Performance Bounds for Sampling and Remote Estimation of Gauss-Markov Processes over a Noisy Channel with Random Delay

Tasmeen Zaman Ornee and Yin Sun
Dept. of ECE, Auburn University, Auburn, AL

***Abstract*—In this study, we generalize a problem of sampling a scalar Gauss Markov Process, namely, the Ornstein-Uhlenbeck (OU) process, where the samples are sent to a remote estimator and the estimator makes a causal estimate of the observed *real-time* signal. In recent years, the problem is solved for stable OU processes. We present solutions for the optimal sampling policy that exhibits a smaller estimation error for both stable and unstable cases of the OU process along with a special case when the OU process turns to a Wiener process. The obtained optimal sampling policy is a threshold policy. However, the thresholds are different for all three cases. Later, we consider additional noise with the sample when the sampling decision is made beforehand. The estimator utilizes noisy samples to make an estimate of the current signal value. The mean-square error (mse) is changed from previous due to noise and the additional term in the mse is solved which provides performance upper bound and room for a pursuing further investigation on this problem to find an optimal sampling strategy that minimizes the estimation error when the observed samples are noisy. Numerical results show performance degradation caused by the additive noise.**

***Index Terms*—Ornstein-Uhlenbeck process, sampling policy, threshold policy, noisy sample.**

I. Introduction

The problem of sampling an Ornstein-Uhlenbeck (OU) process is recently addressed in [1] and another problem of sampling a Wiener process in [2]. However, the optimal sampling policy provided in [1] is only for the stable scenario. In practice, real-time applications of OU processes consider both stable and unstable cases [3]. Therefore, a sampling problem that considers only the stable scenario is insufficient for practical and more dynamical systems, and a generalization of this problem that considers both stable and unstable cases is necessary.

Moreover, a real-time system often consists of noise along with the signal process. Therefore, the analysis based on noisy observation of samples to minimize signal estimation error is practically much more important in real-time networked control and communication systems. In this paper, we generalize a sampling problem of a scalar Gauss-Markov process, named the OU process by considering both stable and unstable scenarios. Later on, we consider noisy samples of OU process and compute the mse from which we establish estimation performance bounds of mse. The optimal sampling policy for noisy samples is not provided in this work but will be considered in our future study.

This work was supported in part by NSF grant CCF-1813050, ONR grant N00014-17-1-2417, and ARO grant W911NF-21-1-0244

The OU process is defined as the solution to the following stochastic differential equation (SDE) [4], [5]

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t, \tag{1}$$

where μ, θ, and $\sigma > 0$ are parameters and W_t represents a Wiener process. In case of stable OU process, $\theta > 0$ [1]. In (1), if $\theta \to 0$, and $\sigma = 1$, X_t reduces to a Wiener process. If $\theta < 0$, then X_t becomes an unstable OU process. Examples and properties of OU processes are explained in [1].

First, we aim to find an optimal sampling strategy that minimizes the mse. The samples of the OU process pass through a channel in first-come, first-serve (FCFS) strategy. A remotely located estimator utilizes these causally received samples to make an estimate $\hat{X}_t$ of X_t. We obtain lower bound of mse in the absence of any additional noise in the system. Second, our goal is to find the expression of mse with the presence of noise in the system. This analysis provides an upper bound of mse when the estimator receives noisy samples. We summarize the contributions of this paper as follows:

- The optimal sampling problem in the absence of noise is formulated and the solved optimal sampling policy is a threshold policy on *instantaneous estimation error*. The structure of the thresholds $v(\beta)$ of a parameter β are different for the three cases: $\theta > 0$ (Stable OU process), $\theta = 0$ (Wiener process), and $\theta < 0$ (Unstable OU process). The value of β is equal to the optimum value of the time-average expected estimation error. The computation of β remains the same irrespective of the signal models.
- Further, we consider noisy samples and obtain an explicit expression for mse. From the expression, we establish a performance upper bound of mse.
- Our results hold for general *i.i.d.* transmission time distributions of the queueing server with a finite mean.

A. Related Work

The results in this paper are tightly connected to the area of remote estimation, e.g., [1], [2], [6]–[15]. Optimal sampling policy of Wiener processes with a zero channel delay was studied in [8], [10], whereas we consider random *i.i.d.* channel delay. A discrete-time optimal stopping problem was solved by using Dynamic programming in [8] to find the optimal sampling policy of OU processes. In [1], an optimal sampler

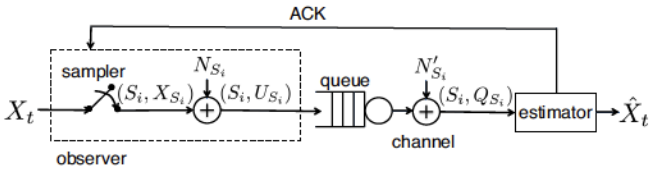

Fig. 1: System model.

of stable OU processes is obtained analytically where the sampling is suspended when the server is busy and is reactivated once the server becomes idle. The optimal sampling policy for Wiener processes in [2] and stable OU processes in [1] is a special case of ours. Remote estimation of Wiener processes with random two-way delay was considered in [14].

In [13], a jointly optimal sampler, quantizer, and estimator design were found for a class of continuous-time Markov processes under a bit-rate constraint. In [15], the quantization and coding schemes on the estimation performance are studied. We consider noisy channels with random delay to establish performance bounds. A recent survey on remote estimation systems was presented in [16].

II. Model and Problem Formulation

A. System Model

We consider a continuous-time remote estimation system that is illustrated in Fig. 1, where an observer takes samples from an OU process X_t. After sampling, additional noises from the sampler and the channel are added to the samples. Then, the noisy samples are sent to the estimator. The channel is modeled as a single-server FIFO queue with *i.i.d.* service times. The samples undergo random service times in the channel due to fading, interference, congestions, etc. We also consider that at a time, only one sample can be delivered through the channel.

The operation of the system starts at time instant $t = 0$. The generation time of the i-th sample is S_i, which satisfy $S_i \leq S_{i+1}$ for all i. Then, i-th sample undergoes a random service time Y_i, and is delivered to the estimator at time D_i, where $S_i + Y_i \leq D_i$, $D_i + Y_{i+1} \leq D_{i+1}$, and $0 < \mathbb{E}[Y_i] < \infty$ hold for all i. The i-th sample packet (S_i, X_{S_i}) contains the sample value X_{S_i} and its sampling time S_i. Suppose that after sampling, noise N_{S_i} is being added to the sample X_{S_i} and the noisy observation of the sample X_{S_i} is denoted by U_{S_i}. Hence,

$$U_{S_i} = X_{S_i} + N_{S_i}, \tag{2}$$

where N_{S_i} is the additive noise with zero mean and variance b_1. Each sample packet (S_i, U_{S_i}) contains the sampling time S_i and the noisy sample U_{S_i}. If channel noise N'_{S_i} with zero mean and variance b_2 is added to the sample during its transmission through the channel, then the sample value becomes

$$Q_{S_i} = U_{S_i} + N'_{S_i}. \tag{3}$$

Initially, at $t = 0$, the state of the system is assumed to hold $S_0 = 0$, and $D_0 = Y_0$. The initial state of the OU process X_0 is a finite constant. The process parameters μ, θ, and σ in (1) are known at both the sampler and estimator.

Let, the idle/busy state of the server at time t is denoted by $I_t \in \{0, 1\}$. We also assume that an acknowledgement is immediately sent back to the sampler whenever a sample is delivered and this operation has zero delay. By this assumption, the sampler is aware of the idle/busy state of the server and the available information at time t can be given by $\{X_s, I_s : 0 \leq s \leq t\}$.

B. Sampling Policies

The sampling time S_i is a finite stopping time with respect to the filtration $\{\mathcal{F}_t^+, t \geq 0\}$ (a non-decreasing and right-continuous family of σ-fields) of the information that is available at the sampler such that [17]

$$\{S_i \leq t\} \in \mathcal{F}_t^+, \forall t \geq 0. \tag{4}$$

Let $\pi = (S_1, S_2, ...)$ denote a sampling policy and Π denote the set of *causal* sampling policies that satisfy two conditions: (i) Each sampling policy $\pi \in \Pi$ satisfies (4) for all i. (ii) The sequence of inter-sampling times $\{T_i = S_{i+1} - S_i, i = 0, 1, \ldots\}$ forms a *regenerative process* [1, Section IIB]: An increasing sequence $0 \leq l_1 < l_2 < \ldots$ of almost surely finite random integers exists such that the post-l_k process $\{T_{l_k+i}, i = 0, 1, \ldots\}$ is independent of the pre-l_k process $\{T_i, i = 0, 1, \ldots, l_k - 1\}$ and has same distribution as the post-l_0 process $\{T_{l_0+i}, i = 0, 1, \ldots\}$; We further assume that $\mathbb{E}[l_{k+1} - l_k] < \infty$, $\mathbb{E}[S_{l_1}] < \infty$, and $0 < \mathbb{E}[S_{l_{k+1}} - S_{l_k}] < \infty$, $k = 1, 2, \ldots$

C. MMSE Estimator

In this section, we provide the MMSE estimator for noisy samples of the OU process.

By using the expression of OU process for stable scenario [18, Eq. (3)] and the strong Markov property of the OU process [19, Eq. (4.3.27)], a solution to (1) for $t \in [S_i, \infty)$ given by the following three cases:

$$X_t = \begin{cases} X_{S_i} e^{-\theta(t-S_i)} + \mu\left[1 - e^{-\theta(t-S_i)}\right] & \\ \quad + \frac{\sigma}{\sqrt{2\theta}} e^{-\theta(t-S_i)} W_{e^{2\theta(t\ \ S_i)}-1}, & \text{if } \theta > 0, \\ \sigma W_t, & \text{if } \theta = 0, \\ X_{S_i} e^{-\theta(t-S_i)} + \mu\left[1 - e^{-\theta(t-S_i)}\right] & \\ \quad + \frac{\sigma}{\sqrt{-2\theta}} e^{-\theta(t-S_i)} W_{1-e^{2\theta(t\ \ S_i)}}, & \text{if } \theta < 0. \end{cases} \tag{5}$$

The estimator uses causally received samples to formulate an estimate $\hat{X}_t$ of the real-time signal value X_t at any time $t \geq 0$. The available information at the estimator has two parts: (i) $M_t = \{(S_i, Q_{S_i}, D_i) : D_i \leq t\}$, which contains the sampling time S_i, noisy sample value Q_{S_i}, and delivery time D_i of the samples that have been delivered by time t and (ii) no sample has been received after the last delivery time $\max\{D_i : D_i \leq t\}$. Similar to [1], [2], [8], [20], we assume that the estimator neglects the second part of information. Then, as shown in [15], the MMSE estimator for $t \in [D_i, D_{i+1})$, $i = 0, 1, 2, \ldots$ for all of the cases in (5) is given as follows

$$\begin{aligned} \hat{X}_t =& \mathbb{E}[X_t | M_t] \\ =& Q_{S_i} e^{-\theta(t-S_i)} + \mu\left[1 - e^{-\theta(t-S_i)}\right]. \end{aligned} \tag{6}$$

D. Performance Metric

We evaluate the performance of remote estimation by the time-average mean square error which is expressed as follows:

$$\mathsf{mse} = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T (X_t - \hat{X}_t)^2 dt\right]. \tag{7}$$

A lower bound of (7) can be obtained when the additive noises are not considered ($N_{S_i} = 0, N'_{S_i} = 0$). On the other hand, an upper bound can be found by taking both the noises into account. Moreover, we formulate the following optimal sampling problem that minimizes the time-average mean-squared estimation error over an infinite time-horizon when no noise is considered.

$$\mathsf{mse}_{\text{opt-wn}} = \min_{\pi\in\Pi} \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T (X_t - \hat{X}_t)^2 dt\right], \tag{8}$$

where $\mathsf{mse}_{\text{opt-wn}}$ is the optimum value of (8) without noise. We do not provide the optimal sampling policy in the presence of noises in this study, but it will be considered in our future work.

III. Main Results

In this section, we first present the lower bounds for mse in (7) for different conditions on the OU process parameter θ. Second, we provide the optimal sampling policy for minimizing the expected estimation error defined in (8). Later, we present upper bound for mse in (7).

A. Lower Bounds for mse

Let us consider an OU process with initial state $O_0 = 0$ and parameter $\mu = 0$, which can be expressed as

$$O_t = \begin{cases} \frac{\sigma}{\sqrt{2\theta}} e^{-\theta t} W_{e^{2\theta t}-1}, & \text{if } \theta > 0, \\ \sigma W_t, & \text{if } \theta = 0, \\ \frac{\sigma}{\sqrt{-2\theta}} e^{-\theta t} W_{1-e^{2\theta t}}, & \text{if } \theta < 0. \end{cases} \tag{9}$$

Before presenting the optimal sampler without noise, let us define the following parameter:

$$\mathsf{mse}_{Y_i} = \begin{cases} \frac{\sigma^2}{2\theta}\mathbb{E}[1 - e^{-2\theta Y_i}], & \text{if } \theta \neq 0, \\ \sigma^2 \mathbb{E}[Y_i], & \text{if } \theta = 0, \end{cases} \tag{10}$$

where mse_{Y_i} is the lower bound of mse. We will also need to use the following two functions

$$G(x) = \frac{\sqrt{\pi}}{2}\frac{e^{x^2}}{x}\,\text{erf}(x),\ x \in [0,\infty), \tag{11}$$

$$K(x) = \frac{\sqrt{\pi}}{2}\frac{e^{-x^2}}{x}\,\text{erfi}(x),\ x \in [0,\infty), \tag{12}$$

where if $x = 0$, both $G(x)$ and $K(x)$ are defined as their right limits $G(0) = \lim_{x\to 0^+} G(x) = 1$, and $K(0) = \lim_{x\to 0^+} K(x) = 1$. Furthermore, erf$(\cdot)$ and erfi$(\cdot)$ are the error function and imaginary error function respectively, defined as

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2} dt, \text{erfi}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{t^2} dt. \tag{13}$$

Note that $G(x)$ is strictly increasing on $x \in [0,\infty)$ [1], whereas $K(x)$ is strictly decreasing on $x \in [0,\infty)$. Hence, their inverses $G^{-1}(\cdot)$ and $K^{-1}(\cdot)$ are properly defined.

First, we consider that the system has no noise, i.e., $N_{S_i} = 0$ and $N'_{S_i} = 0$. Therefore, from (2) and (3), we get, $X_{S_i} = U_{S_i} = Q_{S_i}$. Then, the following theorem illustrates that the optimal sampling policy is a threshold policy and the threshold is found for all the three cases of the OU process parameter θ.

Theorem 1. *If the Y_i's are i.i.d. with $0 < \mathbb{E}[Y_i] < \infty$, then $(S_1(\beta), S_2(\beta), \ldots)$ with a parameter β is an optimal solution to* (8)*, where*

$$S_{i+1}(\beta) = \inf\left\{t \geq D_i(\beta) : \left|X_t - \hat{X}_t\right| \geq v(\beta)\right\}, \tag{14}$$

$D_i(\beta) = S_i(\beta) + Y_i$, and $v(\beta)$ is given by

$$v(\beta) = \begin{cases} \frac{\sigma}{\sqrt{\theta}} G^{-1}\left(\frac{\frac{\sigma^2}{2\theta} - \mathsf{mse}_{Y_i}}{\frac{\sigma^2}{2\theta} - \beta}\right), & \text{if } \theta > 0, \\ \sqrt{3(\beta - \mathbb{E}[Y_i])}, & \text{if } \theta = 0, \\ \frac{\sigma}{\sqrt{-\theta}} K^{-1}\left(\frac{\frac{\sigma^2}{2\theta} - \mathsf{mse}_{Y_i}}{\frac{\sigma^2}{2\theta} - \beta}\right), & \text{if } \theta < 0, \end{cases} \tag{15}$$

where $G^{-1}(\cdot)$ is the inverse function of $G(\cdot)$ in (11)*, $K^{-1}(\cdot)$ is the inverse function of $K(\cdot)$ in* (12)*, and β is the unique root of*

$$\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} (X_t - \hat{X}_t)^2 dt\right] - \beta\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)] = 0. \tag{16}$$

The optimal objective value to (8) *is then given by*

$$\mathsf{mse}_{\text{opt-wn}} = \frac{\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} (X_t - \hat{X}_t)^2 dt\right]}{\mathbb{E}[D_{i+1}(\beta) - D_i(\beta)]}. \tag{17}$$

In [1], it is proved that the optimal sampling policy for stable OU process, i.e., when $\theta > 0$ is a threshold policy. The threshold obtained in [1] coincides with $v(\beta)$ in (15) for the case of $\theta > 0$. For $\theta = 0$, the threshold is obtained for $\sigma = 1$ which represents a Wiener process [2]. For $\theta < 0$, the proof procedure works in the same way as explained in [1] for stable OU processes. The threshold $v(\beta)$ is obtained by solving similar free boundary problems explained in [1] and the optimality of (17) for $\theta < 0$ is thus guaranteed. However, the threshold structure is different for all the three cases in Theorem 1. The function $K(x)$ in (12) is related to the function $G(x)$ in (11) as follows

$$K(x) = G(jx), \tag{18}$$

where j is the imaginary number represented by $j = \sqrt{-1}$. Therefore, the threshold $v(\beta)$ for $\theta < 0$ can be expressed by the following equation as well:

$$v(\beta) = j^{-1}\frac{\sigma}{\sqrt{-\theta}} G^{-1}\left(\frac{\frac{\sigma^2}{2\theta} - \mathsf{mse}_{Y_i}}{\frac{\sigma^2}{2\theta} - \beta}\right). \tag{19}$$

Though the threshold functions $v(\beta)$ varies with signal structure, the computation of the parameter β remains the same for all cases and the uniqueness of the root of (16) is proved in [1]. The decision of taking a new sample defined in (14) works in the same way as explained in [1].

B. Upper Bounds for mse

Suppose that the additive noise in the sampler and channel exist in the system, i.e., $N_{S_i} \neq 0, N'_{S_i} \neq 0$. Moreover, the sampler follows the sampling strategy obtained in (14). The OU process O_t is a Gauss-Markov process. When noises get incorporated with O_t, it does not remain Markov. The analysis presented in our previous study [1] was based on the strong Markov property of the OU processes. Due to the non-Markovian structure of noisy samples of OU processes, finding an optimal sampling policy requires different analytical tools. Due to lack of space, we do not provide the optimal sampling policy with the presence of noise, but it will be considered in our future study.

Because the noises N_{S_i} and N'_{S_i} are independent of the sampling times and the observed OU process, by utilizing (2), (3), (5), (6), and (17), the mse at the estimator which is an upper bound of (7) can be expressed as

$$\begin{aligned}\mathsf{mse} =& \frac{\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}(X_t-\hat{X}_t)^2dt\right]}{\mathbb{E}[D_{i+1}(\beta)-D_i(\beta)]}\\ =& \frac{\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}(O_{t-S_i}-(N_{S_i}+N'_{S_i})e^{-\theta(t-S_i)})^2dt\right]}{\mathbb{E}[D_{i+1}(\beta)-D_i(\beta)]}\\ =& \frac{\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}O^2_{t-S_i}dt\right]}{\mathbb{E}[D_{i+1}(\beta)-D_i(\beta)]}\\ &+\frac{\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}(N_{S_i}+N'_{S_i})^2e^{-2\theta(t-S_i)}dt\right]}{\mathbb{E}[D_{i+1}(\beta)-D_i(\beta)]}, \end{aligned} \tag{20}$$

where (20) follows due to the fact that the OU process O_t has initial state $O_0 = 0$ and the noises N_{S_i} and N'_{S_i} with zero mean are independent of the observed OU process and sampling times.

To compute (20), the first fractional term remains the same as the $\mathsf{mse}_{\text{opt-wn}}$ in (17) with $N_{S_i} = 0$ and $N'_{S_i} = 0$. For stable OU processes, the associated $\mathsf{mse}_{\text{opt-wn}}$ is computed in [1, Lemma 1]. The expression of O^2_{t-Si} is the same for both stable and unstable OU processes. Therefore, the solution for (20) holds for all three cases in (9). For computing the second term, as N_{S_i} and N'_{S_i} are independent of the observed OU process and the sampling times, the numerator of the second fractional term in (20) can be written as:

$$\begin{aligned}&\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}(N_{S_i}+N'_{S_i})^2e^{-2\theta(t-S_i)}dt\right]\\ =&\mathbb{E}[(N_{S_i}+N'_{S_i})^2]\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}e^{-2\theta(t-S_i)}dt\right].\end{aligned} \tag{21}$$

Then, we have the following lemma for the last term in (21).

Lemma 1. *It holds that*

$$\begin{aligned}&\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)}e^{-2\theta(t-S_i)}dt\right]\\ =&\frac{1}{2\theta}\mathbb{E}\left[e^{-2\theta Y_i}\left\{1-\min\left(1,\frac{{}_1F_1\left(1,\frac{1}{2},\frac{\theta}{\sigma^2}O^2_{Y_i}\right)}{{}_1F_1\left(1,\frac{1}{2},\frac{\theta}{\sigma^2}v^2(\beta)\right)}\right)\right\}\right]\mathbb{E}[e^{-2\theta Y_{i+1}}].\end{aligned} \tag{22}$$

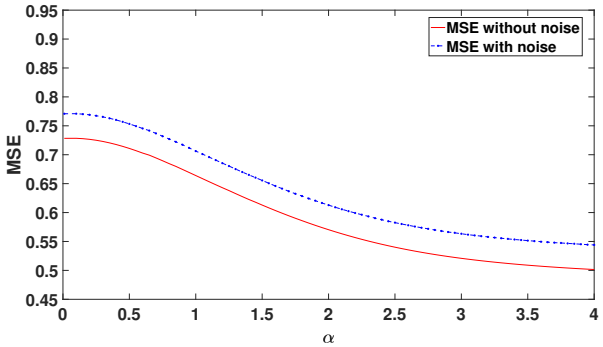

Fig. 2: MSE vs. the scale parameter α of *i.i.d.* normalized log-normal service time distribution with $\mathbb{E}[Y_i] = 1$, where the parameters of the OU process are $\sigma = 1$ and $\theta = 0.5$.

Proof. See Appendix A. □

By using Lemma 1 and the expressions obtained in [1, Lemma 1], all the associated expectations in (20) can be obtained by Monte Carlo simulations of scalar random variables O_{Y_i} and Y_i, which does not require to directly simulate the entire random process $\{O_t, t \geq 0\}$.

IV. Numerical Results

Figure 2 illustrates the MSE of *i.i.d* normalized log-normal service time, where $Y_i = e^{\alpha X_i}/\mathbb{E}[e^{\alpha X_i}]$, and $\alpha > 0$ is the scale parameter of log-normal distribution. The $(X_1, X_2, \dots)$ are *i.i.d.* Gaussian random variables, where $\mathbb{E}[X_i] = 0$ and $\text{Var}(X_i) = 1$. The maximum throughput of the queue is 1 as $\mathbb{E}[Y_i] = 1$. Both of the noises N_{S_i} and N'_{S_i} are considered to have 0 mean and variance 0.1. With the growth of the scale parameter α, the tail of the log-normal distribution becomes heavier. The MSE with noise curve shows performance degradation as the additional term due to noise added with the mse without noise.

V. Conclusion

In this paper, we have explained the optimal sampling strategies for minimizing the *instantaneous estimation error* for three different cases of scalar Gauss-Markov signal processes. The optimal sampler exhibits a threshold policy and by using causal knowledge of the signal values, a smaller estimation error has been obtained. The optimal threshold has been changed with signal structure. For noisy samples, the additional term added in the mse due to noise is found. An optimal sampler design for noisy samples of Gauss-Markov processes will be considered in our future study.

References

[1] T. Z. Ornee and Y. Sun, "Sampling and remote estimation for the Ornstein-Uhlenbeck processes through queues: Age of information and beyond," 2020, accepted by *IEEE Transactions on Networking*.

[2] Y. Sun, Y. Polyanskiy, and E. Uysal, "Sampling of the wiener process for remote estimation over a channel with random delay," *IEEE Trans. Inf. Theory*, vol. 66, no. 2, pp. 1118–1135, 2020.

[3] O. E. Barndorff-Nielsen and N. Shephard, *Modelling by Levy processes for financial econometrics*. Birkhauser, 2001, pp. 283–318.

[4] G. E. Uhlenbeck and L. S. Ornstein, "On the theory of the Brownian motion," *Phys. Rev.*, vol. 36, pp. 823–841, Sept. 1930.

[5] J. L. Doob, "The Brownian movement and stochastic equations," *Annals of Mathematics*, vol. 43, no. 2, pp. 351–369, 1942.

[6] 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.
[7] 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.
[8] M. Rabi, G. V. Moustakides, and J. S. Baras, "Adaptive sampling for linear state estimation," *SIAM Journal on Control and Optimization*, vol. 50, no. 2, pp. 672–702, 2012.
[9] 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, pp. 2246–2260, Sept. 2013.
[10] K. Nar and T. Başar, "Sampling multidimensional Wiener processes," in *IEEE CDC*, Dec. 2014, pp. 3426–3431.
[11] X. Gao, E. Akyol, and T. Başar, "Optimal communication scheduling and remote estimation over an additive noise channel," *Automatica*, vol. 88, pp. 57 – 69, 2018.
[12] J. Chakravorty and A. Mahajan, "Remote estimation over a packet-drop channel with Markovian state," *IEEE Trans. Auto. Control*, vol. 65, no. 5, pp. 2016–2031, 2020.
[13] N. Guo and V. Kostina, "Optimal causal rate-constrained sampling for a class of continuous Markov processes," in *IEEE ISIT*, 2020.
[14] C.-H. Tsai and C.-C. Wang, "Unifying AoI minimization and remote estimation: Optimal sensor/controller coordination with random two-way delay," in *IEEE INFOCOM*, 2020.
[15] A. Arafa, K. Banawan, K. G. Seddik, and H. Vincent Poor, "Timely estimation using coded quantized samples," in *IEEE ISIT*, 2020, pp. 1812–1817.
[16] V. Jog, R. J. La, and N. C. Martins, "Channels, learning, queueing and remote estimation systems with a utilization-dependent component," 2019, coRR, abs/1905.04362.
[17] R. Durrett, *Probability: Theory and Examples*, 4th ed. Cambridge Univerisity Press, 2010.
[18] R. A. Maller, G. Müller, and A. Szimayer, "Ornstein-Uhlenbeck processes and extensions," in *Handbook of Financial Time Series*, T. Mikosch, J.-P. Kreiß, R. A. Davis, and T. G. Andersen, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 421–437.
[19] G. Peskir and A. N. Shiryaev, *Optimal Stopping and Free-Boundary Problems*. Basel, Switzerland: Birkhäuswer Verlag, 2006.
[20] T. Soleymani, S. Hirche, and J. S. Baras, "Optimal information control in cyber-physical systems," *IFAC-PapersOnLine*, vol. 49, no. 22, pp. 1 – 6, 2016.
[21] K. R. Ghusinga, V. Srivastava, and A. Singh, "Driving an ornstein-uhlenbeck process to desired first-passage time statistics," in *18th European Control Conference (ECC)*, 2019, pp. 869–874.

APPENDIX A
PROOF OF LEMMA 1

In order to prove Lemma 1, we need to consider the following two cases:

Case 1: If $|X_{D_i(\beta)} - \hat{X}_{D_i(\beta)}| = |O_{Y_i}| \geq v(\beta)$, then $S_{i+1}(\beta) = D_i(\beta)$. Hence,

$$D_i(\beta) = S_i(\beta) + Y_i, \tag{23}$$
$$D_{i+1}(\beta) = S_{i+1}(\beta) + Y_{i+1} = D_i(\beta) + Y_{i+1}. \tag{24}$$

Let us consider the following equation:

$$\begin{aligned}
&\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} e^{-2\theta(t-S_i)}dt\Big|O_{Y_i}=q, Y_i=y, |O_{Y_i}|\geq v(\beta)\right]\\
=&\mathbb{E}\left[\int_{Y_i}^{Y_i+Y_{i+1}} e^{-2\theta s}ds\Big|O_{Y_i}=q, Y_i=y, |O_{Y_i}|\geq v(\beta)\right]\\
=&\mathbb{E}\left[\frac{1}{2\theta}e^{-2\theta y}(1-e^{-2\theta Y_{i+1}})\Big|O_{Y_i}=q, Y_i=y, |O_{Y_i}|\geq v(\beta)\right]\\
=&\frac{1}{2\theta}e^{-2\theta y}\mathbb{E}\left[1-e^{-2\theta Y_{i+1}}\Big|O_{Y_i}=q, Y_i=y, |O_{Y_i}|\geq v(\beta)\right]\\
=&\frac{1}{2\theta}e^{-2\theta y}\Big\{1-\mathbb{E}[e^{-2\theta Y_{i+1}}]\Big\},
\end{aligned} \tag{25}$$

where (25) holds due to the fact that Y_{i+1} is independent of O_{Y_i} and Y_i.

Case 2: If $|X_{D_i(\beta)} - \hat{X}_{D_i(\beta)}| = |O_{Y_i}| < v(\beta)$, then

$$\begin{aligned}
&\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} e^{-2\theta(t-S_i)}dt\Big|O_{Y_i}=q, Y_i=y, |O_{Y_i}|< v(\beta)\right]\\
=&\mathbb{E}\left[\int_{Y_i}^{Y_i+Z_i+Y_{i+1}} e^{-2\theta s}ds\Big|O_{Y_i}=q, Y_i=y, |O_{Y_i}|< v(\beta)\right]\\
=&\mathbb{E}\left[\frac{1}{2\theta}e^{-2\theta y}(1-e^{-2\theta Z_i}e^{-2\theta Y_{i+1}})\Big|O_{Y_i}=q, Y_i=y\right]\\
=&\frac{1}{2\theta}e^{-2\theta y}\mathbb{E}\left[1-e^{-2\theta Z_i}e^{-2\theta Y_{i+1}}\Big|O_{Y_i}=q, Y_i=y\right]\\
=&\frac{1}{2\theta}e^{-2\theta y}\left\{1-\mathbb{E}\left[e^{-2\theta Z_i}\Big|O_{Y_i}=q, Y_i=y\right]\mathbb{E}[e^{-2\theta Y_{i+1}}]\right\}\\
=&\frac{1}{2\theta}e^{-2\theta y}\left\{1-\mathbb{E}[e^{-2\theta Z_i}|O_{Y_i}=q]\mathbb{E}[e^{-2\theta Y_{i+1}}]\right\},
\end{aligned} \tag{26}$$

where the last equation in (26) holds because Z_i is conditionally independent of Y_i given O_{Y_i}. Next, we need to compute $\mathbb{E}[e^{-2\theta Z_i}|O_{Y_i}=q]$, where Z_i is a hitting time of the time-shifted OU process O_{t+Y_i} given as

$$\begin{aligned}
&Z_i =\\
&\inf\{t : O_{t+Y_i}\notin(-v(\beta), v(\beta))|O_{Y_i}=q\in(-v(\beta), v(\beta))\}.
\end{aligned} \tag{27}$$

By using the characteristic function of the hitting time of the OU process in [21, Eq. 15a], we get that

$$\mathbb{E}[e^{-2\theta Z_i}|O_{Y_i}=q] = \frac{{}_1F_1\big(1,\frac{1}{2},\frac{\theta}{\sigma^2}q^2\big)}{{}_1F_1\big(1,\frac{1}{2},\frac{\theta}{\sigma^2}v^2(\beta)\big)}. \tag{28}$$

Therefore, (26) becomes

$$\begin{aligned}
&\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} e^{-2\theta(t-S_i)}dt\Big|O_{Y_i}=q, Y_i=y, |O_{Y_i}|< v(\beta)\right]\\
=&\frac{1}{2\theta}e^{-2\theta y}\left\{1-\frac{{}_1F_1\big(1,\frac{1}{2},\frac{\theta}{\sigma^2}q^2\big)}{{}_1F_1\big(1,\frac{1}{2},\frac{\theta}{\sigma^2}v^2(\beta)\big)}\mathbb{E}[e^{-2\theta Y_{i+1}}]\right\}.
\end{aligned} \tag{29}$$

By combining (25) and (29), we get that

$$\begin{aligned}
&\mathbb{E}\left[\int_{D_i(\beta)}^{D_{i+1}(\beta)} e^{-2\theta(t-S_i)}dt\Big|O_{Y_i}=q, Y_i=y\right]\\
=&\frac{1}{2\theta}e^{-2\theta y}\left[1-\min\left\{1,\frac{{}_1F_1\big(1,\frac{1}{2},\frac{\theta}{\sigma^2}q^2\big)}{{}_1F_1\big(1,\frac{1}{2},\frac{\theta}{\sigma^2}v^2(\beta)\big)}\right\}\mathbb{E}[e^{-2\theta Y_{i+1}}]\right].
\end{aligned} \tag{30}$$

Finally, by taking the expectation over O_{Y_i} and Y_i in (30), Lemma 1 is proven.