Efficient Sampling for Keeping Track of an Ornstein-Uhlenbeck Process

Maben Rabi, John S. Baras, and George V. Moustakides

***Abstract*— We consider estimation and tracking problems in sensor networks with constraints in the hierarchy of inference making, on the sharing of data and inter-sensor communications. We identify as a typical representative for such problems tracking of a process when the number and type of measurements are constrained. As the simplest representative of such problems, which still encompasses all the key issues involved, we analyze efficient sampling schemes for *tracking* an Ornstein-Uhlenbeck process. We consider sampling based on time, based on amplitude (event-triggered sampling) and optimal sampling (optimal stopping). We obtain the best sampling rule in each case as the solution to a constrained optimization problem. We compare the performances of the various sampling schemes and show that the event-triggered sampling performs close to optimal. Implications and extensions are discussed.**

I. Motivation

Estimation and tracking by sensor networks, employing various hierarchies of decision making and data sharing under communication constraints have recently attracted increasing interest [1], [4], [5], [2], [3], [7]. These constraints are important in systems where it is significantly cheaper to gather measurements than to transmit processed information within the system. As examples, we are encouraged to pay attention to Sensor networks, Robotic systems, etc. A generic system can be represented as a set of nodes which exchange information using limited resources. In many sensor networks, each packet transmitted by a node drains precious battery power. In robotic systems and automobile monitoring systems, nodes share a channel for all communication and so, there are limits on the information transmission rates. In these systems, depending on the task at hand, the nodes have to use their available communication budget wisely. This directly translates to the problem where the inference must be made subject to constraints on the number and type of measurements made. The simplest such problem arises when the number and type of samples of the processes to be estimated are constrained. This is the main problem analyzed in this paper.

In our setting, a *Sensor* makes continuous observations of a Gaussian signal process. It transmits at times it chooses, samples of its observations to a *Supervisor* which uses this stream of samples to maintain a filtered (real-time) estimate of the signal. We study the tracking performance of an efficient sampling scheme which is event-triggered. Extensions of such problems arise in sensor networks because of the limited capacity of a remote sensor node to communicate to the supervisor. For simplicity of exposition, we consider here the signal to be the scalar Ornstein-Uhlenbeck process.

This research was supported by the United States Army Research Office under the ODDR&E MURI01 Program Grant No. DAAD19-01-1-0465 to the center for Networked Communication and Control systems (through Boston University), and under ARO grant No. DAAD1920210319.

Maben Rabi and John S. Baras are at the Institute for Systems Research and the Department of Electrical and Computer Engineering, University of Maryland, College PArk, MD 20742. USA.

George V. Moustakides is at the Department of Computer Enginering and Informatics, University of Patras, Patras, Greece.

II. Preliminaries

We are interested in estimating an Ornstein-Uhlenbeck process x_t inside a prescribed interval $[0,T]$. If $\hat{x}_t$ the estimate, we measure the quality of the estimate by the following average integral squared error

$$\mathcal{J} = \mathbb{E}\left[\int_0^T (x_t - \hat{x}_t)^2\, dt.\right]$$

For the process x_t we have the following sde

$$dx_t = -ax_t dt + dw_t$$

where w_t is a standard Wiener process and x_0 is a *zero mean* random variable with pdf $f(x_0)$. Positive values of a give rise to a stable process, negative to an unstable and finally $a = 0$ to the Wiener process.

The estimate $\hat{x}_t$ relies on knowledge about x_t acquired during the time interval $[0,T]$. The type of information we are interested in, are *samples* obtained by sampling x_t at k time instances $0 \leq \tau_1 \leq \tau_2 \leq \cdots \leq \tau_k \leq T$. If we use the minimum mean square error estimate given by

$$\hat{x}_t = \begin{cases} 0 & t \in [0, \tau_1) \\ x_{\tau_n} e^{-a(t-\tau_n)} & t \in [\tau_n, \tau_{n+1}), \end{cases}$$

the performance measure becomes

$$\mathcal{J}(\tau_1, \ldots, \tau_k) = \mathbb{E}\Bigg[\int_0^{\tau_1} x_t^2\, dt + \sum_{n=1}^{k-1} \int_{\tau_n}^{\tau_{n+1}} (x_t - \hat{x}_t)^2\, dt + \int_{\tau_k}^{T} (x_t - \hat{x}_t)^2\, dt\Bigg]. \quad (1)$$

The goal here is to find sampling policies that are optimal in the sense that they solve the following optimization problem:

$$\inf_{\tau_1,\ldots,\tau_k} \mathcal{J}(\tau_1, \ldots, \tau_k).$$

The sampling times are allowed to be *random* but they must be *stopping times*. The reason is that we would like our decision whether to sample or not at time t to rely only on the observed process up to time t.

For the remainder of this paper, and in order to clarify the concepts and computations involved, we treat the single sample case. The multiple sample case just described will be treated elsewhere.

III. The single sample case

Let us limit ourselves to the single sample case where, for simplicity, we drop the subscript from the unique sampling instance τ_1. In this special case the performance measure in (1) takes the form

$$\begin{aligned}\mathcal{J}(\tau) &= \mathbb{E}\left[\int_0^\tau x_t^2 + \int_\tau^T (x_t-\hat{x}_t)^2\,dt\right]\\ &= \mathbb{E}\left[\int_0^T x_t^2 - 2\int_\tau^T x_t\hat{x}_t\,dt + \int_\tau^T (\hat{x}_t)^2\,dt\right].\end{aligned}$$

Now notice that the second term can be written as follows

$$\begin{aligned}\mathbb{E}\left[\int_\tau^T x_t\hat{x}_t\,dt\right] &= \mathbb{E}\left[\int_\tau^T \mathbb{E}[x_t|\mathcal{F}_\tau]\hat{x}_t\,dt\right]\\ &= \mathbb{E}\left[\int_\tau^T (\hat{x}_t)^2\,dt\right],\end{aligned}$$

where we have used the strong Markov property of x_t that for $t>\tau$ we have $\mathbb{E}[x_t|\mathcal{F}_\tau] = x_\tau e^{-a(t-\tau)} = \hat{x}_t$. Because of this observation the performance measure $\mathcal{J}(\tau)$ takes the form

$$\begin{aligned}\mathcal{J}(\tau) &= \mathbb{E}\left[\int_0^T x_t^2\,dt - \int_\tau^T (\hat{x}_t)^2\,dt\right]\\ &= \frac{e^{-2aT}-1+2aT}{4a^2}\\ &\quad + \mathbb{E}\left[x_0^2\frac{1-e^{-2aT}}{2a} - x_\tau^2\frac{1-e^{-2a(T-\tau)}}{2a}\right]\\ &= T^2\left\{\frac{e^{-2aT}-1+2aT}{4(aT)^2}\right.\\ &\quad \left. + \mathbb{E}\left[\frac{x_0^2}{T}\frac{1-e^{-2aT}}{2(aT)} - \frac{x_\tau^2}{T}\frac{1-e^{-2(aT)(1-\tau/T)}}{2(aT)}\right]\right\}\\ &= T^2\left\{\frac{e^{-2\bar{a}}-1+2\bar{a}}{4\bar{a}^2}\right.\\ &\quad \left. + \mathbb{E}\left[\bar{x}_0^2\frac{1-e^{-2\bar{a}}}{2\bar{a}} - \bar{x}_{\bar{\tau}}^2\frac{1-e^{-2\bar{a}(1-\bar{\tau})}}{2\bar{a}}\right]\right\},\end{aligned}$$

where

$$\bar{t} = \frac{t}{T};\quad \bar{a} = aT;\quad \bar{x}_{\bar{t}} = \frac{x_{\frac{t}{T}}}{\sqrt{T}}. \tag{2}$$

It is interesting to note that

$$d\bar{x}_{\bar{t}} = -\bar{a}\bar{x}_{\bar{t}}d\bar{t} + dw_{\bar{t}}.$$

This suggests that, without loss of generality, we can limit ourselves to the normalized case $T=1$ since the case $T\neq 1$ can be reduced to it by using the transformations in (2). The performance measure we are finally considering is

$$\begin{aligned}\mathcal{J}(\tau) &= \frac{e^{-2a}-1+2a}{4a^2}\\ &\quad + \mathbb{E}\left[x_0^2\frac{1-e^{-2a}}{2a} - x_\tau^2\frac{1-e^{-2a(1-\tau)}}{2a}\right];\quad \tau\in[0,1].\end{aligned} \tag{3}$$

We will also need the following expression

$$\begin{aligned}\mathcal{J}(\tau,x_0) &= \frac{e^{-2a}-1+2a}{4a^2} + x_0^2\frac{1-e^{-2a}}{2a}\\ &\quad - \mathbb{E}\left[x_\tau^2\frac{1-e^{-2a(1-\tau)}}{2a}\Big|x_0\right];\quad \tau\in[0,1].\end{aligned} \tag{4}$$

Clearly, $\mathcal{J}(\tau) = \mathbb{E}[\mathcal{J}(\tau,x_0)]$, where the last expectation is with respect to the statistics of the initial condition x_0.

Next we are going to consider three different classes of admissible sampling strategies and we will attempt to find the optimal within each class that minimizes the performance measure in (3). The classes we are interested in are: a) deterministic sampling; b) threshold sampling and c) optimal sampling.

A. Optimal deterministic sampling

Let us first minimize (3) over the class of deterministic sampling times. The performance measure then takes the form

$$\begin{aligned}\mathcal{J}(\tau) &= \frac{e^{-2a}-1+2a}{4a^2} + \sigma^2\frac{1-e^{-2a}}{2a}\\ &\quad -\frac{1}{4a^2}\left\{1-(1-2a\sigma^2)e^{-2a\tau}\right\}\left\{1-e^{-2a}e^{2a\tau}\right\};\ \tau\in[0,1]\end{aligned} \tag{5}$$

where σ^2 denotes the variance of the initial condition. Clearly $\mathcal{J}(\tau)$ is minimized when we maximize the last term in the previous expression. It is a simple exercise to verify that the optimal sampling time satisfies

$$\begin{aligned}\tau_o &= \arg\max_\tau \left\{1-(1-2a\sigma^2)e^{-2a\tau}\right\}\left\{1-e^{-2a}e^{2a\tau}\right\}\\ &= \begin{cases}\frac{1}{2} + \frac{\log(1-2a\sigma^2)}{4a} & \text{for } \sigma^2 \leq \frac{1-e^{-2a}}{2a},\\ 0 & \text{otherwise.}\end{cases}\end{aligned} \tag{6}$$

In other words, if the initial variance is greater than the value $(1-e^{-2a})/2a$ then it is better to sample in the beginning. The corresponding optimum performance becomes

$$\begin{aligned}\mathcal{J}(\tau_o) &= \frac{e^{-2a}-1+2a}{4a^2}\\ &\quad - \frac{1}{4a^2}(e^{-a}-\sqrt{1-2a\sigma^2})^2\mathbb{1}_{\left\{\sigma^2\leq\frac{1-e^{-2a}}{2a}\right\}}.\end{aligned} \tag{7}$$

B. Optimal threshold sampling

Here we consider a threshold η and we sample the process x_t whenever $|x_t|$ exceeds η for the first time. If we call τ_η the sampling instance

$$\tau_\eta = \inf_{0\leq t}\{t : |x_t| \geq \eta\}.$$

then it is clear that we can have $\tau_\eta > 1$. We therefore define our sampling time as the minimum of the two, that is, $\tau = \min\{\tau_\eta, 1\}$. Of course sampling at time $\tau = 1$, has absolutely no importance since from (3) we can see that such a sampling produces no contribution in the performance measure. Another important detail in threshold sampling is the fact that whenever $|x_0| \geq \eta$ then we sample at the beginning.

Our goal here is, for given parameter a and pdf $f(x_0)$ to find the threshold η that will minimize the performance measure $\mathcal{J}(\tau)$. As in the previous case let us analyze $\mathcal{J}(\tau)$. We first need to compute $\mathcal{J}(\tau, x_0)$ for given threshold η. From (4) we have

$$\mathcal{J}(\tau,x_0) = \frac{e^{-2a}-1+2a}{4a^2} + \left\{x_0^2\frac{1-e^{-2a}}{2a} - \eta^2\mathbb{E}\left[\frac{1-e^{-2a(1-\tau)}}{2a}\Big|x_0\right]\right\}\mathbb{1}_{\{|x_0|<\eta\}}. \tag{8}$$

We first note that our expression captures the fact that we sample in the beginning whenever $|x_0| \geq \eta$. Whenever this does not happen, that is, on the event $\{|x_0| < \eta\}$ we apply our threshold sampling. If $|x_t|$ reaches the threshold η before the limit time 1, then we sample and $x_\tau = \pm\eta$, therefore

$$x_\tau^2\frac{1-e^{-2a(1-\tau)}}{2a} = \eta^2\frac{1-e^{-2a(1-\tau)}}{2a}.$$

If however $|x_t|$ does not reach the threshold before time 1, then we sample at $t = 1$ and we have

$$x_\tau^2\frac{1-e^{-2a(1-\tau)}}{2a}\Big|_{\tau=1} = 0 = \eta^2\frac{1-e^{-2a(1-\tau)}}{2a}\Big|_{\tau=1},$$

Manipulating the last term in (8) we obtain

$$\begin{aligned}\mathcal{J}(\tau,x_0) = &\frac{e^{-2a}-1+2a}{4a^2}\\ &-(\eta^2-x_0^2)\left[\frac{1-e^{-2a}}{2a}\right]\mathbb{1}_{\{|x_0|<\eta\}}\\ &+\eta^2e^{-2a}\mathbb{E}\left[\int_0^\tau e^{2at}\,dt\Big|x_0\right]\mathbb{1}_{\{|x_0|<\eta\}}.\end{aligned}$$

The only term that needs special attention in the previous formula is the last one, for which we must find a computational recipe. Consider a function $U(x,t)$ defined on the orthogonal region $|x| \leq \eta,\ 0 \leq t \leq 1$. We require $U(x,t)$ to satisfy the following pde and boundary conditions

$$\frac{1}{2}U_{xx} - axU_x + U_t + e^{2at} = 0;\quad U(\pm\eta,t) = U(x,1) = 0. \tag{9}$$

If we apply standard Itô calculus on $U(x_t,t)$ we have

$$\begin{aligned}\mathbb{E}[U(x_\tau,\tau)|x_0] - U(x_0,0) &= \mathbb{E}\left[\int_0^\tau dU(x_t,t)\Big|x_0\right]\\ &= \mathbb{E}\left[\int_0^\tau\left\{\frac{1}{2}U_{xx} - axU_x + U_t\right\}dt\Big|x_0\right]\\ &= -\mathbb{E}\left[\int_0^\tau e^{2at}dt\Big|x_0\right].\end{aligned}$$

Notice that at the time of sampling, x_τ is either at the boundary $x_\tau = \pm\eta$ in which case $U(x_\tau,\tau) = U(\pm\eta,\tau) = 0$, or we have reached the limit $t = 1$ with $|x_1| < \eta$, thus we sample at $\tau = 1$ which yields $U(x_\tau,\tau) = U(x_1,1) = 0$. We thus conclude that $\mathbb{E}[\int_0^\tau e^{2at}dt|x_0] = U(x_0,0)$.

With the help of the function $U(x_0,0)$ we can write $\mathcal{J}(\tau,x_0)$ as

$$\begin{aligned}\mathcal{J}(\tau,x_0) = &\frac{e^{-2a}-1+2a}{4a^2}\\ &-\left\{(\eta^2-x_0^2)\frac{1-e^{-2a}}{2a} + \eta^2e^{-2a}U(x_0,0)\right\}\mathbb{1}_{\{|x_0|<\eta\}}.\end{aligned}$$

Averaging this over x_0 yields the following performance measure

$$\begin{aligned}\mathcal{J}(\tau) = &\frac{e^{-2a}-1+2a}{4a^2}\\ &-\frac{1-e^{-2a}}{2a}\mathbb{E}\left[(\eta^2-x_0^2)\mathbb{1}_{\{|x_0|<\eta\}}\right]\\ &-\eta^2e^{-2a}\mathbb{E}\left[U(x_0,0)\mathbb{1}_{\{|x_0|<\eta\}}\right].\end{aligned}$$

To find the optimal threshold and the corresponding optimum performance we need to minimize $\mathcal{J}$ over η. This optimization can be performed numerically as follows: for every η we compute $U(x_0,0)$ by solving the pde in (9); then we perform the averaging over x_0; we then compute the performance measure for different values of η and select the one that yields the minimum $\mathcal{J}(\tau)$.

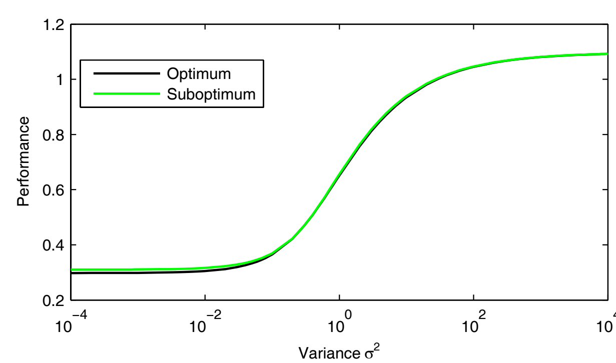

Fig. 1. Relative performance of optimal (variable) threshold and suboptimal constant threshold sampling scheme, as a function of the variance (σ^2) of the initial condition.

In order to observe certain key properties of the optimal thresholding scheme let us consider the case $a = 1$ with a zero mean Gaussian initial value x_0 of variance σ^2. Fig. 1(a) depicts the optimum performance $\mathcal{J}(\tau)$ as a function of the variance σ^2 and Fig. 2(b) the corresponding optimal threshold η. From Fig. 2(b) we observe that the optimal threshold is between two limiting values η_0, η_∞. The interesting point is that both these values are *independent* of the actual density function $f(x_0)$, as long as the pdf is from a *unimodal* family of the form: $f(x) = h(x/\sigma)/\sigma,\ \sigma \geq 0$ where, $h(\cdot)$ is an unimodal pdf with unit variance and with both its mean and mode being zero. Indeed for such a pdf, variance tending to 0, means that the density $f(x_0)$ tends to a Dirac delta

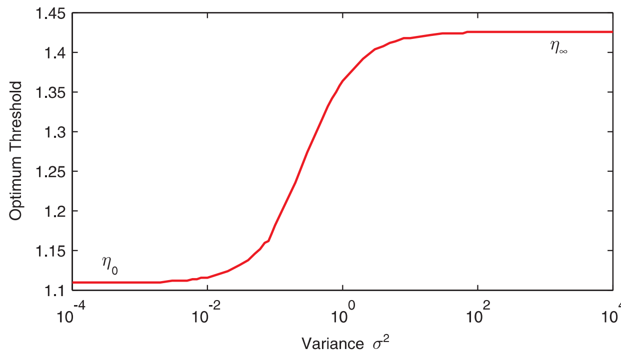

Fig. 2. Optimal threshold as a function of the initial variance σ^2, with $a = 1$.

function at zero. The performance measure in (III-B) then takes the simple form

$$\mathcal{J}(\tau) = \frac{e^{-2a} - 1 + 2a}{4a^2} - \eta^2 \left\{ \frac{1 - e^{-2a}}{2a} + e^{-2a} U(0,0) \right\}$$

which, if minimized with respect to η, yields η_0. If now we let the variance $\sigma^2 \to \infty$ then every unimodal function becomes almost flat with value $f(0)$ inside each finite interval $[-\eta, \eta]$. The corresponding performance measure then takes the form

$$\begin{aligned}\mathcal{J}(\tau) \approx & \frac{e^{-2a} - 1 + 2a}{4a^2} \\ & - f(0) \int_{-\eta}^{\eta} \frac{1 - e^{-2a}}{2a} (\eta^2 - x_0^2)\, dx_0 \\ & - f(0) \int_{-\eta}^{\eta} \eta^2 e^{-2a} U(x_0, 0)\, dx_0 .\end{aligned}$$

To optimize the previous expression it is sufficient to optimize the last integral, which is independent of the actual pdf $f(x_0)$. This optimization will yield η_∞.

Threshold sampling has another interesting property. If instead of using the optimal threshold η which is a function of the initial pdf and the variance σ^2, we use the *constant* threshold $\eta_o = 0.5(\eta_0 + \eta_\infty)$, then the resulting sampling policy is clearly suboptimal. However as we can see from Fig. 1 the performance of the suboptimal scheme is practically indistinguishable from that of the optimal. Having a sampling scheme which is (nearly) optimal for a large variety of pdfs (unimodal functions) and practically any variance value, is definitely a very desirable characteristic. We would like to stress that this property breaks when $f(x_0)$ is not unimodal and also when a takes upon large negative values (i.e. the process is strongly unstable).

C. Optimal sampling

In this section we are interested in sampling strategies that are optimal in the sense that they minimize the performance measure (3) among *all* possible sampling policies (stopping times) τ. Unlike the previous sampling scheme, the optimal sampling rule is completely *independent* of the pdf $f(x_0)$. From (3) it is clear that in order to minimize the cost $\mathcal{J}(\tau)$ it is sufficient to perform the following maximization

$$V(\tau) = \sup_{0 \le \tau \le 1} \mathbb{E}\left[x_\tau^2 \frac{1 - e^{-2a(1-\tau)}}{2a} \right]. \tag{10}$$

Using standard optimal stopping theory [6] let us define the optimum cost to go (Snell envelope) as follows

$$V_t(x) = \sup_{t \le \tau \le 1} \mathbb{E}\left[x_\tau^2 \frac{1 - e^{-2a(1-\tau)}}{2a} \Big| x_t = x \right]. \tag{11}$$

If one has the function $V_t(x)$ then it is straightforward to find the optimal sampling policy. Unfortunately this function is usually very difficult to obtain analytically, we therefore resort to numerical approaches. By discretizing time with step $\delta = 1/N$, we define a sequence of (conditionally with respect to x_0) Gaussian random variables $x_1, \ldots, x_N$, that satisfy the AR(1) model

$$x_n = e^{-a\delta} x_{n-1} + w_n,\ \ w_n \sim \mathcal{N}\left(0, \frac{1 - e^{-2a\delta}}{2a}\right); 1 \le n \le N.$$

As it is indicated, w_n are i.i.d. Gaussian random variables.

Sampling in discrete time means selecting a sample x_ν from the set of $N + 1$ sequentially available random variable $x_0, \ldots, x_N$, with the help of a stopping time $\nu \in \{0, 1, \ldots, N\}$. As in (11) we can define the optimum cost to go which can be analyzed as below. For $n = N, N-1, \ldots, 0$,

$$\begin{aligned}V_n(x) &= \sup_{n \le \nu \le N} \mathbb{E}\left[x_\nu^2 \frac{1 - e^{-2a\delta(N-\nu)}}{2a} \Big| x_n = x \right] \\ &= \max\left\{ x^2 \frac{1 - e^{-2a\delta(N-n)}}{2a}, \mathbb{E}[V_{n+1}(x_{n+1}) | x_n = x] \right\}.\end{aligned}$$

Equ. (12) provides a (backward) recurrence relation for the computation of the cost function $V_n(x)$. Notice that for values of x for which the l.h.s. in (12) exceeds the r.h.s. we stop and sample, otherwise we continue to the next time instant. We can prove by induction that the optimal policy is a *time-varying threshold* one. Specifically for every time n there exists a threshold λ_n such that if $|x_n| \ge \lambda_n$ we sample, otherwise we go to the next time instant. The numerical solution of the recursion presents no special difficulty. If $V_n(x)$ is sampled in x then this function is represented as a vector. In the same way we can see that the conditional expectation is reduced to a simple matrix-vector product. Using this idea we can compute numerically the evolution of the threshold λ_t with time. Fig. 3 depicts examples of threshold time evolution for values of the parameter $a = -1, 0, 1$.

Using $V_n(x)$ the final optimum cost can be computed from (3) as

$$\mathcal{J}(\tau) = \frac{e^{-2a} - 1 + 2a}{4a^2} - \mathbb{E}\left[V_0(x_0) - x_0^2 \frac{1 - e^{-2a}}{2a} \right].$$

Since from the recursion we know that $V_0(x_0) = x_0^2(1 - e^{-2a})/2a$ for $|x_0| \ge \lambda_0$, we conclude that we can also write

$$\begin{aligned}\mathcal{J}(\tau) = & \frac{e^{-2a} - 1 + 2a}{4a^2} \\ & - \mathbb{E}\left[\left\{ V_0(x_0) - x_0^2 \frac{1 - e^{-2a}}{2a} \right\} \mathbb{1}_{\{|x_0| \le \lambda_0\}} \right].\end{aligned} \tag{12}$$

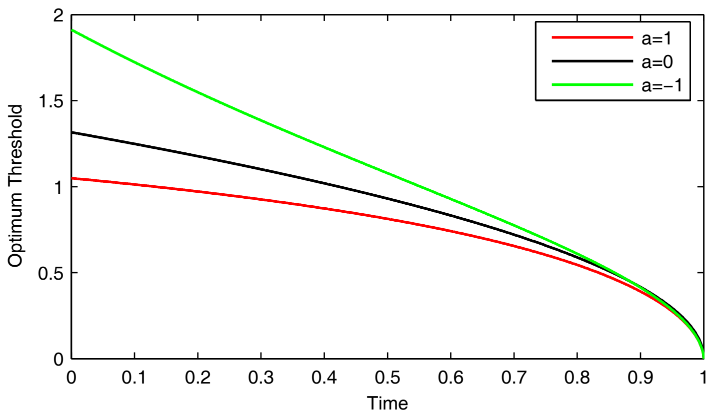

Fig. 3. Time evolution of the optimal threshold λ_t for parameter values $a = 1, 0, -1$.

1) The Wiener case: Let us now focus on the case $a = 0$ which gives rise to a Wiener process. We consider this special case because it is possible to obtain an analytic solution for the optimization problem. For $a = 0$ the optimization in (10) takes the form

$$V(\tau) = \sup_{0\leq\tau\leq 1} \mathbb{E}\left[x_\tau^2(1-\tau)\right].$$

Consider the following function of t and x

$$\mathcal{V}_t(x) = A\left\{\frac{1}{2}(1-t)^2 + x^2(1-t) + \frac{x^4}{6}\right\} \quad (13)$$

where $A = \sqrt{3}/(1+\sqrt{3})$. Using standard Itô calculus, if x_t is a standard Wiener process, we can show that

$$\mathbb{E}[\mathcal{V}_\tau(x_\tau)|x_0] - \mathcal{V}_0(x_0) = \mathbb{E}\left[\int_0^\tau d\mathcal{V}_t(x_t)|x_0\right] = 0 \quad (14)$$

for any stopping time τ. Notice now that

$$\mathcal{V}_t(x) - x^2(1-t) = A\left(\frac{x^2}{\sqrt{6}} - \frac{1-t}{\sqrt{2}}\right)^2 \geq 0. \quad (15)$$

Combining (14) and (15) we conclude that for any stopping time τ

$$\mathcal{V}_0(x_0) = \mathbb{E}[\mathcal{V}_\tau(x_\tau)|x_0] \geq \mathbb{E}[x_\tau^2(1-\tau)|x_0].$$

This relation suggests that the performance of any stopping time τ is upper bounded by $\mathcal{V}_0(x_0)$. Consequently if we can find a stopping time with performance equal to this value then it will be optimal. In fact such a stopping time exists. From the previous relation the last inequality becomes an equality if at the time of sampling τ we have $\mathcal{V}_\tau(x_\tau) = x_\tau^2(1-\tau)$. From (15) we conclude that this can happen iff $|x_\tau|$ is such that the rhs in (15) is exactly 0 which happens if $x_\tau^2/\sqrt{6} = (1-\tau)/\sqrt{2}$. This suggests that the optimal threshold for the Wiener process is the following function of time

$$\lambda_t = \sqrt[4]{3}\sqrt{1-t}.$$

The optimal value of the performance measure, from (12) and letting $a \to 0$, becomes

$$\mathcal{J}(\tau) = \frac{1}{2} - \mathbb{E}\left[\left\{\mathcal{V}_0(x_0) - x_0^2\right\} \mathbb{1}_{\{|x_0|\leq\lambda_0\}}\right],$$

where $\mathcal{V}_t(x)$ is defined in (13).

IV. Comparisons

We have seen that the best sampling strategy is an event-trigerred one. Below, we will see graphically that a simpler event-trigerred strategy based on a constant threshold, is largely of good performance compared to the time-trigerred one, thus providing more ammunition to the ideas of [1] Let us now compare the performance of the three sampling schemes (deterministic, constant thresholding and optimal) for values of the parameter $a = 10, 1, 0, -1$. Regarding threshold sampling we apply the suboptimal version which uses a constant threshold. For the pdf of the initial value x_0 we assume zero mean Gaussian with variance σ^2 ranging from 10^{-4} to 10^4. Figures (4,5,6,7) depict the relative performances of the three schemes with the graphs being normalized so that the maximum is 1. In (a), (b) where a is positive (stable process) the performance of the threshold policy is very close to the optimum and the gain, compared to deterministic sampling, is more important. When however we go to values of a that give rise to unstable processes, threshold sampling starts diverging from the optimal, as in (c) and (d) and, although not shown here, when a is less than -5 (strongly unstable process) deterministic sampling can even perform better than threshold sampling.

References

[1] K. J. Åström, B. Bernhardsson, *Comparison of Riemann and Lebesgue sampling for first order stochastic systems*, Proceedings of the 41st IEEE Conference on Decision and Control, Las Vegas NV, (2002).

[2] L. Chen, M. J. Wainwright, M. Cetin, and A. S. Willsky, *Data assocoation based on optimization in graphical models with applications to sensor networks*, Mathematical and Computer Modeling(Special issue om Optimization and Control for Military applications), (2005).

[3] H. Gharavi, and, S. Kumar, (Eds.)., Special issue on sensor networks and applications, Proceedings of the IEEE, 91(8), (2003).

[4] M. Jones, S. Mehrotra, and J. H. Park, *Tasking distributed sensor networks* , International Journal of High Performance Computer applications, 16, no:3 (2002), pp. 243–257.

[5] S. Kumar, D. Shepherd, and, F. Zhao, *Collaborative Signal and Information Processing in Micro-Sensor Networks*, IEEE Signal Processing Magazine, 19(2), (2002) pp. 13–14. B. Øksendal, *Stochastic differential equations*, Universitext, Springer-Verlag, Berlin, sixth ed., 2003. An introduction with applications.

[6] A. N. Shiryaev, *Optimal stopping rules*, Springer-Verlag, 1978. tranlated from the Russian *Statisticheskii posledovatelnyi analiz.*

[7] F. Zhao, J. Shin, and, J. Reich, *Information-Driven Dynamic Sensor Collaboration for Tracking Applications*, IEEE Signal Processing Magazine, 19(2) (2002), pp. 61-72.

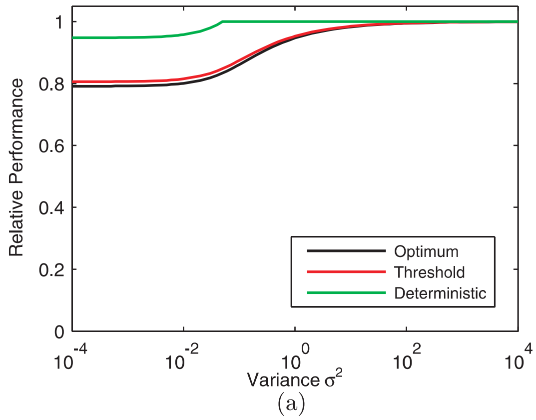

Fig. 4. Relative performance of Optimal, Threshold and Deterministic samplers as a function of initial variance σ^2 and parameter value $a = 10$.

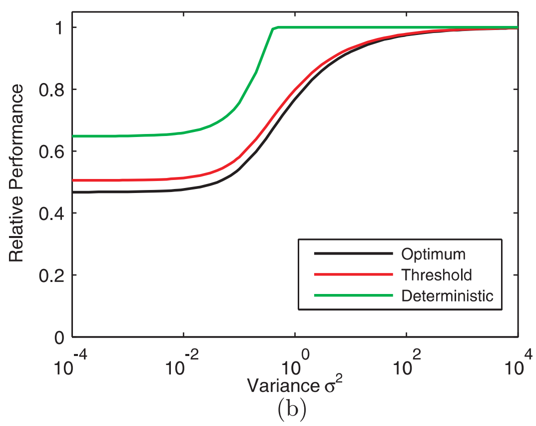

Fig. 5. Relative performance of Optimal, Threshold and Deterministic samplers as a function of initial variance σ^2 and parameter value $a = 1$.

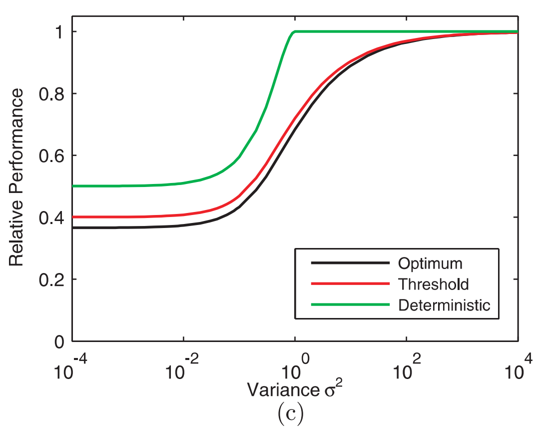

Fig. 6. Relative performance of Optimal, Threshold and Deterministic samplers as a function of initial variance σ^2 and parameter value $a = 0$.

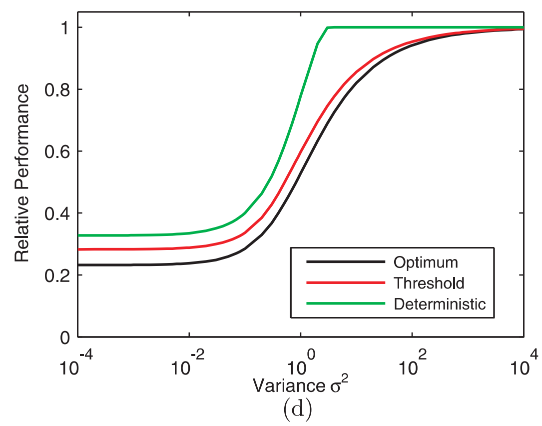

Fig. 7. Relative performance of Optimal, Threshold and Deterministic samplers as a function of initial variance σ^2 and parameter value $a = -1$.