

Optimal Stopping for Event-triggered sensing and actuation

Maben Rabi, Karl H. Johansson, and Mikael Johansson

School of Electrical Engineering
Royal Institute of Technology (KTH)
100 44 Stockholm, Sweden
firstname.lastname@ee.kth.se

Abstract—Novel event-triggered sensing and actuation strategies are presented for networked control systems with limited communication resources. Two architectures are considered: one with the controller co-located with the sensor and one with the control co-located with the actuator. A stochastic control problem with an optimal stopping rule is shown to capture two interesting instances of these architectures. The solution of the problem leads to a parametrization of the control alphabet as piecewise constant commands. The execution of the control commands is triggered by stopping rules for the sensor. In simple situations, it is possible to analytically derive the optimal controller. Examples illustrate how the new event-based control and sensing strategies outperform conventional time-triggered schemes.

I. INTRODUCTION

This paper studies the problem of adaptive sensor sampling and switching control. In a networked control problem, the controller is remotely located to the sensor, or the actuator, or sometimes, to both.

When the sensor is remote to the controller (see figure (1)), continuous transmission of the sensor measurements to the controller are impossible; only a sampled version of the measurements can be forwarded to the controller. The control signal is generated with some compressed information about the state process. It is generated based on the samples received, and is forced to be piecewise deterministic. For the actuator, even though it is co-located with the controller, for the sake of ease of actuation, it is sometimes useful to further restrict the control signal to be piece-wise constant.

If, on the other hand, the controller is co-located with the sensor (see figure (2)), the actuator would be remotely located and the control signal can only be updated at instants the actuator receives information from the controller. Thus the control signal is again forced to be piece-wise deterministic. For the sake of ease of implementation, again, we are led to consider piece-wise constant control.

Thus, on a given finite interval, the control signal can change values no more than a prescribed number of times. The control task is to obtain good stabilization performance while not exceeding the limits on switching rate.

A. Problem Formulation

We now formulate the Optimal control problem of choosing sample and hold controls with a fixed number of allowed

This research was partially funded by the Swedish Research Council, the Swedish Foundation for Innovation Systems, and the European Commission.

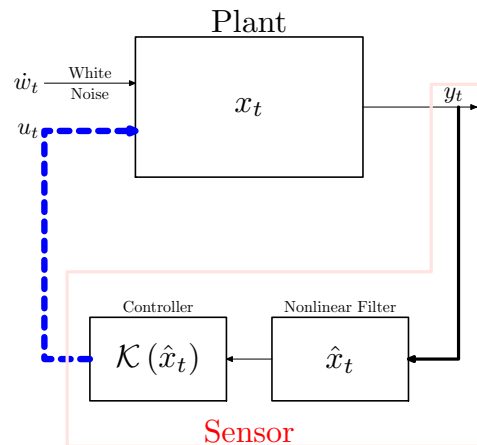


Fig. 1. Schematic of Control station being co-located with the Sensor. The dashed blue link represents restricted communication.

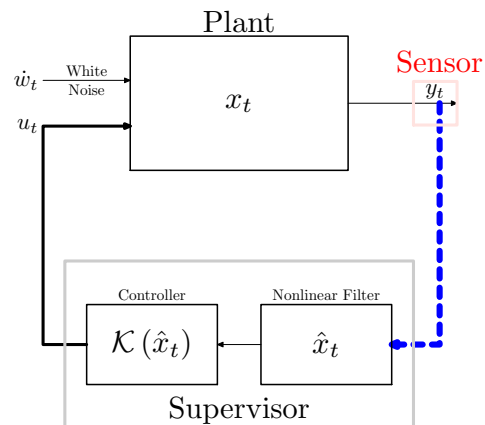


Fig. 2. Schematic of Control station being co-located with the Actuator. The dashed blue link represents restricted communication.

samples. On the time interval $[0, T]$, consider a scalar controlled diffusion process x_t governed by the SDE:

$$dx_t = \mu(x_t, t, u_t) dt + \sigma(x_t, t, u_t) dB_t, \quad (1)$$

where, B_t is a standard Brownian motion process, the functions μ, σ are Lipschitz continuous, and the control process u_t , satisfies the usual causality property of being measurable w.r.t the x -process, thus ensuring existence and uniqueness of the state process x_t [3], [9]. In this paper, we will assume that a sensor measures the state signal continuously and with

negligible measurement noise. This and the the causality of the u_t process means that the times of discontinuity have to be stopping times with respect to the x -process.

$$u_t = U_N \cdot 1_{\{\tau_N \leq t < T\}} + \sum_{i=0}^{N-1} U_i \cdot 1_{\{\tau_i \leq t < \tau_{i+1}\}},$$

where the times τ_i are the sensor sampling times, or equivalently, the control level switch times. We seek a feedback

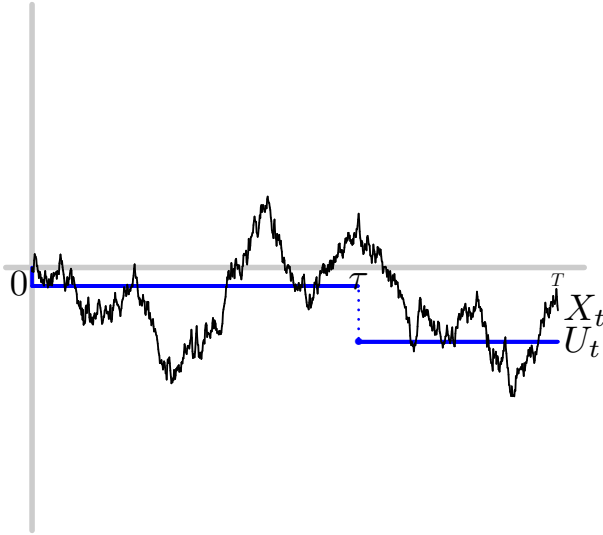


Fig. 3. Piece-wise constant control with one allowed switch

control policy $\mathcal{U}_N : \mathcal{F}_t^x \rightarrow u_t$ that meets the restrictions described above and minimizes the objective function:

$$J_{\text{General}} = \mathbb{E} \left[\int_0^T c(x_s) ds \right], \quad (2)$$

where the function $c(\cdot)$ is non-negative.

B. Relation to previous work

Kushner [7] has considered a variation of the finite horizon LQG control problem where the measurement sampling instants are also decision variables which are to be chosen offline. But in his formulation, only the terminal state and the control energy are being optimized. In our work where we permit an online choice of sampling times, we attack the harder problem of minimizing the deviation of the state at all times, not just the deviation of the terminal state. We however, do not penalize the control energy, because, although there are some qualitative differences from the version penalizing the control effort also, our methods of analysis and design can be extended to cover that situation as well.

Åström and Bernhardsson [2] treat a problem of minimum variance stabilization of a scalar linear plant using state-resetting impulse controls. They provides a comparison of periodic impulse control and impulse control triggered by Δ -excursions of the state from the origin. By explicit calculations, it shows that for the same average rate of impulse invocation, the level-triggered scheme provides lower average

stabilization error. This article provides a similar comparison when the control is not of impulse type, but is piecewise continuous. In fact the class of control signals considered is piecewise constant, these being natural candidates for situations in networked control.

The problem of jointly optimal sampling and control is much more difficult than the jointly optimal sampling and filtering problem tackled in earlier work [10], [11]. In the estimation problem, the error signal is reset to zero at sampling times. In the control problem however, the state signal does not get reset to zero like the error signal does for the estimation problem. Thus, no reduction to repeating the same sampling policy is possible. In practical terms, this means that the feedback control signal as well as the sampling policy should be ‘aggressive’ when the state wanders away from the origin.

This problem differs in its information pattern from similar ones addressed in the Stochastic Control literature. The works [6], [5], [8] seek combined control and stopping policies with both of these adapted to the same signal process, namely the state process. In our problem on the other hand, the stopping policy is allowed to depend on the x -signal process while the control signal is adapted to the sampled sequence. The work of [1] discusses the LQG control performance under Poisson sampling. A deterministic version of control for event-triggered sampling is presented in [4].

II. SPECIAL CASE: BROWNIAN MOTION STARTING AT THE ORIGIN

In this section, we will describe a special case for which the optimal control and its performance can be explicitly computed in closed form.

Specialize equation (1) of state evolution to the case of controlled Brownian motion on the interval $[0, T]$:

$$dx_t = dB_t + u_t dt; \quad x_0 = 0.$$

We are able to provide explicit computations because of the manageable statistics of the model, the quadratic performance criterion as well as the special initial condition. Consider the problem of using zero order hold control to keep the state as close to the origin as possible with exactly one switch allowed. Because of the Markov property of diffusions, the control levels depend only on the last received sample. Then, the control waveform has the precise form:

$$u_t = \begin{cases} U_0 = U_0(x_0, T) & \text{if } 0 \leq t < \tau, \\ U_1 = U_1(x_\tau, T - \tau) & \text{if } \tau \leq t \leq T, \end{cases}$$

where, the switch time τ is a stopping time [9] w.r.t. the x -process, one which is restricted to fall inside the interval $[0, T]$. The mechanism that determines this control waveform comprises of policies which determine the combination of variables U_0, U_1 , and τ .

We will choose the control objective function introduced in Equation (2) to be quadratic in the state, and hence,

$$J = \mathbb{E} \left[\int_0^T x_s^2 ds \right]. \quad (3)$$

A. Choice of terminal control

We will now decompose this aggregate quadratic cost as follows:

$$\begin{aligned} J &= \mathbb{E} \left[\int_0^T x_s^2 ds \right], \\ &= \mathbb{E} \left[\int_0^\tau x_s^2 ds \right] + \mathbb{E} \left[\int_\tau^T x_s^2 ds \right]. \end{aligned}$$

The last term of the expression above represents the part of the control error incurred from the switch time τ up to the end time T , and, it can be expressed as a terminal cost for a combined optimal control and stopping problem. Let,

$$L(x_\tau, U_1(x_\tau, T - \tau), T - \tau) = \mathbb{E} \left[\int_\tau^T x_s^2 ds \middle| x_\tau, U_1, \tau \right].$$

We will now express this quantity as a terminal cost incurred upon stopping at τ . Adopting the notation $\delta := T - \tau$ for the duration of the terminal stage, and, suppressing some arguments of functions for the sake of convenience, we can write,

$$\begin{aligned} \mathbb{E}[L] &= \mathbb{E} \left[\mathbb{E} \left[\int_\tau^T x_s^2 ds \middle| \tau, x_\tau, U_1 \right] \right], \\ &= \mathbb{E} \left[x_\tau^2 \delta + U_1^2 \frac{\delta^3}{3} + x_\tau U_1 \delta^2 + \frac{\delta^2}{2} \right], \\ &= \mathbb{E} \left[\frac{1}{4} x_\tau^2 \delta + \delta \cdot \left(\frac{\delta U_1}{\sqrt{3}} + \frac{x_\tau \sqrt{3}}{2} \right)^2 + \frac{\delta^2}{2} \right]. \end{aligned}$$

In the last expression above, we have completed squares and this allows us to see that, the optimal choice of terminal control level U_1^* is always the linear feedback law:

$$U_1^*(x_\tau, T - \tau) = -\frac{3x_\tau}{2(T - \tau)}. \quad (4)$$

The optimal choice of U_0 on the other hand, is closely tied with the the choice of τ . In fact, for the best control performance, these two variables have to be picked jointly. In view of the expression for optimal U_1 given by equation (4), minimizing the quadratic objective function (3) is now achieved by minimizing the objective function

$$J_0 = \mathbb{E} \left[\int_0^\tau x_s^2 ds \right] + \mathbb{E} \left[\frac{1}{4} x_\tau^2 (T - \tau) + \frac{(T - \tau)^2}{2} \right]. \quad (5)$$

Next, we will describe the optimal combined choice if the switch time is restricted to be deterministic.

B. Optimal deterministic switching

Recall that the optimal choice of τ is event-triggered and hence causally dependent on the x -process. However, in this part of the section, we will present an easy calculation of the optimal deterministic choice of switching time. We will use the computed performance for comparisons with the event-triggered counterpart.

Denote the candidate deterministic switch time by θ . Like the random variable τ , the time θ is required to fall within the interval $[0, T]$.

We know now that the optimal choice of U_1^* , is the linear law given by equation (4). What remains of the optimal control problem is to choose the pair (U_0, θ) to minimize the objective function J_0 . Taking into account the zero initial condition, Equation (5) can be rewritten as follows:

$$J_0 = U_0^2 \frac{\theta^3}{3} + \frac{\theta^2}{2} + (U_0^2 \theta^2 + \theta) \frac{T - \theta}{4} + \frac{(T - \theta)^2}{2}.$$

Since the variables θ , $T - \theta$, and, U_0^2 are all nonnegative, every term on the RHS of the last equation is non-negative. This means that for any choice of θ , the optimal choice of U_0 is zero. This choice gives the following quadratic form of the performance:

$$J_0(U_0 = 0, \theta) = \frac{\theta^2 + (T - \theta)^2}{2} + \frac{\theta(T - \theta)}{4},$$

from which, we can see that the optimal deterministic switch time θ^* is $T/2$. The corresponding control performance cost is $\frac{2.5T^2}{8}$.

C. Optimum switching

Now, we will compute the optimal choice of the pair (U_0, τ) , one which will involve an event-triggered choice for τ . We will also see that the optimal choice of U_0 will be zero just like in the case of deterministic switch times.

The optimization problem has no restriction on τ except the requirement of causal dependence on the state trajectory. It is a combined optimal control and stopping problem which seeks an admissible pair (U_0, τ) that minimizes the cost:

$$\begin{aligned} J_0 &= \mathbb{E} \left[\int_0^\tau x_s^2 ds + L(x_\tau, U_1^*(x_\tau, T - \tau), T - \tau) \right], \\ &= \mathbb{E} \left[\int_0^\tau x_s^2 ds \right] + \mathbb{E} \left[\frac{1}{4} x_\tau^2 (T - \tau) + \frac{(T - \tau)^2}{2} \right], \\ &= \frac{T^2}{2} + \frac{U_0^2 T^3}{3} - \mathbb{E} \left[\left(\frac{x_\tau \sqrt{3}}{2} + \frac{U_0 (T - \tau)}{\sqrt{3}} \right)^2 (T - \tau) \right]. \end{aligned}$$

This cost captures the aggregate squared deviation of the state process away from the origin. When the initial state is exactly zero, it makes intuitive sense to set the initial control level U_0 to be zero. This is because the standard Brownian motion process is a martingale with zero mean; it is costly to steer the mean of the process using nonzero U_0 , away from the origin.

We will now sketch a more rigorous way of showing that for a zero initial condition, the optimal pair of the

initial control level and the switch time is such that the initial control level is zero. Firstly we need the uniqueness of optimal U_0 . This result can actually be obtained from the considerations of the next section. But we can also view the optimization problem as a nested sequence of an optimal stopping problem and an optimal control problem which involves only U_0 . Essentially because of the well-behavedness of the quadratic cost function and the tame dynamics of the Brownian motion process, we have a unique solution to the optimal control problem.

On the other hand, if (U_0^*, τ) is an optimal choice, then the pair $(-U_0^*, \tau)$ provides exactly the same performance because the x -process possesses even-symmetry. However, if U_0^* is unique, then $U_0^* = -U_0^*$, which leads us to the conclusion that $U_0^* = 0$.

This further reduces the expression for the objective function as below:

$$J_0 = \frac{T^2}{2} - \frac{3}{4} \mathbb{E} [x_\tau^2 (T - \tau)].$$

Thus we have reduced the optimization task to a stopping problem with a time-varying terminal reward. This problem can be solved explicitly using standard methods of optimal stopping as has been carried out for some closely related estimation problems [11], [10].

D. An explicit solution for the Optimal stopping problem

We seek an admissible stopping time τ that maximizes the expected reward function:

$$\mathbb{E} [x_\tau^2 (T - \tau)]. \quad (6)$$

Suppose that we can find a smooth function $g(x, t)$ for which

$$\begin{cases} g(x, t) & \geq x^2(T - t), \\ \frac{1}{2}g_{xx} + g_t & = 0, \end{cases} \quad (7)$$

where, the inequality is tight in the sense that for every $T-t$, there exists at least one x for which the inequality becomes an equality. Then, by the Itô change of variable formula [9], we have that for every stopping time τ with bounded expectation and satisfying

$$0 \leq t \leq \tau \leq T,$$

we get,

$$\begin{aligned} \mathbb{E} [x_\tau^2 (T - \tau) | \mathcal{F}_t^x] & \leq \mathbb{E} [g(x_\tau, \tau) | \mathcal{F}_t^x], \\ & = g(x_t, t) + \mathbb{E} \left[\int_t^\tau dg(x_\tau, \tau) \right], \\ & = g(x_t, t) + \mathbb{E} \left[\int_t^\tau \left\{ \frac{1}{2}g_{xx} + g_t \right\} dt \right], \\ & = g(x_t, t). \end{aligned}$$

Thus at any time t on the interval $[0, T]$, we have $g(x_t, t)$ as an upper bound for the expected reward for stopping at a later time. Moreover, as we will see shortly, this bound is achievable so that it is also the maximum achievable expected

reward. This process is called the *Snell envelope* (S_t) for the optimal stopping problem:

$$S_t = g(x_t, t).$$

The inequality in (7) connecting the instantaneous reward function and the function g is tight. Hence, there exists a stopping time with bounded expectation which will achieve this upper bound for the maximum expected reward. It is precisely the earliest instant when the reward collected by stopping equals the Snell envelope process:

$$\tau^* = \inf \left\{ t \leq \nu \leq T \mid g(x_\nu, \nu) = x_\nu^2 (T - \nu) \right\}.$$

We will now furnish an explicit solution to the sufficiency conditions (7). Consider as the candidate solution, the fourth Hermite polynomial [9] of the ratio $\frac{x_t}{\sqrt{\delta}}$, multiplied by the scaling factor $A\delta^2$. Hermite polynomials are solutions to the PDE in the equation group (7). So, we have:

$$g(x_t, t) = A\delta^2 \left(\frac{1}{6} \left(\frac{x_t}{\sqrt{\delta}} \right)^4 + \left(\frac{x_t}{\sqrt{\delta}} \right)^2 + \frac{1}{2} \right).$$

This provides the following expression for the Snell envelope:

$$S_t = A \left(\frac{x_t^4}{6} + x_t^2 (T - t) + \frac{(T - t)^2}{2} \right).$$

We require the inequality:

$$S_t = g(x_t, t) \geq x_t^2 (T - t),$$

to be tight. Indeed the choice $A = \frac{\sqrt{3}}{1+\sqrt{3}}$ makes the quantity $h = g - x^2(T - t)$ a perfect square, thus providing the desired tightness:

$$h(x_t, t) = \frac{1}{2(1+\sqrt{3})} \cdot \left(\frac{x_t^4}{3} - 2\frac{x_t^2}{\sqrt{3}}(T - t) + (T - t)^2 \right)$$

Thus the optimal stopping rule is the symmetric quadratic envelope:

$$\tau^* = \inf \left\{ t \mid x_t^2 \geq \sqrt{3}(T - t) \right\}.$$

The expected control performance cost incurred by the optimal switching scheme is then $T^2/8$, which is nearly a third of the the cost of using deterministic switching.

III. NON-ZERO INITIAL CONDITION

In the preceding section, under the assumption of starting the state process at the origin, we were able to present explicit calculations of the performance the optimal switching and control strategies. We were also able to describe analytically the containment envelope for the state process used to determine the switching time.

In this section, we will relax the requirement that the initial condition should be zero and otherwise attack the same optimization problem. However, there is no longer the symmetry effected by starting at the origin. We will see that,

like in the previous case, the optimal switching time is the first exit time from an envelope for the x -process. We will be able to determine a structural property of this envelope - at all times, its two arms are on different sides of the origin. But the obtained description of the envelope itself is not analytic. It is computed numerically using backward dynamic programming.

The optimal value of the initial control level is non-trivial to determine. The provision of optimal switch times still a step forward in itself, because, one could use a simple linear rule or another heuristic for the choice of initial control level.

A. Optimum deterministic switching

First we restrict the switch time to be deterministic. Recall that the best choice of U_1 is still provided by equation (4). This leads us to pick the pair (U_0, θ) that minimizes the cost:

$$J_0 = x_0^2 T + x_0 U_0 T^2 + \frac{U_0^2 T^3}{3} + \frac{T^2}{2} - \mathbb{E} \left[\left(\frac{x_\theta \sqrt{3}}{2} + \frac{U_0 (T - \theta)}{\sqrt{3}} \right)^2 (T - \theta) \right].$$

As an optimization problem this is quite straightforward. However, the solution for the optimal control strategy is not linear. First, we freeze U_0 and pick the best sampling time θ . This choice can be made by minimizing a cubic polynomial in θ . Equating the derivative of this cubic polynomial to zero, we get a quadratic equation whose roots if real and within the interval $[0, T]$ are also candidates for the optimal θ . However, the end points of the intervals themselves are candidates for the optimal choice of θ . In computations with MATLAB, a lot of effort is saved by actually performing a line search along the interval $[0, T]$ for the optimum θ .

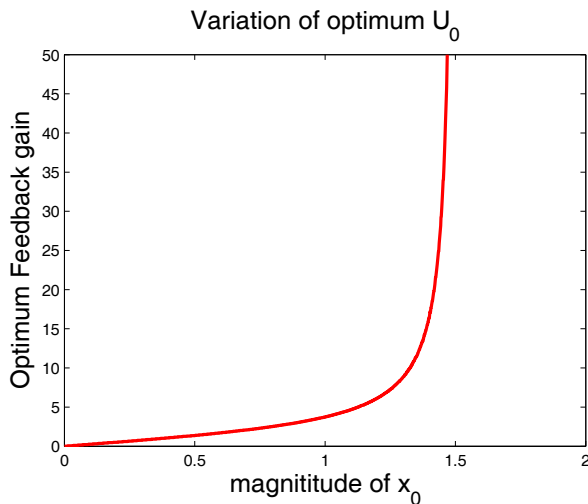


Fig. 4. Optimal choice of initial control level as a function of the initial condition.

For choosing the best values of U_0 corresponding to x_0 , we perform another line search. The outcome of this computation is depicted in figure (4). Notice that the law for U_0^* is not linear ! In fact, the value of U_0^* rapidly escapes to

infinity. The reason is that the objective function penalizes the magnitude of the state process but does not directly prohibit large amplitudes of the control signal. Hence, if the initial condition is too large, the efficient control and sampling strategy is to use a large amplitude control signal for a very short time. This brings the state close to zero very quickly. In fact when a large amplitude control waveform is used, to prevent adverse effects of overshoot on the other side of the origin, it must be used only for a relatively short period. This is indeed the idea that figures (4, 5) convey. This also affects the dependence of the performance on the initial condition as shown in figure (6).

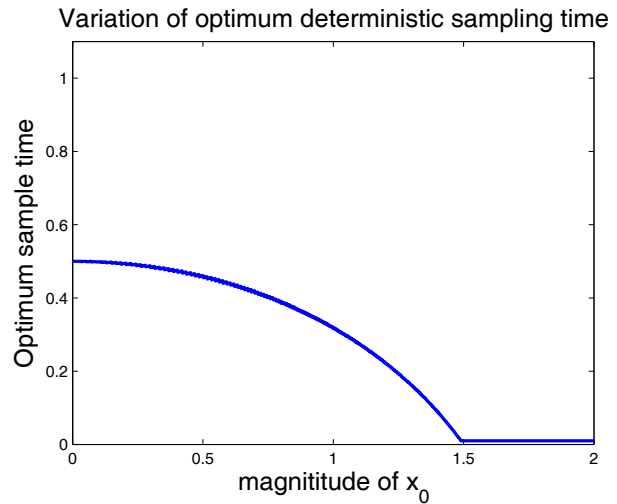


Fig. 5. Optimal choice of deterministic switch time as a function of the initial condition.

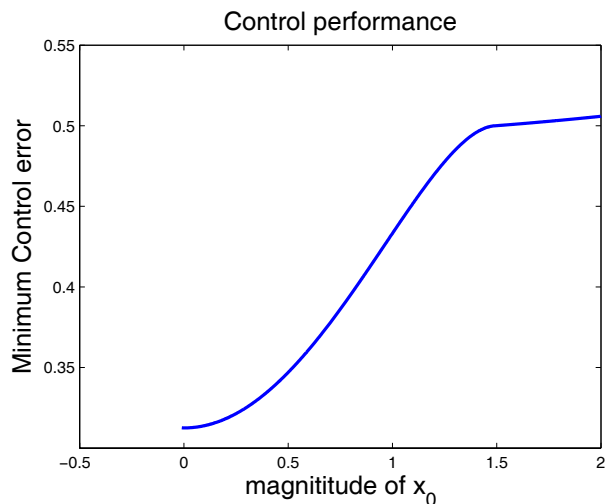


Fig. 6. Optimal choice of deterministic switch time as a function of the initial condition.

B. Optimum sampling

Now we turn our attention to the problem of optimal switching of the control signal. For a fixed nonzero U_0 , we

seek a stopping time which minimizes the expected cost:

$$\mathbb{E} \left[\left(x_\tau + \frac{2U_0(T-\tau)}{3} \right)^2 (T-\tau) \right].$$

As before we seek a sufficiency condition for the Snell envelope process. We are thus seeking a smooth function g such that

$$\begin{cases} g(x, t) & \geq \left(x_\tau + \frac{2U_0(T-\tau)}{3} \right)^2 (T-\tau), \\ \frac{1}{2}g_{xx} + g_t & = 0, \end{cases} \quad (8)$$

where, we require the inequality to be tight. This validity of this set of conditions for sufficiency can be proved in exactly the same manner as for those of (7).

To find candidate Snell envelopes, we need to resort to backward dynamic programming. Because of the Markov property of the controlled process, we obtain an optimal stopping time which is a first exit time from an envelope for the x -process. This envelope represents a time-varying set of thresholds for triggering the sensor sampling. We carried out the backward dynamic program numerically. In figure (7), we

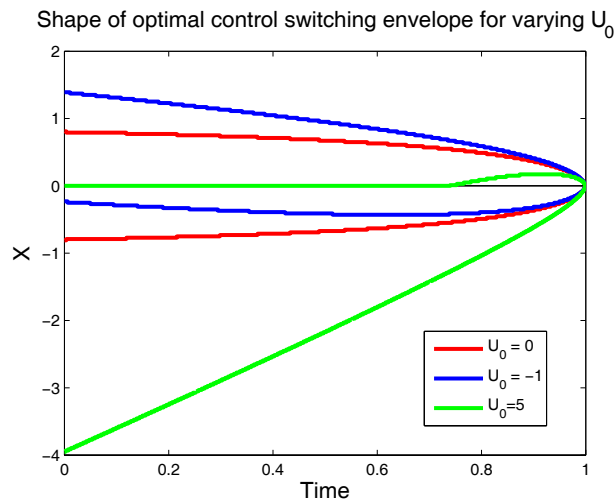


Fig. 7. Optimal time-varying thresholds for the triggering of the switch time for different values of U_0 .

see some shapes for the optimal envelopes. The results of numerical computations show some notable features. Firstly, for very large values of the initial control level, the optimal envelope becomes a triangular wedge with the time axis ($x = 0$) being one of the boundaries of the wedge. Secondly, the two branches of the envelope are always on either side of the origin, and lastly, we want to point out that neither of the branches has to be a monotonic function of time; the lower branch of the envelope corresponding to $U_0 = 0$ increases first and then decreases.

The brute force search of the optimal initial control given the initial condition is numerically intensive. We are currently searching for an efficient solution for it.

IV. CONCLUSIONS

We have formulated the problem of dynamic sensor sampling and choice of switching control as a combined optimal control and stopping problem. In a rather special case, we have furnished a complete solution. When we generalize the situation further, even the deterministic sampling problem displays interesting nonlinear solutions. The optimal sensor sampling time is a first crossing time of a time-varying envelope. The event-triggered sampling and control strategies we have designed offer promising efficiency gains over their counterparts with deterministic scheduling.

We should firstly complete the design process by devising an efficient procedure for finding optimal U_0 . Then the extension to multiple switches and to other models for the state should be straightforward. We should also investigate the situations where piecewise deterministic waveforms are allowed for the control signal instead of merely piece-wise constant ones.

REFERENCES

- [1] Michel Adès, Peter E. Caines, and Roland P. Malhamé. Stochastic optimal control under Poisson-distributed observations. *IEEE Trans. Automat. Control*, 45(1):3–13, 2000.
- [2] Karl Johan Åström and Bo Bernhardsson. Comparison of Riemann and Lebesgue sampling for first order stochastic systems. In *Proceedings of the 41st IEEE conference on Decision and Control (Las Vegas NV, 2002)*, pages 2011–2016. IEEE Control Systems Society, 2002.
- [3] Wendell H. Fleming and Raymond W. Rishel. *Deterministic and stochastic optimal control*. Springer-Verlag, Berlin, 1975.
- [4] Dimitris Hristu-Varsakelis and Panganamala R. Kumar. Interrupt-based feedback control over a shared communication medium. In *Proceedings of the 41st IEEE conference on Decision and Control and European Control Conference (Las Vegas, 2002)*, pages 3223–3228. IEEE Control Systems Society, 2002.
- [5] Ioannis Karatzas and Hui Wang. Utility maximization with discretionary stopping. *SIAM J. Control Optim.*, 39(1):306–329 (electronic), 2000.
- [6] Ioannis Karatzas and Ingrid-Mona Zamfirescu. Martingale approach to stochastic control with discretionary stopping. *Appl. Math. Optim.*, 53(2):163–184, 2006.
- [7] Harold J. Kushner. On the optimum timing of observations for linear control systems with unknown initial state. *IEEE Trans. Automatic Control*, AC-9:144–150, 1964.
- [8] Hiroaki Morimoto. Variational inequalities for combined control and stopping. *SIAM J. Control Optim.*, 42(2):686–708 (electronic), 2003.
- [9] Bernt Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.
- [10] Maben Rabi. *Packet based Inference and Control*. PhD thesis, University of Maryland, College Park, September 2006.
- [11] Maben Rabi, George V. Moustakides, and John S. Baras. Multiple sampling for estimation on a finite horizon. In *Proceedings of the 45th IEEE conference on Decision and Control (San Diego, CA, 2006)*, pages 1351–1357. IEEE Control Systems Society, 2006.