Remote Estimation of the Wiener Process over a Channel with Random Delay

Yin Sun, Yury Polyanskiy, and Elif Uysal-Biyikoglu
Dept. of ECE, the Ohio State University, Columbus, OH
Dept. of EECS, Massachusetts Institute of Technology, Cambridge, MA
Dept. of EEE, Middle East Technical University, Ankara, Turkey

Abstract*—In this paper, we consider a problem of sampling a Wiener process, with samples forwarded to a remote estimator via a channel that consists of a queue with random delay. The estimator reconstructs a real-time estimate of the signal from causally received samples. Motivated by recent research on age-of-information, we study the optimal sampling strategy that minimizes the mean square estimation error subject to a sampling frequency constraint. We prove that the optimal sampling strategy is a threshold policy, and find the optimal threshold. This threshold is determined by the sampling frequency constraint and how much the Wiener process varies during the channel delay. An interesting consequence is that even in the absence of the sampling frequency constraint, the optimal strategy is *not* zero-wait sampling in which a new sample is taken once the previous sample is delivered; rather, it is optimal to wait for a non-zero amount of time after the previous sample is delivered, and then take the next sample. Further, if the sampling times are independent of the observed Wiener process, the optimal sampling problem reduces to an age-of-information optimization problem that has been recently solved. Our comparisons show that the estimation error of the optimal sampling policy is much smaller than those of age-optimal sampling, zero-wait sampling, and classic uniform sampling.

I. Introduction

Consider a system with two terminals (see Fig. 1): An observer measuring a Wiener process W_t and an estimator, whose goal is to provide the best-guess $\hat{W}_t$ for the current value of W_t. These two terminals are connected by a channel that transmits time-stamped samples of the form (S_i, W_{S_i}), where the sampling times S_i satisfy $0 \leq S_1 \leq S_2 \leq \ldots$ The channel is modeled as a work-conserving FIFO queue with random *i.i.d.* delay Y_i, where $Y_i \geq 0$ is the channel delay (i.e., transmission time) of sample i.[1] The observer can choose the sampling times S_i causally subject to an average sampling frequency constraint

$$\liminf_{n\to\infty} \frac{1}{n}\mathbb{E}[S_n] \geq \frac{1}{f_{\max}},$$

where $f_{\max}$ is the maximum allowed sampling frequency.

Unless it arrives at an empty system, sample i needs to wait in the queue until its transmission starts. Let G_i be the transmission starting time of sample i such that $S_i \leq G_i$. The delivery time of sample i is $D_i = G_i + Y_i$. The initial value $W_0 = 0$ is known by the estimator for free, represented by $S_0 = D_0 = 0$. At time t, the estimator forms $\hat{W}_t$ using causally received samples with $D_i \leq t$. By minimum mean square error (MMSE) estimation,

$$\begin{aligned}\hat{W}_t =&\mathbb{E}[W_t | W_{S_j}, D_j \leq t]\\ =&W_{S_i}, \text{ if } t \in [D_i, D_{i+1}), \ i = 0, 1, 2, \ldots,\end{aligned} \tag{1}$$

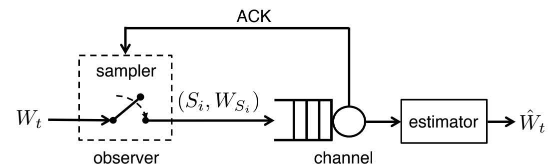


Fig. 1: System model.

as illustrated in Fig. 2. We measure the quality of remote estimation via the MMSE:

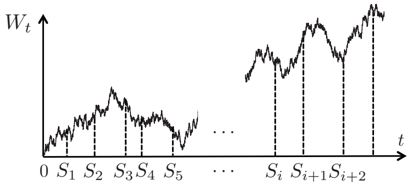
$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T (W_t - \hat{W}_t)^2 dt\right].$$

In this paper, we study the optimal sampling strategy that achieves the fundamental tradeoff between $f_{\max}$ and MMSE. The contributions of this paper are summarized as follows:

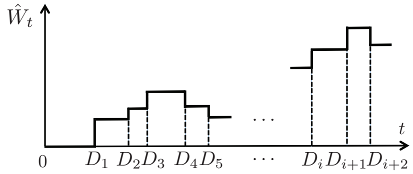
- The optimal sampling problem for minimizing the MMSE subject to the sampling frequency constraint is solved exactly. We prove that the optimal sampling strategy is a threshold policy, and find the optimal threshold. This threshold is determined by $f_{\max}$ and the amount of signal variation during the channel delay (i.e., random transmission time of a sample). Our threshold policy has an important difference from the previous threshold policies studied in, e.g., [1]–[10]: In our model, each sample waiting in the queue for its transmission opportunity unnecessarily becomes stale. We have proven that it is suboptimal to take a new sample when the channel is busy. Consequently, the threshold should be *disabled* whenever there is a packet in transmission.
- An unexpected consequence of our study is that even in the absence of the sampling frequency constraint (i.e., $f_{\max} = \infty$), the optimal strategy is *not* zero-wait sampling in which a new sample is generated once the previous sample is delivered; rather, it is optimal to wait a positive amount of time after the previous sample is delivered, and then take the next sample.
- If the sampling times are independent of the observed Wiener process, the optimal sampling problem reduces to an age-of-information optimization problem solved in [11], [12]. The asymptotics of the MMSE-optimal and age-

The research was supported in part by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-09-39370, by ONR grant N00014-17-1-2417, and by TUBITAK.

[1] By "work-conserving", we meant that the channel is kept busy whenever there exist some generated samples that are not delivered to the estimator.

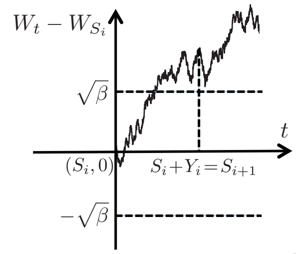


(a) Wiener process W_t and its samples.

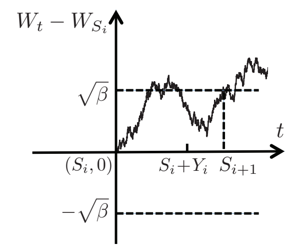


(b) Estimate process $\hat{W}_t$ using causally received samples.

Fig. 2: Sampling and remote estimation of the Wiener process.



(a) If $|W_{S_i+Y_i} - W_{S_i}| \geq \sqrt{\beta}$, sample $i+1$ is taken at time $S_{i+1} = S_i + Y_i$.



(b) If $|W_{S_i+Y_i} - W_{S_i}| < \sqrt{\beta}$, sample $i+1$ is taken at time t that satisfies $t \geq S_i + Y_i$ and $|W_t - W_{S_i}| = \sqrt{\beta}$.

Fig. 3: Illustration of the threshold policy (4).

optimal sampling policies at low/high channel delay or low/high sampling frequencies are studied.
- Our theoretical and numerical comparisons show that the MMSE of the optimal sampling policy is much smaller than those of age-optimal sampling, zero-wait sampling, and classic uniform sampling.

II. Related Work

On the one hand, the results in this paper are closely related to the recent age-of-information studies, e.g., [11]–[20], where the focus was on queueing and channel delay, without a signal model. The discovery that the zero-wait policy is not always optimal for minimizing the age-of-information can be found in [11]–[13]. The sub-optimality of a work-conserving scheduling policy was also observed in [19], which considered scheduling updates to different users with unreliable channels. One important observation in our study is that the behavior of the optimal update policy changes dramatically after adding a signal model.

On the other hand, the paper can be considered as a contribution to the rich literature on remote estimation, e.g., [1]–[10], [21], by adding a queueing model. Optimal transmission scheduling of sensor measurements for estimating a stochastic process was recently studied in [9], [10], where the samples are transmitted over a channel with additive noise. In the absence of channel delay and queueing (i.e., $Y_i = 0$), the problems of sampling Wiener process and Gaussian random walk were addressed in [1], [7], [8], where the optimality of threshold policies was established. To the best of our knowledge, [7] is the closest study with this paper. Because there is no queueing and channel delay in [7], the problem analyzed therein is a special case of ours.

III. Main Result

Let $\pi = (S_0, S_1, \ldots)$ represent a sampling policy, and Π be the set of *causal* sampling policies which satisfy the following conditions: (i) The *information* that is available for determining the sampling time S_i includes the history of the Wiener process $(W_t : t \in [0, S_i])$, the history of channel states $(I_t : t \in [0, S_i])$, and the sampling times of previous samples $(S_0, \ldots, S_{i-1})$, where $I_t \in \{0, 1\}$ is the idle/busy state of the channel at time t. (ii) The inter-sampling times $\{T_i = S_{i+1} - S_i, i = 0, 1, \ldots\}$ form a *regenerative process* [22, Section 6.1]: There exist integers $0 \leq k_1 < k_2 < \ldots$ such that the post-k_j process $\{T_{k_j+i}, i = 0, 1, \ldots\}$ has the same distribution as the post-k_1 process $\{T_{k_1+i}, i = 0, 1, \ldots\}$ and is independent of the pre-k_j process $\{T_i, i = 0, 1, \ldots, k_j - 1\}$; in addition, $\mathbb{E}[S_{k_1}^2] < \infty$ and $0 < \mathbb{E}[(S_{k_{j+1}} - S_{k_j})^2] < \infty$ for $j = 1, 2, \ldots$[2] We assume that the Wiener process W_t and the channel delay Y_i are mutually independent and do not change according to the sampling policy. We also assume $\mathbb{E}[Y_i^2] < \infty$.

A sampling policy $\pi \in \Pi$ is said to be *signal-independent* (*signal-dependent*), if π is (not) independent of the Wiener process $\{W_t, t \geq 0\}$. Example policies in Π include:

1. *Uniform sampling*: The inter-sampling times are constant, such that for some $\beta \geq 0$,
$$S_{i+1} = S_i + \beta. \tag{2}$$
2. *Zero-wait sampling* [11]–[14]: A new sample is generated once the previous sample is delivered, i.e.,
$$S_{i+1} = S_i + Y_i. \tag{3}$$
3. *Threshold policy in signal variation:* The sampling times are given by
$$S_{i+1} = \inf\left\{t \geq S_i + Y_i : |W_t - W_{S_i}| \geq \sqrt{\beta}\right\}, \tag{4}$$
which is illustrated in Fig. 3. If $|W_{S_i+Y_i} - W_{S_i}| \geq \sqrt{\beta}$, sample $i+1$ is generated at the time $S_{i+1} = S_i + Y_i$ when sample i is delivered; otherwise, if $|W_{S_i+Y_i} - W_{S_i}| < \sqrt{\beta}$, sample $i+1$ is generated at the earliest time t such that $t \geq S_i + Y_i$ and $|W_t - W_{S_i}|$ reaches the threshold $\sqrt{\beta}$. It is worthwhile to emphasize that even if there exists time $t \in [S_i, S_i + Y_i)$ such that $|W_t - W_{S_i}| \geq \sqrt{\beta}$, no sample is taken at such time t, as depicted in Fig. 3. In other words, the threshold-based control is disabled during $[S_i, S_i + Y_i)$ and is reactivated at time $S_i + Y_i$. This is a key difference from previous studies on threshold policies [1]–[10].
4. *Threshold policy in time variation* [11]–[13]: The sampling times are given by
$$S_{i+1} = \inf\left\{t \geq S_i + Y_i : t - S_i \geq \beta\right\}. \tag{5}$$

The optimal sampling problem for minimizing the MMSE

[2] Really, we assume that T_i is a regenerative process because we analyze the time-average MMSE in (6), but operationally a nicer definition is $\limsup_{n\to\infty} \mathbb{E}[\int_0^{D_n} (W_t - \hat{W}_t)^2 dt]/\mathbb{E}[D_n]$. These two definitions are equivalent when T_i is a regenerative process.

subject to a sampling frequency constraint is formulated as

$$\min_{\pi\in\Pi} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T (W_t-\hat{W}_t)^2dt\right] \tag{6}$$

$$\text{s.t.} \liminf_{n\to\infty} \frac{1}{n}\mathbb{E}[S_n] \geq \frac{1}{f_{\max}}. \tag{7}$$

Problem (6) is a constrained continuous-time Markov decision problem with a continuous state space. Somewhat to our surprise, we were able to exactly solve (6):

Theorem 1. *There exists $\beta \geq 0$ such that the sampling policy* (4) *is optimal to* (6)*, and the optimal β is determined by solving*[3]

$$\mathbb{E}[\max(\beta, W_Y^2)] = \max\left(\frac{1}{f_{\max}}, \frac{\mathbb{E}[\max(\beta^2, W_Y^4)]}{2\beta}\right), \tag{8}$$

where Y is a random variable with the same distribution as Y_i. The optimal value of (6) *is then given by*

$$\mathsf{mmse}_{\text{opt}} \triangleq \frac{\mathbb{E}[\max(\beta^2, W_Y^4)]}{6\mathbb{E}[\max(\beta, W_Y^2)]} + \mathbb{E}[Y]. \tag{9}$$

Proof. See Section IV. □

The optimal policy in (4) and (8) is called the "MMSE-optimal" policy. Note that one can use the bisection method or other one-dimensional search method to solve (8) with quite low complexity. Interestingly, this optimal policy does not suffer from the "curse of dimensionality" issue encountered in many Markov decision problems.

Notice that the feasible policies in Π can use the complete history of the Wiener process $(W_t : t \in [0, S_{i+1}])$ to determine S_{i+1}. However, the MMSE-optimal policy in (4) and (8) only requires recent knowledge of the Wiener process $(W_t - W_{S_i} : t \in [S_i + Y_i, S_{i+1}])$ to determine S_{i+1}.

Moreover, according to (8), the threshold $\sqrt{\beta}$ is determined by the maximum sampling frequency $f_{\max}$ and the distribution of the signal variation W_Y during the channel delay Y. It is worth noting that W_Y *is a random variable that tightly couples the source process and the channel delay*. This is different from the traditional wisdom of information theory where source coding and channel coding can be treated separately.

A. Signal-Independent Sampling and the Age-of-Information

Let $\Pi_{\text{sig-independent}} \subset \Pi$ denote the set of signal-independent sampling policies, defined as

$$\Pi_{\text{sig-independent}} = \{\pi \in \Pi : \pi \text{ is independent of } W_t, t \geq 0\}.$$

For each $\pi \in \Pi_{\text{sig-independent}}$, the MMSE (6) can be written as

$$\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T \Delta(t)dt\right], \tag{10}$$

where

$$\Delta(t) = t - S_i,\ t \in [D_i, D_{i+1}),\ i = 0, 1, 2, \ldots, \tag{11}$$

is the *age-of-information* [14], that is, the time difference between the generation time of the freshest received sample and the current time t. If the policy space in (6) is restricted from Π to $\Pi_{\text{sig-independent}}$, (6) reduces to the following age-of-information optimization problem [11], [12]:

$$\min_{\pi\in\Pi_{\text{sig-independent}}} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T \Delta(t)dt\right] \tag{12}$$

$$\text{s.t.} \liminf_{n\to\infty} \frac{1}{n}\mathbb{E}[S_n] \geq \frac{1}{f_{\max}}.$$

Problem (6) is significantly more challenging than (12), because in (6) the sampler needs to make decisions based on the evolution of the signal process W_t, which is not required in (12). More powerful techniques than those in [11], [12] are developed in Section IV and our technical report [23] to solve (6).

Theorem 2. *[11], [12] There exists $\beta \geq 0$ such that the sampling policy* (5) *is optimal to* (12)*, and the optimal β is determined by solving*

$$\mathbb{E}[\max(\beta, Y)] = \max\left(\frac{1}{f_{\max}}, \frac{\mathbb{E}[\max(\beta^2, Y^2)]}{2\beta}\right). \tag{13}$$

The optimal value of (12) *is then given by*

$$\mathsf{mmse}_{\text{age-opt}} \triangleq \frac{\mathbb{E}[\max(\beta^2, Y^2)]}{2\mathbb{E}[\max(\beta, Y)]} + \mathbb{E}[Y]. \tag{14}$$

The sampling policy in (5) and (13) is referred to as the "age-optimal" policy. Because $\Pi_{\text{sig-independent}} \subset \Pi$,

$$\mathsf{mmse}_{\text{opt}} \leq \mathsf{mmse}_{\text{age-opt}}. \tag{15}$$

In the following, the asymptotics of the MMSE-optimal and age-optimal sampling policies at low/high channel delay or low/high sampling frequencies are studied.

B. Low Channel Delay or Low Sampling Frequency

Let $Y_i = dX_i$ represent the scaling of the channel delay Y_i with d, where $d \geq 0$ and the X_i's are *i.i.d.* positive random variables. If $d \to 0$ or $f_{\max} \to 0$, we can obtain from (8) that

$$\beta = \frac{1}{f_{\max}} + o\left(\frac{1}{f_{\max}}\right), \tag{16}$$

where $f(x) = o(g(x))$ as $x \to a$ means that $\lim_{x\to a} f(x)/g(x) = 0$. Hence, the MMSE-optimal policy becomes

$$S_{i+1} = \inf\left\{t \geq S_i : |W_t - W_{S_i}| \geq \sqrt{\frac{1}{f_{\max}}}\right\}, \tag{17}$$

and the optimal value of (6) becomes $\mathsf{mmse}_{\text{opt}} = 1/(6f_{\max}) + o(1/f_{\max})$. The sampling policy (17) was also obtained in [7] for the case that $Y_i = 0$ for all i.

If $d \to 0$ or $f_{\max} \to 0$, one can show that the age-optimal policy in (5) and (13) becomes uniform sampling (2) with $\beta = 1/f_{\max} + o(1/f_{\max})$, and the optimal value of (12) is $\mathsf{mmse}_{\text{age-opt}} = 1/(2f_{\max}) + o(1/f_{\max})$. Therefore,

$$\lim_{d\to 0} \frac{\mathsf{mmse}_{\text{opt}}}{\mathsf{mmse}_{\text{age-opt}}} = \lim_{f_{\max}\to 0} \frac{\mathsf{mmse}_{\text{opt}}}{\mathsf{mmse}_{\text{age-opt}}} = \frac{1}{3}. \tag{18}$$

C. High Channel Delay or Unbounded Sampling Frequency

If $d \to \infty$ or $f_{\max} \to \infty$, the MMSE-optimal policy for solving (6) is given by (4) where β is determined by solving

$$2\beta\mathbb{E}[\max(\beta, W_Y^2)] = \mathbb{E}[\max(\beta^2, W_Y^4)]. \tag{19}$$

[3] If $\beta \to 0$, the last terms in (8) and (13) are determined by L'Hôpital's rule.

Similarly, if $d \to \infty$ or $f_{\max} \to \infty$, the age-optimal policy for solving (12) is given by (5) where β is determined by solving

$$2\beta\mathbb{E}[\max(\beta, Y)] = \mathbb{E}[\max(\beta^2, Y^2)]. \tag{20}$$

In these limits, the ratio between $\mathsf{mmse}_{\text{opt}}$ and $\mathsf{mmse}_{\text{age-opt}}$ depends on the distribution of Y.

When the sampling frequency is unbounded, i.e., $f_{\max} = \infty$, one logically reasonable policy is the zero-wait policy in (3) [11]–[14]. This zero-wait policy achieves the maximum throughput and the minimum queueing delay of the channel. Surprisingly, this zero-wait policy *does not always* minimize the age-of-information in (12) and *almost never* minimizes the MMSE in (6), as stated below:

Theorem 3. *If $f_{\max} = \infty$, the zero-wait policy is optimal for solving* (6) *if and only if $Y = 0$ with probability one.*

Theorem 4. *[12] If $f_{\max} = \infty$, the zero-wait policy is optimal for solving* (12) *if and only if*

$$\mathbb{E}[Y^2] \leq 2 \operatorname{ess\,inf} Y \ \mathbb{E}[Y], \tag{21}$$

where $\operatorname{ess\,inf} Y = \sup\{y \in [0, \infty) : \Pr[Y < y] = 0\}$.

Theorems 3 and 4 are proven in our technical report [23].

IV. Proof Sketch of the Main Result

A proof sketch of Theorem 1 is provided here, and the detailed proof is relegated to our technical report [23]:

We first provide a lemma that is crucial for simplifying (6).

Lemma 1. *In the optimal sampling problem* (6) *for minimizing the MMSE of the Wiener process, it is suboptimal to take a new sample before the previous sample is delivered.*

In recent studies on age-of-information [11], [12], Lemma 1 was intuitive and hence was used without a proof: If a sample is taken when the channel is busy, it needs to wait in the queue until its transmission starts, and becomes stale while waiting. A better method is to wait until the channel becomes idle, and then generate a new sample, as stated in Lemma 1. However, this lemma is not intuitive in the MMSE minimization problem (6): The proof of Lemma 1 relies on the strong Markov property of Wiener process, which may not hold for other signal processes.

By Lemma 1, we only need to consider a sub-class of sampling policies $\Pi_1 \subset \Pi$ defined by

$$\Pi_1 = \{\pi \in \Pi : S_{i+1} = G_{i+1} \geq D_i \text{ for all } i\}.$$

This completely eliminates the waiting time wasted in the queue, and hence the queue should always be kept empty. Let $Z_i = S_{i+1} - D_i \geq 0$ represent the *waiting time* between the delivery time D_i of sample i and the generation time S_{i+1} of sample $i+1$. Then, $S_i = Z_0 + \sum_{j=1}^{i-1}(Y_j + Z_j)$ and $D_i = \sum_{j=0}^{i-1}(Z_j + Y_{j+1})$ for each $i = 1, 2, \ldots$ If $(Y_1, Y_2, \ldots)$ is given, $(S_0, S_1, \ldots)$ is uniquely determined by $(Z_0, Z_1, \ldots)$. Hence, one can also use $\pi = (Z_0, Z_1, \ldots)$ to represent a sampling policy.

Because T_i is a regenerative process, (6) can be reformulated as the following Markov decision problem:

$$\mathsf{mmse}_{\text{opt}} \triangleq \min_{\pi \in \Pi_1} \lim_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}\left[\int_{D_i}^{D_{i+1}} (W_t - W_{S_i})^2 dt\right]}{\sum_{i=0}^{n-1} \mathbb{E}\left[Y_i + Z_i\right]} \tag{22}$$

$$\text{s.t. } \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}\left[Y_i + Z_i\right] \geq \frac{1}{f_{\max}}, \tag{23}$$

Define the σ-fields $\mathcal{F}_t^s = \sigma(W_{s+v} - W_s : v \in [0, t])$ and $\mathcal{F}_t^{s+} = \cap_{v>t}\mathcal{F}_v^s$, as well as the filtration (i.e., a non-decreasing and right-continuous family of σ-fields) $\{\mathcal{F}_t^{s+}, t \geq 0\}$ of the time-shifted Wiener process $\{W_{s+t} - W_s, t \in [0, \infty)\}$. Let $\mathfrak{M}_s$ denote the set of square-integrable stopping times of $\{W_{s+t} - W_s, t \in [0, \infty)\}$, i.e.,

$$\mathfrak{M}_s = \{\tau \geq 0 : \{\tau \leq t\} \in \mathcal{F}_t^{s+}, \mathbb{E}\left[\tau^2\right] < \infty\}.$$

After some manipulations, we formulate the Lagrangian dual problem of (22), which can be decomposed into a sequence of *mutually independent* per-sample control problem (24):

$$Z_i = \min_{\tau \in \mathfrak{M}_{S_i+Y_i}} \mathbb{E}\left[\frac{1}{2}(W_{S_i+Y_i+\tau} - W_{S_i})^4 - \beta(Y_i + \tau) \middle| W_{S_i+Y_i} - W_{S_i}, Y_i\right], \tag{24}$$

where $\beta \geq 0$ is the dual variable. In the proof of (24), we have used the following lemma:

Lemma 2. *Let $\tau \geq 0$ be a stopping time of the Wiener process W_t with $\mathbb{E}[\tau^2] < \infty$, then*

$$\mathbb{E}\left[\int_0^{\tau} W_t^2 dt\right] = \frac{1}{6}\mathbb{E}\left[W_\tau^4\right]. \tag{25}$$

Using the optimal stopping rule developed in [24], we get

Theorem 5. *An optimal solution to* (24) *is*

$$Z_i = \inf\left\{t \geq 0 : |W_{S_i+Y_i+t} - W_{S_i}| \geq \sqrt{\beta}\right\}. \tag{26}$$

Notice that (26) is equivalent to (4).

Theorem 6. *The following assertions are true:*

(a). The duality gap between (22) *and its Lagrangian dual problem is zero.*

(b). A common optimal solution to (6) *and* (22) *is given by* (4) *and* (8).

By this, Theorem 1 is proven.

V. Numerical Results

In this section, we evaluate the estimation performance achieved by the following four sampling policies:

1. *Uniform sampling:* The policy in (2) with $\beta = f_{\max}$.
2. *Zero-wait sampling [11]–[14]:* The sampling policy in (3), which is feasible when $f_{\max} \geq \mathbb{E}[Y_i]$.
3. *Age-optimal sampling [11], [12]:* The sampling policy in (5) and (13), which is the optimal solution to (12).
4. *MMSE-optimal sampling:* The sampling policy in (4) and (8), which is the optimal solution to (6).

Let $\mathsf{mmse}_{\text{uniform}}$, $\mathsf{mmse}_{\text{zero-wait}}$, $\mathsf{mmse}_{\text{age-opt}}$, and $\mathsf{mmse}_{\text{opt}}$, be the MMSEs of uniform sampling, zero-wait sampling, age-

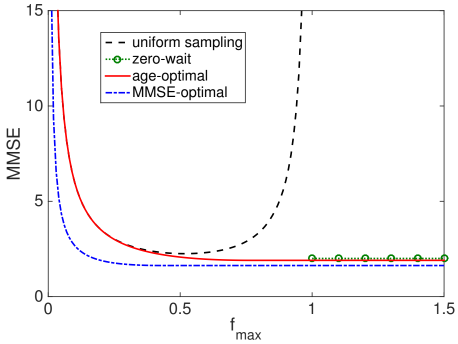


Fig. 4: MMSE vs. $f_{\max}$ tradeoff for *i.i.d.* exponential channel delay.

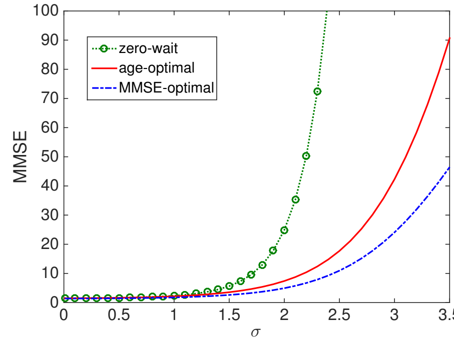


Fig. 5: MMSE vs. the scale parameter σ of *i.i.d.* log-normal channel delay for $f_{\max} = 1.5$.

optimal sampling, MMSE-optimal sampling, respectively. According to (15), as well as the facts that uniform sampling is feasible for (12) and zero-wait sampling is feasible for (12) when $f_{\max} \geq \mathbb{E}[Y_i]$, we can obtain

$$\mathsf{mmse}_{\text{opt}} \leq \mathsf{mmse}_{\text{age-opt}} \leq \mathsf{mmse}_{\text{uniform}},$$
$$\mathsf{mmse}_{\text{opt}} \leq \mathsf{mmse}_{\text{age-opt}} \leq \mathsf{mmse}_{\text{zero-wait}}, \text{ when } f_{\max} \geq \mathbb{E}[Y_i],$$

which fit with our numerical results below.

Figure 4 depicts the tradeoff between MMSE and $f_{\max}$ for *i.i.d.* exponential channel delay with mean $\mathbb{E}[Y_i] = 1/\mu = 1$. Hence, the maximum throughput of the channel is $\mu = 1$. In this setting, $\mathsf{mmse}_{\text{uniform}}$ is characterized by eq. (25) of [14], which was obtained using a D/M/1 queueing model. For small values of $f_{\max}$, age-optimal sampling is similar with uniform sampling, and hence $\mathsf{mmse}_{\text{age-opt}}$ and $\mathsf{mmse}_{\text{uniform}}$ are of similar values. However, as $f_{\max}$ approaches the maximum throughput 1, $\mathsf{mmse}_{\text{uniform}}$ increases to infinite. This is because the queue length in uniform sampling is large at high sampling frequencies, and the samples become stale during their long waiting times in the queue. On the other hand, $\mathsf{mmse}_{\text{opt}}$ and $\mathsf{mmse}_{\text{age-opt}}$ decrease with respect to $f_{\max}$. The reason is that the set of feasible policies satisfying the constraints in (6) and (12) becomes larger as $f_{\max}$ grows, and hence the optimal values of (6) and (12) are decreasing in $f_{\max}$. Moreover, the gap between $\mathsf{mmse}_{\text{opt}}$ and $\mathsf{mmse}_{\text{age-opt}}$ is large for small values of $f_{\max}$. The ratio $\mathsf{mmse}_{\text{opt}}/\mathsf{mmse}_{\text{age-opt}}$ tends to $1/3$ as $f_{\max} \to 0$, which is in accordance with (18). As we expected, $\mathsf{mmse}_{\text{zero-wait}}$ is larger than $\mathsf{mmse}_{\text{opt}}$ and $\mathsf{mmse}_{\text{age-opt}}$ when $f_{\max} \geq 1$.

Figure 5 illustrates the MMSE of *i.i.d.* log-normal channel delay for $f_{\max} = 1.5$, where $Y_i = e^{\sigma X_i}/\mathbb{E}[e^{\sigma X_i}]$, $\sigma > 0$ is the scale parameter of log-normal distribution, and $(X_1, X_2, \ldots)$ are *i.i.d.* Gaussian random variables with zero mean and unit variance. Because $\mathbb{E}[Y_i] = 1$, the maximum throughput of the channel is 1. Because $f_{\max} > 1$, $\mathsf{mmse}_{\text{uniform}}$ is infinite and hence is not plotted. As the scale parameter σ grows, the tail of the log-normal distribution becomes heavier and heavier. We observe that $\mathsf{mmse}_{\text{zero-wait}}$ grows quickly with respect to σ and is much larger than $\mathsf{mmse}_{\text{opt}}$ and $\mathsf{mmse}_{\text{age-opt}}$. In addition, the gap between $\mathsf{mmse}_{\text{opt}}$ and $\mathsf{mmse}_{\text{age-opt}}$ increases as σ grows.

ACKNOWLEDGEMENT

The authors are grateful to Ness B. Shroff and Roy D. Yates for their careful reading of the paper and valuable suggestions.

REFERENCES

[1] K. J. Astrom and B. M. Bernhardsson, "Comparison of Riemann and Lebesgue sampling for first order stochastic systems," in *IEEE CDC*, 2002.
[2] B. Hajek, K. Mitzel, and S. Yang, "Paging and registration in cellular networks: Jointly optimal policies and an iterative algorithm," *IEEE Trans. Inf. Theory*, vol. 54, no. 2, pp. 608–622, Feb 2008.
[3] O. C. Imer and T. Başar, "Optimal estimation with limited measurements," *International Journal of Systems Control and Communications*, vol. 2, no. 1-3, pp. 5–29, 2010.
[4] G. M. Lipsa and N. C. Martins, "Remote state estimation with communication costs for first-order LTI systems," *IEEE Trans. Auto. Control*, vol. 56, no. 9, pp. 2013–2025, Sept 2011.
[5] A. Nayyar, T. Başar, D. Teneketzis, and V. V. Veeravalli, "Optimal strategies for communication and remote estimation with an energy harvesting sensor," *IEEE Trans. Auto. Control*, vol. 58, no. 9, 2013.
[6] J. Wu, Q. Jia, K. H. Johansson, and L. Shi, "Event-based sensor data scheduling: Tradeoff between communication rate and estimation quality," *IEEE Trans. Auto. Control*, vol. 58, no. 4, 2013.
[7] K. Nar and T. Başar, "Sampling multidimensional Wiener processes," in *IEEE CDC*, Dec. 2014, pp. 3426–3431.
[8] J. Chakravorty and A. Mahajan, "Distortion-transmission trade-off in real-time transmission of Gauss-Markov sources," in *IEEE ISIT*, 2015.
[9] X. Gao, E. Akyol, and T. Başar, "Optimal estimation with limited measurements and noisy communication," in *IEEE CDC*, 2015.
[10] ——, "On remote estimation with multiple communication channels," in *American Control Conference*, 2016, pp. 5425–5430.
[11] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," in *IEEE Infocom*, 2016.
[12] ——, "Update or wait: How to keep your data fresh," submitted to *IEEE Trans. Inf. Theory*, 2016, https://arxiv.org/abs/1601.02284.
[13] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *IEEE ISIT*, 2015.
[14] S. Kaul, R. D. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *IEEE Infocom*, 2012.
[15] C. Kam, S. Kompella, and A. Ephremides, "Age of information under random updates," in *IEEE ISIT*, 2013.
[16] B. T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu, "Age of information under energy replenishment constraints," in *ITA*, 2015.
[17] A. M. Bedewy, Y. Sun, and N. B. Shroff, "Optimizing data freshness, throughput, and delay in multi-server information-update systems," in *IEEE ISIT*, 2016.
[18] ——, "Age-optimal information updates in multihop networks," in *IEEE ISIT*, 2017.
[19] I. Kadota, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Minimizing the age of information in broadcast wireless networks," in *Allerton Conference*, 2016.
[20] B. T. Bacinoglu and E. Uysal-Biyikoglu, "Scheduling status updates to minimize age of information with an energy harvesting sensor," in *IEEE ISIT*, 2017.
[21] R. Mehra, "Optimization of measurement schedules and sensor designs for linear dynamic systems," *IEEE Trans. Auto. Control*, 1976.
[22] P. J. Haas, *Stochastic Petri Nets: Modelling, Stability, Simulation*. New York, NY: Springer New York, 2002.
[23] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the Wiener process over a channel with random delay," Jan. 2017, http://arxiv.org/abs/1701.06734.
[24] A. N. Shiryaev, *Optimal Stopping Rules*. New York: Springer-Verlag, 1978.