

# Optimal Self-Driven Sampling for Estimation Based on Value of Information

Touraj Soleymani, *Student Member, IEEE*, Sandra Hirche, *Senior Member, IEEE*,  
and John S. Baras, *Life Fellow, IEEE*

**Abstract**—Consider an observer (reporter) who desires to inform optimally a distant agent regarding a physical stochastic process in the environment while the directed communication of the observer to the agent has a price. We define a metric, from a task oriented perspective, for the information transferred from the observer to the agent. We develop a framework for optimizing an augmented cost function which is a convex combination of the transferred information and the paid price over a finite horizon. We suppose that the decision making takes place inside a source encoder, and that the sampling schedule is the decision variable. Moreover, we assume that no measurement at the current time is available to the observer for the decision making. We derive the optimal self-driven sampling policy using dynamic programming, and we show that this policy corresponds to a self-driven sampling policy based on a quantity that is in fact the value of information at each time instant. In addition, we use a semi-definite programming relaxation to provide a suboptimal sampling policy. Numerical and simulation results are presented for a simple unstable system.

**Index Terms**—Approximate Dynamic Programming, Estimation, Self-Driven Sampling, Value of Information.

## I. INTRODUCTION

Consider an observer (reporter) who desires to inform optimally a distant agent regarding a physical stochastic process in the environment. In practice, the directed communication of the observer to the agent has a price due to the associated energy consumption. For instance, in a simple scenario the energy for transmitting a single bit by radio over a distance of 100 meters is the same as that for executing 3000 instruction lines [1]. Consequently, the observer should transmit only measurements that are of more valuable *information* to the agent. Applications of this study include surveillance and reconnaissance, planetary exploration, wearable sensing, teleoperation, and many other examples of networked systems.

We define information transferred from the observer to the agent as the change in the knowledge possessed by the agent about the state of the process. We develop a framework for optimizing an augmented cost function which is a convex combination of the transferred information and the paid price over a finite horizon. We suppose that the decision making takes place inside a *source encoder*, and that the

sampling schedule is the decision variable. Moreover, we assume that no measurement at the current time is available to the observer for the decision making.

Nonuniform sampling is a new paradigm for improving the performance of systems in signal processing, estimation, and control when communication effects and its limitations matter [2], [3]. A special and important case in nonuniform sampling is *event-driven sampling*. In event-driven sampling for estimation, a measurement is sampled whenever an event is generated. The mechanism that generates such an event is called event-triggering mechanism. In the last few years, several event-triggering mechanisms have been proposed including ones based on the error between the current measurement and the last transmitted measurement [4], [5], on the error between the current measurement and the predicted measurement [6], and on the Kullback-Leibler divergence between the prior and the posterior conditional distributions [7].

In a seminal work, Åström and Bernhardsson [8] show that event-driven sampling can outperform periodic sampling with respect to the estimation error of a scalar linear system under a sampling rate constraint. Rabi *et al.* [9] study optimal sampling as a stopping time problem for a scalar system under a finite transmission budget constraint. Molin and Hirche [10] investigate the optimal design for sampling in a scalar system with a communication cost by considering a two-player problem. Moreover, Sijs and Lazar [11] study event-driven sampling for the estimation problem with an asymptotic bound on the estimation error covariance.

Event-driven sampling can reduce communication cost while providing a good estimation performance. However, it requires the observer to monitor the process constantly. Roughly similar to [12], we define *self-driven sampling* for estimation as a technique in which the observer is self-contained in decision making and can compute the sampling schedule without monitoring the process constantly. An example of a self-triggering mechanism for sampling in the estimation problem is the one based on the covariance of the estimation error [13]. In [14], we proposed an optimized LMI-based self-triggering mechanism in which the parameter of the mechanism is found by solving an optimization problem. However, we showed that the sampling policy specified by this mechanism is not necessarily optimal. In this paper, we do not make any assumption on the structure of the self-triggering mechanism. In contrast, we derive the optimal sampling policy using dynamic programming (DP), and we show that this policy corresponds to a self-driven

T. Soleymani and S. Hirche are with the Institute for Information-Oriented Control, Technische Universität München, D-80333 Munich, Germany (email: {touraj, hirche}@tum.de).

J. S. Baras is with the Institute for Systems Research, University of Maryland College Park, MD 20742, USA (email: baras@umd.edu).

T. Soleymani and J. S. Baras are also with the Institute for Advanced Study, Technische Universität München, D-85748 Garching, Germany.

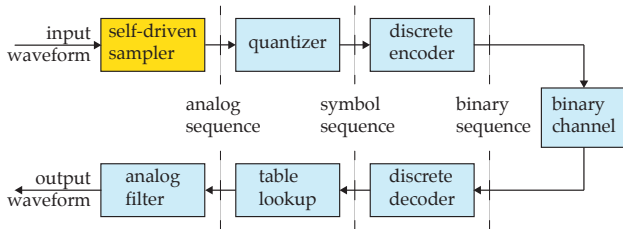


Fig. 1. Three layers of source encoding/decoding. The focus of this study is on designing an optimal self-driven sampler.

sampling policy based on a quantity that is in fact the *value of information* (VOI) at each time instant.

The outline of this paper is as follows. After an introduction on notation, the problem formulation is presented in Section II. In Section III, we obtain the optimal sampling policy and the corresponding self-driven policy. We illustrate numerical and simulation results in Section IV. Finally, concluding remarks are made in Section V.

### A. Notations

In this paper, we represent an  $n$  dimensional vector with  $x = [x_1, \dots, x_n]^T$  where  $x_i$  is its  $i$ th component. We write  $x^T$  to denote the transpose of the vector  $x$ . The identity matrix with dimension  $n$  is denoted by  $I_n$ . We use  $x_{1:k}$  to denote the sequence  $\{x_1, \dots, x_k\}$ . We write  $\delta_{t-t'}$  and  $\delta_{ss'}$  to denote the Dirac and Kronecker delta functions, respectively. We write  $p(x)$  to denote the probability distribution of the stochastic variable  $x$ . The normal distribution with mean  $\mu$  and covariance  $\sigma^2$  is denoted by  $N(\mu, \sigma^2)$ . For matrices  $A$  and  $B$ , we write  $A \succ 0$  and  $B \succeq 0$  to mean that  $A$  and  $B$  are positive definite and positive semi-definite, respectively.

## II. PROBLEM FORMULATION

### A. Physical Process and Source Encoder Models

We use a continuous-discrete time model [15] to describe our system. This is due to the fact that the state of the process evolves in continuous time, and that the measurements are obtained at discrete instants of time. Consider a stochastic continuous-time physical process in the environment which is generated by the following linear stochastic differential equation:

$$\dot{x}_t = A_t x_t + B_t w_t \quad (1)$$

where  $x_t \in \mathbb{R}^n$  is the state of the system at time  $t$ ,  $A_t$  and  $B_t$  are the state and the input matrices at time  $t$ , and  $w_t \in \mathbb{R}^m$  is a white noise process with

$$E[w_t] = 0, \quad E[w_t w_{t'}^T] = Q_t \delta_{t-t'} \quad (2)$$

where  $Q_t \succ 0$ . The initial state  $x_0$  is assumed to be a Gaussian vector with zero mean and covariance  $P_0$ . An observer who desires to inform a distant agent regarding the state of the process employs a source encoder to transmit its measurements in the form of a sequence of bits. Fig. 1 illustrates in general the three layers of source encoding in digital communication, i.e., sampling, quantization, and encoding.

In particular, first a sampler that is equipped with a sensor samples the process at time instants  $t_s$  for  $s = 1, \dots, M$ . The measurement model of the sensor is given by

$$y_s = C_s x_s + v_s \quad (3)$$

where  $y_s \in \mathbb{R}^p$  is the output of the sampler at time instant  $t_s$ ,  $x_s$  is the state of the process at time instant  $t_s$ ,  $C_s$  is the output matrix at time instant  $t_s$ , and  $v_s$  is a white noise sequence with

$$E[v_s] = 0, \quad E[v_s v_{s'}^T] = R_s \delta_{ss'} \quad (4)$$

where  $R_s \succ 0$ .

The sample  $y_s$  is then quantized by a lattice quantizer. Following the high-resolution quantization assumption, we model the quantizer output by

$$z_s = y_s + n_s \quad (5)$$

where  $z_s$  is the quantizer output at time instant  $t_s$  and  $n_s$  is an uncorrelated white additive noise sequence with

$$E[n_s] = 0, \quad E[n_s n_{s'}^T] = \frac{b^2}{12} I_p \delta_{ss'} \quad (6)$$

where  $b$  is the quantization step-size. We use the convention  $b = 0$  for the ideal quantization. The validity of the additive white noise model for quantization is studied in [16].

Finally, an encoder codes the quantized sample  $z_s$ . The bits encoding the codewords of  $z_s$  are then transmitted to the agent through a noiseless zero-delay binary channel. Therefore, at time  $t \geq t_s$  the set of measurements  $z_{1:s}$  is available to the agent. Through this study, the agent assumes that the observer is trustable and that measurements are never compromised.

*Remark 1:* In this setting, the decision making takes place inside the source encoder, and the decision variable is the sampling schedule  $t_{1:M}$ .

### B. Fisher Information Matrix Dynamics

The conditional distribution  $p(x_t | z_{1:s}) = N(\hat{x}_t, P_t)$ , which is Gaussian given the system's model introduced before, evolves in time due to the system dynamics, and is updated at time instants  $t_s$  due to the measurements.

Consider the transformation  $I_t = P_t^{-1}$  and  $\dot{I}_t = -I_t \dot{P}_t I_t$  where  $I_t$  is the Fisher information matrix (FIM) [17]. Following the Kolmogorov forward equation [15], which in our case reduces to the propagation of the conditional mean and conditional covariance matrix of the state, the rate of change of the FIM during the interval  $[t_{s-1}, t_s)$ , in which there are no measurements, is obtained as

$$\dot{I}_t = -A_t^T I_t - I_t A_t - I_t B_t Q_t B_t^T I_t, \quad t \in [t_{s-1}, t_s) \quad (7)$$

with initial condition  $I_{s-1} = I_{(s-1)^-} + \Delta I_{s-1}$  where we mean by  $I_{s-1}^-$  the left hand side limit of  $I_t$  as  $t$  tends to  $t_s$  from the left. In addition, following Bayes' rule [15], the change in the FIM at the time instant  $t_s$ , when a measurement is available, is obtained as

$$\Delta I_s = I_s - I_{s-1}^- = C_s^T \bar{R}_s^{-1} C_s \triangleq G_s \quad (8)$$

where

$$\bar{R}_s = R_s + \frac{b^2}{12} I_p \quad (9)$$

and  $G_s$  is called quality matrix.

*Remark 2:* The time evolution of the FIM in the continuous-discrete time iterative filtering problem is characterized by the jump dynamics defined by the equations (7), (8). Hence, given the sampling schedule  $t_{1:M}$  and the initial FIM  $I_0$  the whole trajectory of the FIM can be obtained.

### C. Measure of Information

Let the agent's knowledge regarding the state of the process given the measurements at times  $t_s$ ,  $t_s \leq t$ , be represented by the conditional distribution  $p(x_t | z_{1:s}) = N(\hat{x}_t, P_t)$ . The differential entropy [17] of the knowledge is given by

$$H(x_t | z_{1:s}) = \frac{1}{2} \log_2 \left[ (2\pi e)^n \det I_t^{-1} \right] \quad (10)$$

where  $n$  is the dimension of  $x_t$ .

*Definition 1 (Information from agent's perspective):*

Information is the change in the knowledge possessed by the agent about the state of the process given the transmitted measurements over the time interval  $[t_0, t]$  with  $t \geq t_0$ , i.e.,

$$\Pi_t \triangleq H_0 - H(x_t | z_{1:s}) \quad (11)$$

where  $H_0$  is the entropy at the reference time  $t_0$ .

In fact, this information has a destination (i.e., the agent) and is defined with respect to an epoch (i.e., time  $t_0$ ).

*Remark 3:* Stochastic dynamical systems evolve in the direction in which entropy increases. The *dissipated information* is the amount of information loss through the natural evolution of the system. However, the entropy can decrease as a result of measurements. The *mutual information* between the state and a measurement is the amount of information gain about the state provided by that measurement.

For maximizing the cumulative information about the state of the process, from the measurements, which is the integral of the information defined in (11) over the time horizon  $[0, T]$ , we can minimize the following cost function:

$$J_I(I_0, t_{1:M}) = - \int_0^T \log_2 \det I_t \, dt \quad (12)$$

where  $I_0$  is the initial FIM.

### D. Communication Price

The directed communication of the observer to the agent is costly. In general, the communication price depends on the state of the channel and on the measurements. Let the communication price per measurement at time  $t_s$  be denoted by  $\lambda_s$ . Then, the communication cost over the time horizon  $[0, T]$  is given by

$$J_{II}(t_{1:M}) = \sum_{s=1}^M \lambda_s \quad (13)$$

where  $M$  is the number of samples over the time horizon  $[0, T]$ . In the following, we assume that  $\lambda_s = \lambda$  is constant.

### E. Objective

We define the augmented cost function over the time horizon  $[0, T]$  as a convex combination of the cost functions (12) and (13), i.e.,

$$J(I_0, t_{1:M}) = \theta J_I(I_0, t_{1:M}) + (1 - \theta) J_{II}(t_{1:M}). \quad (14)$$

where  $\theta \in [0, 1]$ . The goal of this paper is to solve the following problem.

*Problem 1:* Given the system defined by (7), (8), design a source encoder parametrized by a sampling schedule  $t_{1:M}$  that minimizes the augmented cost function  $J(I_0, t_{1:M})$  defined in (14) over the time horizon  $[0, T]$ .

## III. OPTIMAL SELF-DRIVEN SAMPLING

### A. Optimal Self-Driven Policy

We first discretize the time variable  $t$ . Let  $t_k = k\epsilon$  and  $N = T/\epsilon$  where  $\epsilon$  is the time step-size. We define a logical decision variable  $\delta_k$  at time step  $k$  as

$$\delta_k = \begin{cases} 1, & \text{if } t_k = t_s, \\ 0, & \text{otherwise.} \end{cases} \quad (15)$$

Note that by specifying the decision set  $\delta_{1:N}$ , the sampling schedule  $t_{1:M}$  is determined.

Using the Riemann method and the equation (15), the augmented cost function in discrete time is written as

$$J(I_0, t_{1:M}) = \sum_{k=0}^N g(I_k, \delta_k) \quad (16)$$

where  $g_k(I_k, \delta_k)$  is the stage cost at time step  $k$  defined by

$$g_k(I_k, \delta_k) = -\theta \epsilon \log_2 \det I_k + (1 - \theta) \lambda \delta_k \quad (17)$$

where  $\delta_0 = 0$ . Moreover, using the Euler method and the equation (15) we write the FIM dynamics by the following difference equation

$$I_k = \phi_k(I_{k-1}, \delta_k) \quad (18)$$

where

$$\begin{aligned} \phi_k(I_{k-1}, \delta_k) = & -\epsilon A_{k-1}^T I_{k-1} - \epsilon I_{k-1} A_{k-1} + I_{k-1} \\ & - \epsilon I_{k-1} B_{k-1} Q_{k-1} B_{k-1}^T I_{k-1} + G_k \delta_k. \end{aligned} \quad (19)$$

Now, we can mathematically formulate Problem 1 by the following finite horizon optimization problem

$$\begin{aligned} & \text{minimize} && \sum_{k=0}^N -\theta \epsilon \log_2 \det I_k + (1 - \theta) \lambda \delta_k && (20) \\ & \text{subject to} && I_k = -\epsilon A_{k-1}^T I_{k-1} - \epsilon I_{k-1} A_{k-1} + I_{k-1} \\ & && - \epsilon I_{k-1} B_{k-1} Q_{k-1} B_{k-1}^T I_{k-1} + G_k \delta_k, \end{aligned}$$

with variables  $I_k \succ 0$  and  $\delta_k \in \{0, 1\}$  for all  $k = 1, 2, \dots, N$ , and with initial condition  $I_0$ .

Define the value function  $V_k(I_{k-1})$  at time step  $k$  as

$$\begin{aligned} V_k(I_{k-1}) &= \inf_{\delta_{k:N}} \sum_{l=k}^N g_l(I_l, \delta_l) \\ &= \inf_{\delta_{k:N}} \sum_{l=k}^N g_l(\phi_l(I_{l-1}, \delta_l), \delta_l). \end{aligned} \quad (21)$$

*Proposition 1:* Let  $V_k(I_{k-1})$  be the value function at time step  $k$  with condition

$$V_{N+1}(I_N) = 0. \quad (22)$$

Then, the optimal policy  $\delta_k^*(I_{k-1})$  at time step  $k$  is given by

$$\begin{aligned} \delta_k^*(I_{k-1}) &= \operatorname{argmin}_{\delta_k} \left\{ g_k(\phi_k(I_{k-1}, \delta_k), \delta_k) + V_{k+1}(\phi_k(I_{k-1}, \delta_k)) \right\}. \end{aligned} \quad (23)$$

*Proof:* The optimal decision set  $\delta_{1:N}^*$  is obtained by solving Bellman's optimality equation backward in time from  $k = N$  to  $k = 1$ :

$$\begin{aligned} V_k(I_{k-1}) &= \min_{\delta_k} \left\{ g_k(\phi_k(I_{k-1}, \delta_k), \delta_k) + V_{k+1}(\phi_k(I_{k-1}, \delta_k)) \right\}. \end{aligned} \quad (24)$$

Therefore, the optimal solution  $\delta_k^*$  is given by the argument of the minimum in (24). ■

One can calculate the optimal sampling policy using DP as stated in Proposition 1 backward in time. However, the DP calculation for obtaining the value function becomes intractable when the time space and the state space are large (known as curse of dimensionality). In the following, we aim at providing an approximate solution.

### B. One-Step Lookahead Policy

We would like to reduce the computation complexity required by DP. In addition, we are interested in solving the problem online. To this goal, we make a trade-off between the convenient implementation and adequate performance using approximate dynamic programming (ADP). In fact, in ADP we step forward in time and settle for a suboptimal policy.

We propose a one-step lookahead policy in which the horizon is truncated and the optimal cost-to-go of the remaining horizon is approximated. Assume that  $\tilde{V}_{k+1}(I_k)$  is an approximation of the optimal cost-to-go function  $V_{k+1}(I_k)$ . We can obtain a suboptimal policy  $\delta_k^+(I_{k-1})$  at time step  $k$  by

$$\begin{aligned} \delta_k^+(I_{k-1}) &= \operatorname{argmin}_{\delta_k} \left\{ g_k(\phi_k(I_{k-1}, \delta_k), \delta_k) + \tilde{V}_{k+1}(\phi_k(I_{k-1}, \delta_k)) \right\}. \end{aligned} \quad (25)$$

In the following, we use a semi-definite programming (SDP) relaxation [18] to find a lower bound on the optimal cost-to-go function  $V_{k+1}(I_k)$ .

*Theorem 1:* A lower bound on the optimal cost-to-go function  $V_{k+1}(I_k)$  defined in (21) is given by the solution

of the following SDP:

$$\begin{aligned} &\text{minimize} && \sum_{j=1}^{N-k} -\theta\epsilon \log_2 \det I_j + (1-\theta)\lambda\delta_j && (26) \\ &\text{subject to} && I_j = -\epsilon A_{j-1}^T I_{j-1} - \epsilon I_{j-1} A_{j-1} + I_{j-1} \\ &&& \quad - \epsilon D_{j-1} + G_j \delta_j, \\ &&& \begin{bmatrix} D_j & I_j \\ I_j & (B_j Q_j B_j^T)^{-1} \end{bmatrix} \succeq 0, \end{aligned}$$

with variables  $I_j \succ 0$ ,  $D_j \succ 0$ , and  $\delta_j \in [0, 1]$  for all  $j = 1, 2, \dots, N-k$ , and with initial conditions  $I_0 = I_k$  and  $D_0 = I_k B_k Q_k B_k^T I_k$ .

*Proof:* We can obtain the optimal cost-to-go starting at time step  $k+1$  by solving the following finite horizon optimization problem:

$$\begin{aligned} &\text{minimize} && \sum_{j=k+1}^N -\theta\epsilon \log_2 \det I_j + (1-\theta)\lambda\delta_j && (27) \\ &\text{subject to} && I_j = -\epsilon A_{j-1}^T I_{j-1} - \epsilon I_{j-1} A_{j-1} + I_{j-1} \\ &&& \quad - \epsilon I_{j-1} B_{j-1} Q_{j-1} B_{j-1}^T I_{j-1} + G_j \delta_j, \end{aligned}$$

with variables  $I_j \succ 0$  and  $\delta_j \in \{0, 1\}$  for all  $j = k+1, k+2, \dots, N$ , and with initial condition  $I_k$ . Note that the functions  $-\epsilon \log_2 \det I_j$  and  $\lambda\delta_j$  are convex in  $I_j$  and  $\delta_j$  given that  $I_j \succ 0$  and  $\delta_j \in \{0, 1\}$ . Following the optimization problem in (27), we relax the integrality constraint to  $\delta_j \in [0, 1]$ . Then, we aim at relaxing the quadratic term in the FIM dynamics. Since this equality constraint does not influence the set of admissible decisions  $\delta_j$ , such a relaxation does not violate the feasibility of the problem. By defining  $D_j = I_j B_j Q_j B_j^T I_j$ , we have

$$\begin{aligned} I_j &= -\epsilon A_{j-1}^T I_{j-1} - \epsilon I_{j-1} A_{j-1} + I_{j-1} \\ &\quad - \epsilon D_{j-1} + G_j \delta_j \end{aligned} \quad (28)$$

which is a linear equation. We substitute the new constraint  $D_j = I_j B_j Q_j B_j^T I_j$  with its convex relaxation, i.e.,  $D_j - I_j B_j Q_j B_j^T I_j \succeq 0$ . Then, this nonlinear convex inequality can be converted into an LMI by using the Schur complement:

$$\begin{bmatrix} D_j & I_j \\ I_j & (B_j Q_j B_j^T)^{-1} \end{bmatrix} \succeq 0. \quad (29)$$

Substituting the constraints  $I_j = \phi_j(I_{j-1}, \delta_j)$  and  $\delta_j \in \{0, 1\}$  with the constraints (28), (29), and  $\delta_j \in [0, 1]$  in the optimization problem in (27), and following the convexity of the cost function we obtain the relaxed problem (26) which yields a lower bound for the original problem. ■

*Remark 4:* The lower bound obtained in Theorem 1 can be used as an approximation to the optimal cost-to-go function  $V_{k+1}(I_k)$  defined in (21).

### C. Value of Information and the Self-Triggering Mechanism

Transferring information to the agent on one side has a value as it decreases the entropy. On the other side, the

directed communication of the observer to the agent has a price. Consequently, at each time instant the observer is dealing with the question whether it is worth to transmit a measurement or not. In the following, we quantify this worthiness.

*Definition 2 (Value of information):* The value of information (VOI) is the maximum value that the observer will be willing to pay for the transmission of a measurement  $z_k$  at time instant  $t_k$ , i.e.,

$$\begin{aligned} \text{VoI}_k(I_{k-1}) &= g_k(\phi_k(I_{k-1}, 0), 0) - g_k(\phi_k(I_{k-1}, 1), 0) \\ &+ V_{k+1}(\phi_k(I_{k-1}, 0)) - V_{k+1}(\phi_k(I_{k-1}, 1)). \end{aligned} \quad (30)$$

This is closely connected to the VOI in [19] which is used in a prioritization mechanism for multi-sensor scheduling.

*Proposition 2:* Let  $\text{VoI}_k(I_{k-1})$  be the VOI at time step  $k$ . Then, the self-driven sampling policy specified by the following self-triggering mechanism

$$\delta_k^*(I_{k-1}) = \begin{cases} 1, & \text{if } \text{VoI}_k(I_{k-1}) \geq (1 - \theta)\lambda, \\ 0, & \text{otherwise,} \end{cases} \quad (31)$$

is optimal.

*Proof:* Following (23), the optimal decision variable at time step  $k$  is  $\delta_k^* = 1$  if

$$\begin{aligned} g_k(\phi_k(I_{k-1}, 0), 0) + V_{k+1}(\phi_k(I_{k-1}, 0)) \\ - g_k(\phi_k(I_{k-1}, 1), 1) - V_{k+1}(\phi_k(I_{k-1}, 1)) \geq 0 \end{aligned} \quad (32)$$

and is  $\delta_k^* = 0$  otherwise.

We can write  $g_k(\phi_k(I_{k-1}, 1), 1) = g_k(\phi_k(I_{k-1}, 1), 0) + (1 - \theta)\lambda$ . Substituting this in (32) and from Definition 2, we can obtain the optimal sampling policy as specified in (31). ■

Using the lower bound obtained in Theorem 1, we can calculate the approximate VOI at time step  $k$

$$\begin{aligned} \widetilde{\text{VoI}}_k(I_{k-1}) &= g_k(\phi_k(I_{k-1}, 0), 0) - g_k(\phi_k(I_{k-1}, 1), 0) \\ &+ \widetilde{V}_{k+1}(\phi_k(I_{k-1}, 0)) - \widetilde{V}_{k+1}(\phi_k(I_{k-1}, 1)). \end{aligned} \quad (33)$$

Then, we can implement the following self-triggering mechanism

$$\delta_k^+(I_{k-1}) = \begin{cases} 1, & \text{if } \widetilde{\text{VoI}}_k(I_{k-1}) \geq (1 - \theta)\lambda, \\ 0, & \text{otherwise.} \end{cases} \quad (34)$$

The sampling policy provided by this mechanism is suboptimal, computationally cheap, and easily implementable.

#### D. State Estimate Dynamics

The agent can estimate the state of the process based on the transmitted measurements from the observer. As it is said, at time  $t \geq t_s$  the set of measurements  $z_{1:s}$  is available to the agent. The Kalman filter is the optimal estimator since the conditional distribution is Gaussian. Hence, the dynamics of the state estimate in the intervals  $[t_{s-1}, t_s]$  is

$$\dot{\hat{x}}_t = A_t \hat{x}_t, \quad t \in [t_{s-1}, t_s] \quad (35)$$

with initial condition  $\hat{x}_{s-1} = \hat{x}_{(s-1)^-} + \Delta \hat{x}_{s-1}$ , and the change in the state estimate at the time instant  $t_s$  is

$$\Delta \hat{x}_s = \hat{x}_s - \hat{x}_{s-1} = K_s(z_s - C_s \hat{x}_{s-1}) \quad (36)$$

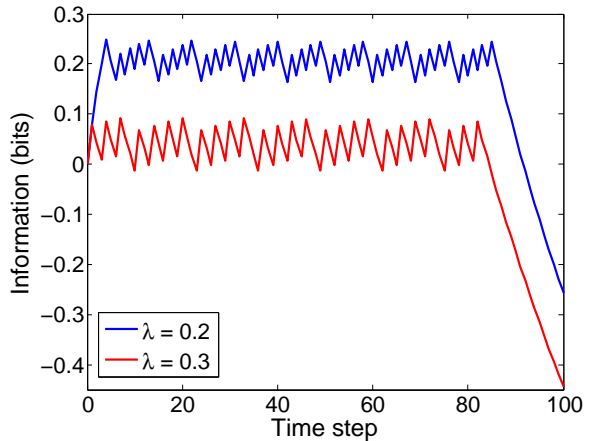


Fig. 2. Information diagrams corresponding to the self-driven sampling for two different communication prices.

where  $K_s$  is the optimal gain given by

$$K_s = I_s^{-1} C_s^T \bar{R}_s^{-1}. \quad (37)$$

Notice that  $I_s$  here is the FIM at time instant  $t_s$  after being updated.

#### IV. AN ILLUSTRATIVE EXAMPLE

In this section, we present numerical and simulation results for a simple unstable system.

*Example 1:* Consider a system with dynamics defined by

$$\begin{aligned} \dot{x}_t &= 0.1x_t + w_t \\ y_k &= \sqrt{0.1}x_k + v_k \\ z_k &= y_k + n_k \end{aligned} \quad (38)$$

where the covariances are  $P_0 = 2$ ,  $Q = 0.5$ ,  $R = 1.245$ , and  $\bar{R} = 1.25$ . It is assumed that the time step-size is  $\epsilon = 0.1$ , the horizon is  $N = 100$ , and  $\theta = 0.5$ . We employ the self-driven sampling policy introduced in (34) for two different communication prices per measurement  $\lambda = 0.2, 0.3$ .

Fig. 2 shows the information diagrams corresponding to the self-driven sampling policy in (34) for  $\lambda = 0.2, 0.3$ . As we can see, the amount of information transferred from the observer to the agent is higher for the lower communication price.

#### V. CONCLUSION

In this work, we developed a framework for obtaining the optimal self-driven sampling policy to be used by an observer who desires to inform a distant agent regarding a physical stochastic process in the environment. We assumed that no measurement at the current time is available to the observer for the decision making. Therefore, the observer should resort to the information that exists in the FIM (or the covariance) of the estimation error. We used DP to find the optimal sampling policy that optimizes an augmented cost function defined as a convex combination of the transferred information and the paid price over a finite horizon. We showed that this policy can be generated by a self-triggering mechanism based on the VOI at each time instant. In

addition, we used an SDP relaxation to find an approximation to the optimal cost-to-go function, and proposed a one-step lookahead policy in order to obtain a suboptimal sampling policy which can cheaply be computed.

#### ACKNOWLEDGMENT

This work has been carried out with the support of the Technische Universität München - Institute for Advanced Study, funded by the German Excellence Initiative. John S. Baras' research was also supported by NSF grant CNS-1035655, and by DARPA (ARO) grant W911NF1410384.

#### REFERENCES

- [1] R. Zheng, J. Hou, and N. Li, "Power management and power control in wireless networks," *Ad Hoc and Sensor Networks. Nova Science Publishers*, 2004.
- [2] H. Ishii and B. A. Francis, *Limited data rate in control systems with networks*, vol. 275. Springer Science & Business Media, 2002.
- [3] F. Marvasti, *Nonuniform sampling: theory and practice*. Springer Science & Business Media, 2012.
- [4] P. G. Otañez, J. R. Moyne, and D. M. Tilbury, "Using deadbands to reduce communication in networked control systems," in *Proc. of the IEEE American Control Conf.*, vol. 4, pp. 3015–3020, 2002.
- [5] M. Miskowicz, "Send-on-delta concept: an event-based data reporting strategy," *Sensors*, vol. 6, no. 1, pp. 49–63, 2006.
- [6] J. Wu, Q.-S. Jia, K. H. Johansson, and L. Shi, "Event-based sensor data scheduling: Trade-off between communication rate and estimation quality," *IEEE Trans. on Automatic Control*, vol. 58, no. 4, pp. 1041–1046, 2013.
- [7] J. W. Marck and J. Sijs, "Relevant sampling applied to event-based state-estimation," in *Proc. of the International Conf. on Sensor Technologies and Applications*, pp. 618–624, 2010.
- [8] K. J. Åström and B. Bernhardsson, "Comparison of Riemann and Lebesgue sampling for first order stochastic systems," in *Proc. of the IEEE Conf. on Decision and Control*, vol. 2, pp. 2011–2016, 2002.
- [9] 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.
- [10] A. Molin and S. Hirche, "An iterative algorithm for optimal event-triggered estimation," in *Proc. of the IFAC Conf. on Analysis and Design of Hybrid Systems*, pp. 64–69, 2012.
- [11] J. Sijs and M. Lazar, "Event based state estimation with time synchronous updates," *IEEE Trans. Automatic Control*, vol. 57, no. 10, pp. 2650–2655, 2012.
- [12] A. Anta and P. Tabuada, "To sample or not to sample: Self-triggered control for nonlinear systems," *IEEE Trans. Automatic Control*, vol. 55, no. 9, pp. 2030–2042, 2010.
- [13] S. Trimpe and R. D'Andrea, "Event-based state estimation with variance-based triggering," *IEEE Trans. on Automatic Control*, vol. 59, no. 12, pp. 3266–3281, 2014.
- [14] T. Soleymani, S. Hirche, and J. S. Baras, "Maximization of information in energy-limited directed communication," in *Proc. of the IEEE European Control Conf.*, pp. –, 2016.
- [15] A. Jazwinski, "Filtering for nonlinear dynamical systems," *IEEE Trans. on Automatic Control*, vol. 11, no. 4, pp. 765–766, 1966.
- [16] D. Marco and D. L. Neuhoff, "The validity of the additive noise model for uniform scalar quantizers," *IEEE Trans. Information Theory*, vol. 51, no. 5, pp. 1739–1755, 2005.
- [17] D. Guo, S. Shamaï, and S. Verdú, "The interplay between information and estimation measures," *Foundations and Trends in Signal Processing*, vol. 6, no. 4, pp. 243–429, 2012.
- [18] S. Boyd and L. Vandenberghe, "Semidefinite programming relaxations of non-convex problems in control and combinatorial optimization," in *Communications, Computation, Control, and Signal Processing*, pp. 279–287, Springer, 1997.
- [19] A. Molin, C. Ramesh, H. Esen, and K. H. Johansson, "Innovations-based priority assignment for control over CAN-like networks," in *Proc. of the IEEE Conf. on Decision and Control*, pp. –, 2015.