is due to that, in this case, the inter-sampling time, on average, is larger than the minimum allowable sampling time dictated by the sampling constraint.*

If the sampling constraint is binding ($f_{\max} < \mu$), the average waiting time monotonically increases with the erasure probability. This is true because no waiting is allowed in between unsuccessful transmissions. To account for the expected large number of back-to-back sample transmissions in the epoch, sampler waits for a relatively larger amount of time at its beginning so that the sampling constraint is satisfied.

IV. OPTIMAL THRESHOLD WAITING AND MINIMUM SUM MSE CHARACTERIZATION

We provide the optimal solution of problem (14) for a *sum MSE* penalty $g\left(\overline{\mathrm{mse}^{[1]}}, \cdots, \overline{\mathrm{mse}^{[K]}}\right) = \sum_{k=1}^{K} \overline{\mathrm{mse}^{[k]}}$, together with a *stationary deterministic* waiting policy, in which the waiting value at the beginning of an epoch is given by a deterministic function $w(\cdot)$ of the previous epoch's total service time, denoted $\tilde{Y} \sim \sum_{k=1}^{K} \sum_{m=1}^{M^{[k]}} Y^{[k]}(m)$. Such choice of waiting policies emerges naturally since the MSE is an increasing function of the AoI, whose value at the start of the epoch is, in turn, an increasing function of $\tilde{Y}$. Stationary deterministic policies have been used in similar contexts in the literature [18]–[20] and shown to perform optimally.

Formally, substituting the above into problem (14), we now aim at solving the following functional optimization problem:

$$\min_{w(\cdot) \geq 0} \quad \frac{\sum_{k=1}^{K} \mathbb{E}\left[\int_{D^{[k]}}^{D^{[k]}+\Gamma} \mathrm{mse}^{[k]}\left(t, \tilde{S}^{[k]}\right) dt\right]}{\mathbb{E}\left[\Gamma\right]}$$

$$s.t. \quad \mathbb{E}\left[w\left(\tilde{Y}\right)\right] \geq \frac{1}{1-\epsilon}\left(\frac{K}{f_{\max}} - \frac{K}{\mu}\right). \quad (19)$$

Theorem 1 provides the solution of problem (19). A proof sketch is given due to space limits. We use the compact vector notation $\boldsymbol{\theta} \triangleq [\theta_1 \ \theta_2 \ \cdots \ \theta_K]$ and $\boldsymbol{\sigma} \triangleq [\sigma_1^2 \ \sigma_2^2 \ \cdots \ \sigma_K^2]$.

Theorem 1 *The optimal waiting policy $w^*(\cdot)$ that solves problem (19) is given by the* threshold *policy*

$$w^*(z) = \left[\tau^*(K, f_{\max}, \epsilon, \boldsymbol{\theta}, \boldsymbol{\sigma}) - z\right]^+, \quad (20)$$

where the optimal threshold $\tau^*(K, f_{\max}, \epsilon, \boldsymbol{\theta}, \boldsymbol{\sigma})$ is given by

$$\tau^* = \max\left\{G_{\boldsymbol{\theta},\boldsymbol{\sigma}}^{-1}(\beta^*), H^{-1}\left(\frac{1}{(1-\epsilon)}\left[\frac{K}{f_{\max}} - \frac{K}{\mu}\right]^+\right)\right\}, \quad (21)$$

where $G_{\boldsymbol{\theta},\boldsymbol{\sigma}}(\tau) \triangleq \sum_{k=1}^{K} \frac{\sigma_k^2}{2\theta_k}\left(1 - \mathbb{E}\left[e^{-2\theta_k Y}\right]e^{-2\theta_k \tau}\right)$, *and* β^* *corresponds to the optimal long-term average sum MSE in this case, and is given by the unique solution of*

$$\sum_{k=1}^{K} \frac{\sigma_k^2}{2\theta_k}\left(H(\tau^*) + \frac{K}{\mu(1-\epsilon)} - \frac{1}{2\theta_k} \cdot \frac{\mu}{2\theta_k+\mu}(1-F_k(\tau^*))\right)$$
$$- \beta^*\left(H(\tau^*) + \frac{K}{\mu(1-\epsilon)}\right) = 0, \quad (22)$$

with $H(\cdot)$, *and* $F_k(\cdot)$ *defined in* (23), *and* (24), *respectively.*

Proof: [*Sketch.*] We apply Dinkelbach's approach [24] to transform the fractional objective function of (19) into a parameterized difference between the numerator and the denominator. This produces the auxiliary optimization problem

$$p(\beta) \triangleq \min_{w(\cdot)\geq 0} \quad \sum_{k=1}^{K} \mathbb{E}\left[\int_D^{D+\Gamma} \mathsf{mse}^{[k]}(t, \tilde{S}^{[k]})dt\right] - \beta\mathbb{E}\left[\Gamma\right]$$
$$\text{s.t.} \quad \mathbb{E}\left[w\left(\tilde{Y}\right)\right] \geq \frac{1}{1-\epsilon}\left(\frac{K}{f_{\max}} - \frac{K}{\mu}\right), \quad (25)$$

from which the optimal solution of problem (19) is given by β^* that uniquely solves $p(\beta^*) = 0$ [24].

Now for a fixed β, one can follow a Lagrangian approach to show that the optimal solution of problem (25) satisfies

$$\sum_{k=1}^{K} \frac{\sigma_k^2}{2\theta_k}\left(1 - \mathbb{E}\left[e^{-2\theta_k Y}\right]e^{-2\theta_k(w^*(z)+z)}\right)$$
$$= \beta + \zeta(1-\epsilon) + \frac{\eta(y)}{f_{\tilde{Y}}(z)}, \quad (33)$$

with $\eta(y)$ and ζ being Lagrange multipliers corresponding to the dual problem, and $f_{\tilde{Y}}(\cdot)$ is the probability density function of the total service time in the epoch. We define the left hand side of the above as $G_{\boldsymbol{\theta},\boldsymbol{\sigma}}(w^*(z)+z)$ given in the theorem, which is an increasing function. Thus, one can uniquely solve for $w^*(z)$ in terms of the Lagrange multipliers.

We then make use of the complementary slackness conditions and some involved mathematical manipulations to characterize the effect of the Lagrange multipliers on the optimal solution. Specifically, we define the function $H(\cdot)$ to denote the average waiting time and characterize it using the sampling constraint (when binding). The function $H(\cdot)$ depends on the distribution of $\tilde{Y}$, given by a convolution of a random number (that is geometrically distributed) of the $\exp(\mu)$ distribution. This gives rise to the incomplete Gamma function used in (23) and (24).

Finally, everything is combined by solving $p(\beta^*) = 0$. ∎

Theorem 1 shows that the sensor only takes a new sample in the epoch only if the previous epoch's *total* service time $\sim \tilde{Y}$ (of all processes) surpasses a certain threshold. It is emphasized in (20) that such threshold depends on the system parameters. This is highlighted in the next section.

V. NUMERICAL RESULTS

We present our numerical results concerning Theorem 1. Fig. 2 studies a 2-process system with $\boldsymbol{\theta} = [0.1 \quad 0.5]$,

and $\boldsymbol{\sigma} = [1 \quad 2]$, and service rate $\mu = 1$. We show the optimal threshold τ^* versus the erasure probability ϵ for $f_{\max} = 0.5, 0.95, , 1.5$. Our results show that for all sampling frequency constraints, the optimal threshold increases as the erasure probability increases. This is because that $G_{\boldsymbol{\theta},\boldsymbol{\sigma}}^{-1}(\cdot)$, and $H^{-1}(\cdot)$ are increasing functions in ϵ. We have three different cases. First, when $f_{\max} = 0.5$, the sampling frequency constraint is binding even at $\epsilon = 0$. The optimal threshold $\tau^* = H^{-1}\left(\frac{1}{1-\epsilon}\left[\frac{K}{f_{\max}} - \frac{K}{\mu}\right]\right) = H^{-1}\left(\frac{2}{1-\epsilon}\right)$. Thus, the optimal threshold is higher than the remaining cases and much steeper. Second, when $f_{\max} = 1.5$, the sampling frequency constraint is inactive as $f_{\max} > \mu$, and $\tau^* = G_{\boldsymbol{\theta},\boldsymbol{\sigma}}^{-1}(\beta^*)$ for all ϵ. Finally, when $f_{\max} = 0.95$, we observe an interesting behavior. When $\epsilon < \epsilon^* = 0.7$, the threshold corresponding to $G_{\boldsymbol{\theta},\boldsymbol{\sigma}}^{-1}(\beta^*)$ is (slightly) higher than the threshold corresponding to $H^{-1}\left(\frac{1}{1-\epsilon}\left[\frac{K}{f_{\max}} - \frac{K}{\mu}\right]\right)$ (which is shown as a dotted curve in Fig. 2), while for $\epsilon > \epsilon^* = 0.7$, the sampling frequency constraint becomes binding and therefore, the optimal threshold is characterized by $H^{-1}(\cdot)$ and becomes more steeper.

In Fig. 3, we consider a symmetric system with K processes, each having $\sigma_k^2 = 1$, and $\theta_k = 0.5$ with service rate $\mu = 1$. We study the optimal threshold versus the number of processes K. Fig. 3 shows that as K increases, the optimal threshold increases. The slope of the curve depends on $f_{\max}$. When $f_{\max} = 0.5$, the sampling frequency constraint is binding, and τ^* linearly increases with K with a steeper slope. When $f_{max} = 1.5 > \mu = 1$ (unconstrained problem), the optimal threshold is slowly increasing with K. For $f_{\max} = 0.95$, the optimal threshold matches the unconstrained solution for $K = 1, 2$. However, when $K > 2$, the sampling frequency constraint becomes binding and the linear-like profile prevails.

In Fig. 4, we consider a 2-process system with $\sigma_1^2 = 2$, $\sigma_2^2 = 1$, $\theta_1 = 0.5$ and $\mu = 1$. We vary $\theta_2 \in [0.1, 1]$ and observe the optimal threshold and the MMSE. When the sampling frequency constraint is binding, e.g., when $f_{\max} = 0.5$, the optimal threshold is independent of θ_2 as the argument of $H^{-1}(\cdot)$ is independent of θ_2. The optimal threshold monotonically decreases as θ_2 increases for $f_{\max} = 1.5$ as the system needs to wait less to track the variations in faster processes. In both cases, the long-term average MMSE is decreasing in θ_2 since the sum of the processes' variances decreases.

REFERENCES

[1] R. D. Yates, Y. Sun, D. R. Brown III, S. K. Kaul, E. Modiano, and S. Ulukus. Age of information: An introduction and survey. *IEEE J. Sel. Areas Commun.*, 39(5):1183–1210, May 2021.
[2] S. K. Kaul, R. D. Yates, and M. Gruteser. Real-time status: How often should one update? In *Proc. IEEE Infocom*, March 2012.
[3] C. Kam, S. Kompella, and A. Ephremides. Age of information under random updates. In *Proc. IEEE ISIT*, July 2013.
[4] A. Baknina, O. Ozel, J. Yang, S. Ulukus, and A. Yener. Sending information through status updates. In *Proc. IEEE ISIT*, June 2018.
[5] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor. Age-minimal transmission for energy harvesting sensors with finite batteries: Online policies. *IEEE Trans. Inf. Theory*, 66(1):534–556, January 2020.
[6] B. T. Bacinoglu, Y. Sun, E. Uysal-Biyikoglu, and V. Mutlu. Optimal status updating with a finite-battery energy harvesting source. *J. Commun. Netw.*, 21(3):280–294, June 2019.

$$H(\tau) = \sum_{\rho=K}^{\infty} \binom{\rho-1}{K-1} \epsilon^{\rho-K}(1-\epsilon)^K \left[\tau\gamma(\mu\tau,\rho) - \frac{\rho}{\mu}\gamma(\mu\tau,\rho+1) \right], \tag{23}$$

$$F_k(\tau) = \sum_{\rho=K}^{\infty} \binom{\rho-1}{K-1} \epsilon^{\rho-K}(1-\epsilon)^K \left[e^{-2\theta_k\tau}\gamma(\mu\tau,\rho) + \left(\frac{\mu}{2\theta_k+\mu}\right)^\rho (1-\gamma((2\theta_k+\mu)\tau,\rho)) \right], \tag{24}$$

where $\gamma(x,y)$ is the normalized incomplete Gamma function defined as $\gamma(x,y) = \frac{1}{(y-1)!}\int_0^x t^{y-1}e^{-t}dt$.

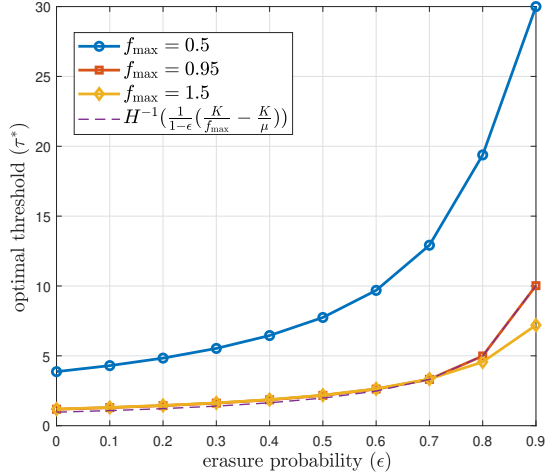Fig. 2. The optimal threshold (τ^*) versus the erasure probability (ϵ) for different sampling frequency constraints ($f_{\max}$).
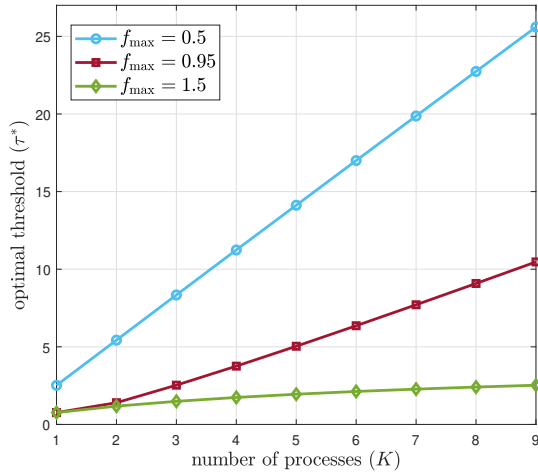

Fig. 4. Optimal threshold (τ^*) and optimal long-term average MMSE versus θ_2 for tracking two processes, where the first process has fixed parameters $\sigma_1^2 = 1$ and $\theta_1 = 1$ with different sampling frequency constraints ($f_{\max}$).

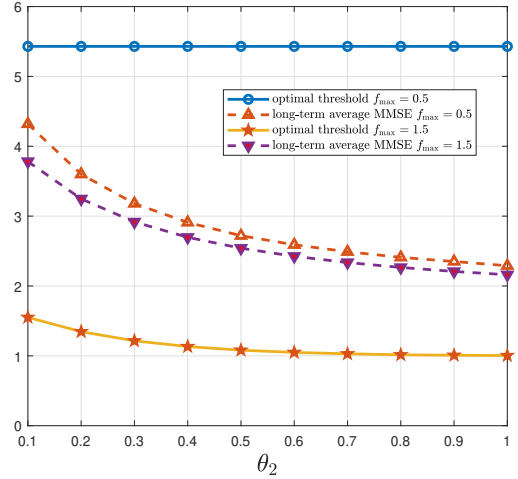Fig. 3. The optimal threshold (τ^*) versus the number of processes (K) for different sampling frequency constraints ($f_{\max}$).

[7] B. Buyukates, A. Soysal, and S. Ulukus. Age of information in multihop multicast networks. *J. Commun. Netw.*, 21(3):256–267, June 2019.

[8] H. H. Yang, A. Arafa, T. Q. S. Quek, and H. V. Poor. Age-based scheduling policy for federated learning in mobile edge networks. Available Online: arXiv:1910.14648.

[9] M. Klugel, M. H. Mamduhi, S. Hirche, and W. Kellerer. AoI-penalty minimization for networked control systems with packet loss. In *Proc. IEEE Infocom*, April 2019.

[10] A. Mitra, J. A. Richards, S. Bagchi, and S. Sundaram. Finite-time distributed state estimation over time-varying graphs: Exploiting the age-of-information. In *Proc. ACC*, July 2019.

[11] J. Chakravorty and A. Mahajan. Remote estimation over a packet-drop channel with Markovian state. *IEEE Trans. Autom. Control*, 65(5):2016–

2031, May 2020.

[12] O. Ayan, M. Vilgelm, M. Klugel, S. Hirche, and W. Kellerer. Age-of-information vs. value-of-information scheduling for cellular networked control systems. In *Proc. IEEE ICCPS*, April 2019.

[13] S. Roth, A. Arafa, H. V. Poor, and A. Sezgin. Remote short blocklength process monitoring: Trade-off between resolution and data freshness. In *Proc. IEEE ICC*, June 2020.

[14] D. Ramirez, E. Erkip, and H. V. Poor. Age of information with finite horizon and partial updates. Available Online: arXiv:1910.00963.

[15] M. Bastopcu and S. Ulukus. Age of information for updates with distortion: Constant and age-dependent distortion constraints. Available Online: arXiv:1912.13493.

[16] M. Bastopcu and S. Ulukus. Partial updates: Losing information for freshness. Available Online: arXiv:2001.11014.

[17] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides. The age of incorrect information: A new performance metric for status updates. Available Online: arXiv:1907.06604.

[18] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu. Remote estimation of the Wiener process over a channel with random delay. *IEEE Trans. Inf. Theory*, 66(2):1118–1135, February 2020.

[19] T. Ornee and Y. Sun. Sampling and remote estimation for the ornstein-uhlenbeck process through queues: Age of information and beyond. *IEEE/ACM Transactions on Networking*, 29(5):1962–1975, 2021.

[20] A. Arafa, K. Banawan, K. G. Seddik, and H. V. Poor. Sample, quantize, and encode: Timely estimation over noisy channels. *IEEE Trans. Commun.*, 69(10):6485–6499, October 2021.

[21] N. Guo and V. Kostina. Optimal causal rate-constrained sampling for a class of continuous Markov processes. *IEEE Trans. Inf. Theory*, 67(12):7876–7890, December 2021.

[22] G. E. Uhlenback and L. S. Ornstein. On the theory of the Brownian motion. *Phys. Rev.*, 36:823–841, September 1930.

[23] J. L. Doob. The Brownian movement and stochastic equations. *Ann. Math.*, 43(2):351–369, 1942.

[24] W. Dinkelbach. On nonlinear fractional programming. *Management Science*, 13(7):492–498, 1967.