Optimal Control with Limited Controls

Orhan C. Imer and Tamer Başar

***Abstract*— We consider a linear discrete-time optimal control problem where the control is limited in terms of the number of times it can be applied. We assume an additive quadratic performance criterion that does not penalize the control directly, and show that for a scalar plant driven by an independent additive Gaussian noise the optimal control is piecewise linear based on some offline computed thresholds on the optimal estimate of the plant state which can be generated by a Kalman filter.**

I. Introduction

Optimal quadratic control of discrete-time linear systems has been extensively studied in the literature [1]. In the standard formulation of the discrete-time linear quadratic Gaussian (LQG) problem, the controller is assumed to be *unlimited* in its actions in terms of the number of times it can be applied. This assumption, though valid in many scenarios, is not an accurate model in other scenarios where each application of control but not the actual value of the control action is expensive. For example, in most wireless control applications the controller is power-limited, and therefore can only transmit a *limited* number of signals to the actuator [2]. Limited control action may also be justified in the control of some macro economic quantities, such as exchange rates, or in controlling inventory levels at a warehouse. Motivated by these observations, we formulate an optimal control problem with limited controls where over a decision horizon of length N, the controller has M time units in which it can act. We assume a linear discrete-time plant model driven by an additive Gaussian noise. The controller has access to some noisy version of the plant state at all decision instances, but it is limited in its actions. In this paper, we show that for an additive quadratic performance criterion that does not penalize the control directly, the optimal control law is piecewise linear based on some off-line computed thresholds on the optimal estimate of the plant state. We also show that the separation of estimation and control holds [3], and therefore the optimal estimate of the plant state can be recursively generated by a Kalman filter.

The rest of the paper is organized as follows. In Section II, we formally define the problem, briefly discuss some of the potential applications, and establish the difference between the open-loop and closed-loop scheduling policies. The solution is derived in Section III using a dynamic programming type argument. We present some numerical solutions in Section IV.

First author is with GE Global Research, One Research Circle, Niskayuna, NY 12309. Second author is with Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, 1308 W. Main Street, Urbana, IL 61801. imer@research.ge.com, tbasar@control.csl.uiuc.edu

Research supported by the NSF grant NSF CCR 00-85917

II. Problem Statement

A. Problem Definition

Consider the scalar plant described by the discrete-time dynamics[1]

$$x_{k+1} = Ax_k + u_k + w_k,\ k = 0,1,\ldots,N-1 \qquad (1)$$

where $x_k \in \mathbf{R}$ is the state, $u_k \in \mathbf{R}$ is the control, and $w_k \in \mathbf{R}$ is a zero-mean i.i.d. Gaussian process with finite variance, σ_w^2, describing the plant noise. In (1), N denotes the decision horizon, and the initial state x_0 is characterized by its probability distribution P_{x_0}. The controller receives at the beginning of each period k an observation of the form:

$$y_k = x_k + v_k,\ k = 0,1,\ldots,N-1$$

where $v_k \in \mathbf{R}$ is the observation noise modelled as an i.i.d. Gaussian process with zero mean, and variance $\sigma_v^2 < \infty$. The noise processes $\{v_k\}$ and $\{w_k\}$, and the initial state x_0 are assumed to be independent. Let I_k denote the information available to the controller at time k. We have $I_0 = y_0$, and $I_k = \left\{y_0^k, u_0^{k-1}\right\},\ k = 1,2,\ldots,N-1$. Consider the class of policies consisting of a sequence of functions $\pi = \{\mu_0, \mu_1, \ldots, \mu_{N-1}\}$, where each function maps the information vector I_k into the control space C_k. The sets C_k are restricted such that the control can map I_k into $C_k = \mathbf{R}$ only a limited number of times. For all other k, C_k is the singleton $C_k = \{0\}$. Such control policies are called *admissible*. We want to find an admissible policy π that minimizes the performance criterion

$$J_\pi = E\left\{x_N^2 + \sum_{k=0}^{N-1} x_k^2\right\}$$

subject to the system equation (1). Note that the restriction on control sets is equivalent to forcing the control to be zero for those times during which the control is not allowed to act. We assume that the control is mapped into the real line $M \leq N$ times. Also, we do not include a direct penalty for control in the cost function, but note that there is an indirect penalty on control in terms of a limitation in the number of times it can act.

B. Example Applications

There are several instances where the nature of the control action is limited in terms of the number of times it can be exercised. In a networked control systems setting, this may be due to a power constraint that limits the number of available transmission opportunities for the wireless controller device

[1] As we will see shortly, considering a scalar system does not lead to much loss of conceptual generality.

for a given decision horizon, see Figure 1. This is similar to the case in [4], [5], where the wireless sensor is limited in power and in turn in the number of transmissions it can make.

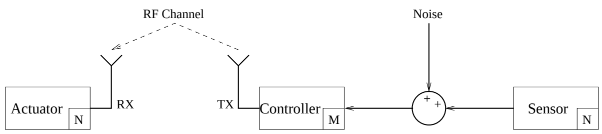

Fig. 1. Optimal control over an RF channel with limited controls.

Limited control action may also be justified in the control of some macro economic quantities, or in controlling inventory levels at a warehouse. For instance, when regulating the value of a financial instrument, an institution, such as a central bank, may only act a limited number of times for a given decision horizon. The financial instrument that we are trying to regulate might be the exchange rate of a particular currency, and we may want to stabilize this rate around an equilibrium value for a certain period of time. The regulatory commission that is in charge of this stabilization may meet regularly, but due to several reasons, it may only act in a limited number of those meetings. In inventory control, the stock level of a particular merchandise needs to be regulated at some desirable level. The inventory manager who is in charge of keeping the stock at the desired level may be limited in terms of the number of orders it can make to restock the warehouse.

C. Open-loop versus Closed-loop Scheduling Policies

In this section, we would like to draw attention to the difference between open-loop and closed-loop control scheduling policies. Note that for a given pair of numbers (M,N), if want to determine M out of N times during which we should apply control *a priori*, the best times to apply control would be the first M time units, and the control during those times would be of the form

$$u_k = \begin{cases} 0 & \text{if} \quad M \le k \le N-1 \\ -AE\{x_k|I_k\} & \text{if} \quad 0 \le k \le M-1 \end{cases}$$

where the conditional expectation of the state can be recursively generated by a Kalman filter. The control is linear in the conditional expectation of the state, because the cost function is quadratic in x_k, and there is no direct cost on control. Now it can be verified that among all open-loop schedules for the controller, the one that leads to the smallest average cost is the one where we apply control during the first M time units. The term *open-loop* refers to the determination of the controller schedule as it relates to what time units it should act. When it does act, the control action itself depends on the current estimate of the state in a closed-loop fashion. In the next section, we determine both the schedule and the actions of the controller in a closed-loop fashion.

III. Derivation of the Solution

In order to derive the solution we let s denote the number of control actions left, and t denote the number of decision instances left. Given M and N, going backward in time, t increases from $t=1$ to $t=N$, while s takes values on the interval $\max\{0, M-(N-t)\} \le s \le \min\{t, M\}$. Thus, for a given t, N, and M such that $1 \le M \le N$, the maximal interval in which s can take values is given by $0 \le s \le t$. We derive the solution by a dynamic programming argument starting with $t=1$, and going back in time, or forward in t to $t=N$. For each t, we consider the potential values s can take. Note that given that we are at the decision stage (s,t), we may go from here to either stage $(s-1,t-1)$ or $(s,t-1)$ depending on whether we decide to act at stage (s,t) or not. By a similar argument, we can see that we must have arrived at stage (s,t) either from stage $(s,t+1)$, or $(s+1,t+1)$. Now, starting with $t=1$, we see that $0 \le s \le 1$. When $s=0$, we must have $u_{(0,1)}=0$, as $(0,1)$ can only lead to $(0,0)$. We can calculate the optimal cost to go from stage $(0,1)$ as $J_{(0,1)} = K_{(0,1)}E\{x_{N-1}^2|y_0^{N-1}\} + \sigma_w^2$, where $K_{(0,1)} = 1+A^2$. When $s=1$, $(1,1)$ can only lead to $(0,0)$ with the control $u_{(1,1)} = -AE\{x_{N-1}|y_0^{N-1}\}$, and the associated cost-to-go

$$\begin{aligned} J_{(1,1)} \;&=\; E\{x_{N-1}^2|y_0^{N-1}\} \\ &\quad +A^2E\{\left(x_{N-1}-E\{x_{N-1}|y_0^{N-1}\}\right)^2|y_0^{N-1}\} \\ &\quad +\sigma_w^2 + A^2E\{\left(x_{N-1}-E\{x_{N-1}|y_0^{N-1}\}\right)^2|y_0^{N-1}\} \end{aligned}$$

Before proceeding with the next stage, observe that the term $E\{\left(x_{N-1}-E\{x_{N-1}|y_0^{N-1}\}\right)^2|y_0^{N-1}\}$ is independent of past controls. This is due to the linearity of both the system and measurement equation. We state this fact as a lemma whose proof can be found in [4].

Lemma 1: For every $k \in [0,N-1]$, $x_k - E\{x_k|I_k\}$ is independent of the control policy being used.

Therefore, we can write the cost to go from $(1,1)$ as

$$J_{(1,1)} = E\{x_{N-1}^2|y_0^{N-1}\} + A^2\sigma_{N-1|N-1}^2 + \sigma_w^2$$

where the error covariance $\sigma_{N-1|N-1}^2$ is defined by

$$\sigma_{k|k-1}^2 = E\{(x_k - E\{x_k|y_0^{k-1}\})^2|y_0^{k-1}\}$$

We next let $t=2$, which implies that $0 \le s \le 2$. When $s=0$, $(0,2)$ can only lead to $(0,1)$, resulting in the control $u_{(0,2)} = 0$, and the associated cost

$$J_{(0,2)} = K_{(0,2)}E\{x_{N-2}^2|y_0^{N-2}\} + K_{(0,1)}\sigma_w^2 + \sigma_w^2$$

where $K_{(0,2)} = 1 + A^2K_{(0,1)}$. When $s=1$, on the other hand, $(1,2)$ may lead to $(1,1)$ or $(0,1)$ depending on whether we apply a control or not. The optimal policy when we do apply control can be found by minimizing the quadratic cost-to-go function and is still linear. We have $u_{(1,2)}^{(0)} = 0$ and $u_{(1,2)}^{(1)} = -AE\{x_{N-2}|y_0^{N-2}\}$. Here the superscripts (1) and (0) denote whether a control action is taken or not. Substituting these controls into the cost function, we obtain the cost to go from stage $(1,2)$ for both cases as follows:

$$\begin{aligned} J_{(1,2)}^{(0)} \;&=\; (1+A^2)E\{x_{N-2}^2|y_0^{N-2}\} + A^2\sigma_{N-1|N-1}^2 + \sum_{n=0}^{1}\sigma_w^2 \\ J_{(1,2)}^{(1)} \;&=\; E\{x_{N-2}^2|y_0^{N-2}\} + K_{(0,1)}A^2\sigma_{N-2|N-2}^2 + K_{(0,1)}\sigma_w^2 \\ &\quad +\sigma_w^2 \end{aligned}$$

Thus, the decision as to whether to control or not depends on the comparison between the cost functions $J^{(0)}_{(1,2)}$ and $J^{(1)}_{(1,2)}$. We look at the difference $J^{(0)}_{(1,2)} - J^{(1)}_{(1,2)}$, which needs to be compared against zero to make the decision as to whether to control or not. Assuming $A \neq 0$, this comparison is equivalent to comparing the following quantity against zero:

$$E\{x^2_{N-2}|y^{N-2}_0\} + \sigma^2_{N-1|N-1} - (1+A^2)\sigma^2_{N-2|N-2} - \sigma^2_w \gtrless 0$$

Now, let $\Delta_{(1,2)}$ denote this difference, i.e., $\Delta_{(1,2)} = J^{(0)}_{(1,2)} - J^{(1)}_{(1,2)}$. We have

$$\begin{aligned}\Delta_{(1,2)} &= A^2 \Big(E\{x^2_{N-2}|y^{N-2}_0\} + \sigma^2_{N-1|N-1} \\ &\quad -(1+A^2)\sigma^2_{N-2|N-2} - \sigma^2_w \Big)\end{aligned}$$

Note that the conditional expectation $E\{x^2_{N-2}|y^{N-2}_0\}$ can be expressed as $E\{x^2_{N-2}|y^{N-2}_0\} = \sigma^2_{N-2|N-2} + \left(E\{x_{N-2}|y^{N-2}_0\}\right)^2$. Substituting this into the expression for $\Delta_{(1,2)}$ yields

$$\Delta_{(1,2)} = A^2 \left(\left(E\{x_{N-2}|y^{N-2}_0\}\right)^2 + \sigma^2_{N-1|N-1} - \sigma^2_{N-1|N-2} \right)$$

Therefore, the decision at $(1,2)$ to control or not is only a function of the optimal estimate of the plant state which in turn can be expressed as a function of the current measurement y_{N-2}, the previous control u_{N-3}, and the previous estimate $E\{x_{N-3}|y^{N-3}_0\}$ through the Kalman filter recursion:

$$E\{x_{N-2}|y^{N-2}_0\} = AE\{x_{N-3}|y^{N-3}_0\} + u_{N-3}$$

Note that, as a function of the conditional estimate $E\{x_{N-2}|y^{N-2}_0\}$, $\Delta_{(1,2)}\left(E\{x_{N-2}|y^{N-2}_0\}\right)$ has a unique minimum $\Delta^{\min}_{(1,2)} = A^2\left(\sigma^2_{N-1|N-1} - \sigma^2_{N-1|N-2}\right) < 0$, which occurs at $E\{x_{N-2}|y^{N-2}_0\} = 0$. Thus, the equation $\Delta_{(1,2)}\left(E\{x_{N-2}|y^{N-2}_0\}\right) = 0$ has two distinct real roots $\tau^+_{(1,2)} = -\tau^-_{(1,2)} = \sqrt{\sigma^2_{N-1|N-2} - \sigma^2_{N-1|N-1}}$. That is, $\Delta_{(1,2)}(\tau^+_{(1,2)}) = \Delta_{(1,2)}(\tau^-_{(1,2)}) = 0$. The numbers $\tau^-_{(1,2)} < \tau^+_{(1,2)}$ are the thresholds on the current estimate of the state, $E\{x_{N-2}|y^{N-2}_0\}$, and we do not apply control if $\left|E\{x_{N-2}|y^{N-2}_0\}\right| \leq \tau^+_{(1,2)}$. Now, the optimal cost-to-go function from stage $(1,2)$ can be written as

$$J_{(1,2)} = \begin{cases} J^{(0)}_{(1,2)} & \text{if} \quad \tau^-_{(1,2)} \leq E\{x_{N-2}|y^{N-2}_0\} \leq \tau^+_{(1,2)} \\ J^{(1)}_{(1,2)} & \text{otherwise} \end{cases}$$

Finally, if $s=2$, $(2,2)$ can only lead to $(1,1)$, with the control law $u_{(2,2)} = -AE\{x_{N-2}|y^{N-2}_0\}$, and the associated cost to go

$$J_{(2,2)} = E\{x^2_{N-2}|y^{N-2}_0\} + A^2\sigma^2_{N-2|N-2} + A^2\sigma^2_{N-1|N-1} + \sum_{n=0}^{1}\sigma^2_w$$

We next let $t=3$, which implies that s must be in the range $0 \leq s \leq 3$. If $s=0$, $(0,3)$ can only lead to $(0,2)$; thus, we have the control $u_{(0,3)} = 0$, and the cost-to-go function

$$J_{(0,3)} = K_{(0,3)}E\{x^2_{N-3}|y^{N-3}_0\} + K_{(0,2)}\sigma^2_w + K_{(0,1)}\sigma^2_w + \sigma^2_w$$

where $K_{(0,3)} = 1 + A^2K_{(0,2)}$. If $s=1$, $(1,3)$ may lead to $(1,2)$ or $(0,2)$. If we do use control at stage $(1,3)$, the control must still minimize a quadratic cost function, leading to $u^{(1)}_{(1,3)} = -AE\{x_{N-3}|y^{N-3}_0\}$ with the associated cost

$$\begin{aligned}J^{(1)}_{(1,3)} &= E\{x^2_{N-3}|y^{N-3}_0\} + K_{(0,2)}A^2\sigma^2_{N-3|N-3} \\ &\quad + K_{(0,2)}\sigma^2_w + K_{(0,1)}\sigma^2_w + \sigma^2_w\end{aligned}$$

If, on the other hand, no control is used at stage $(1,3)$, we have $u^{(0)}_{(1,3)} = 0$. To calculate the cost-to-go from stage $(1,3)$ with zero control, we need to average the cost-to-go function $J_{(1,2)}$ over the statistics of $E\{x_{N-2}|y^{N-2}_0\}$ given y^{N-3}_0:

$$\begin{aligned}J^{(0)}_{(1,3)} &= E\{x^2_{N-3}|y^{N-3}_0\} \\ &\quad + \int_{|\hat{x}_{N-2|N-2}| \leq \tau^+_{(1,2)}} J^{(0)}_{(1,2)} f^{(0)}_{\hat{x}_{N-2|N-2}|y^{N-3}_0} d\hat{x}_{N-2|N-2} \\ &\quad + \int_{|\hat{x}_{N-2|N-2}| > \tau^+_{(1,2)}} J^{(1)}_{(1,2)} f^{(0)}_{\hat{x}_{N-2|N-2}|y^{N-3}_0} d\hat{x}_{N-2|N-2}\end{aligned}$$

where $f^{(0)}_{\hat{x}_{N-2|N-2}|y^{N-3}_0}$ is the conditional density function of $\hat{x}_{N-2|N-2}$ given the available information,[2] and $\hat{x}_{k|k-1}$ denotes $E\{x_k|y^{k-1}_0\}$. First, observe that given the past, i.e., y^{N-3}_0 and the past controls, the conditional estimate, $\hat{x}_{N-2|N-2}$ is Gaussian. We have

$$E\{\hat{x}_{N-2|N-2}|y^{N-3}_0\} = A\hat{x}_{N-3|N-3}$$

since $E\{y_{N-2}|y^{N-3}_0\} = A\hat{x}_{N-3|N-3}$. Also, we can write

$$E\{y^2_{N-2}|y^{N-3}_0\} = A^2\sigma^2_{N-3|N-3} + A^2\hat{x}^2_{N-3|N-3} + \sigma^2_w + \sigma^2_v$$

Now, for the conditional variance, we have

$$\begin{aligned}&E\{\left(\hat{x}_{N-2|N-2} - E\{\hat{x}_{N-2|N-2}|y^{N-3}_0\}\right)^2|y^{N-3}_0\} \\ &= E\{\left(\hat{x}_{N-2|N-2} - A\hat{x}_{N-3|N-3}\right)^2|y^{N-3}_0\}\end{aligned}$$

Writing out $\hat{x}_{N-2|N-2}$ from the Kalman filter recursion

$$\begin{aligned}&E\{\left(\hat{x}_{N-2|N-2} - A\hat{x}_{N-3|N-3}\right)^2|y^{N-3}_0\} \\ &= \frac{(\sigma^2_{N-2|N-3})^2}{(\sigma^2_{N-2|N-3}+\sigma^2_v)^2}E\{(y_{N-2} - A\hat{x}_{N-3|N-3}))^2|y^{N-3}_0\} \\ &= \frac{(\sigma^2_{N-2|N-3})^2}{(\sigma^2_{N-2|N-3}+\sigma^2_v)^2}(A^2\sigma^2_{N-3|N-3} + \sigma^2_w + \sigma^2_v) \\ &= \frac{(\sigma^2_{N-2|N-3})^2}{\sigma^2_{N-2|N-3}+\sigma^2_v}\end{aligned}$$

Therefore, $f_{\hat{x}_{N-2|N-2}|y^{N-3}_0} \sim N(A\hat{x}_{N-3|N-3}, \frac{(\sigma^2_{N-2|N-3})^2}{\sigma^2_{N-2|N-3}+\sigma^2_v})$. Rearranging, we can write the cost to go $J^{(0)}_{(1,3)}$ as

$$\begin{aligned}J^{(0)}_{(1,3)} &= (1+A^2)E\{x^2_{N-3}|y^{N-3}_0\} + K_{(0,1)}A^2\sigma^2_{N-2|N-2} \\ &\quad + K_{(0,1)}\sigma^2_w + \sum_{n=0}^{1}\sigma^2_w \\ &\quad + \int_{|\hat{x}_{N-2|N-2}| \leq \tau^+_{(1,2)}} \Delta_{(1,2)} f^{(0)}_{\hat{x}_{N-2|N-2}|y^{N-3}_0} d\hat{x}_{N-2|N-2}\end{aligned}$$

[2]Clearly, the available information I_k includes past actions. We do not write this explicitly for ease of notation.

Recall that as a function of $\hat{x}_{N-2|N-2}$, $\Delta_{(1,2)}(\hat{x}_{N-2|N-2})$ has a unique minimum at $\hat{x}_{N-2|N-2}=0$, with $\Delta_{(1,2)}(0)<0$. Now, to determine whether we should control or not at stage $(1,3)$, we compare the cost-to-go functions of the two alternatives $J^{(0)}_{(1,3)}$ and $J^{(1)}_{(1,3)}$. Similar to stage $(1,2)$, let us define $\Delta_{(1,3)}$ as $\Delta_{(1,3)} := J^{(0)}_{(1,3)} - J^{(1)}_{(1,3)}$. Substituting the expressions for $J^{(0)}_{(1,3)}$ and $J^{(1)}_{(1,3)}$ yields

$$\begin{aligned}\Delta_{(1,3)} &= A^2\Big(\Psi_{(1,3)}(\hat{x}_{N-3|N-3}) + \hat{x}^2_{N-3|N-3} \\ &\quad + K_{(0,1)}(\sigma^2_{N-2|N-2} - \sigma^2_{N-2|N-3})\Big)\end{aligned}$$

where $\Psi_{(1,3)}(\hat{x}_{N-3|N-3})$ is defined by the integral

$$\Psi_{(1,3)} := \frac{1}{A^2}\int_{|\hat{x}_{N-2|N-2}|\le\tau^+_{(1,2)}} \Delta_{(1,2)}\, f^{(0)}_{\hat{x}_{N-2|N-2}|y_0^{N-3}} d\hat{x}_{N-2|N-2}$$

We next state a property of $\Psi_{(1,3)}(\hat{x}_{N-3|N-3})$ whose proof can be found in [4].

Proposition 1: As a function of $\hat{x}_{N-3|N-3}$, $\Psi_{(1,3)}(\hat{x}_{N-3|N-3})$ is even, and has a unique minimum at $\hat{x}_{N-3|N-3}=0$ with a strictly negative value, i.e., $\Psi^{\min}_{(1,3)}(\hat{x}_{N-3|N-3}) = \Psi_{(1,3)}(0) < 0$.

Using Proposition 1, we can see that as a function of $\hat{x}_{N-3|N-3}$, $\Delta_{(1,3)}(\hat{x}_{N-3|N-3})$ is even, and has a unique minimum at $\hat{x}_{N-3|N-3}=0$. Furthermore, we have

$$\begin{aligned}\Delta^{\min}_{(1,3)}(\hat{x}_{N-3|N-3}) &= \Delta_{(1,3)}(0) \\ &= A^2K_{(0,1)}\Big(\frac{\Psi_{(1,3)}(0)}{K_{(0,1)}} + \sigma^2_{N-2|N-2} \\ &\quad - \sigma^2_{N-2|N-3}\Big) \\ &< 0\end{aligned}$$

Therefore, the equation $\Delta_{(1,3)}(\hat{x}_{N-3|N-3})=0$ has two distinct real roots $\tau^+_{(1,3)} = -\tau^-_{(1,3)} > 0$ with the property that $\Delta_{(1,3)}(\hat{x}_{N-3|N-3}) < 0$ if $|\hat{x}_{N-3|N-3}| < \tau^+_{(1,3)}$, and $\Delta_{(1,3)}(\hat{x}_{N-3|N-3}) \ge 0$, otherwise. The numbers $\tau^-_{(1,3)} < \tau^+_{(1,3)}$ are the thresholds on the current estimate of the state, $\hat{x}_{N-3|N-3}$, and we do not apply control if $|\hat{x}_{N-3|N-3}| \le \tau^+_{(1,3)}$. Now, the optimal cost-to-go function from stage $(1,3)$ can be written as

$$J_{(1,3)} = \begin{cases} J^{(0)}_{(1,3)} & \text{if } |\hat{x}_{N-3|N-3}| \le \tau^+_{(1,3)} \\ J^{(1)}_{(1,3)} & \text{if } |\hat{x}_{N-3|N-3}| > \tau^+_{(1,3)} \end{cases}$$

We next let $s=2$, and observe that from stage $(2,3)$ we can either go to stage $(1,2)$ or $(2,2)$ depending on whether we apply control or not. If we do not use control at stage $(2,3)$, we have $u^{(0)}_{(2,3)}=0$, which leads to the cost-to-go function

$$\begin{aligned}J^{(0)}_{(2,3)} &= (1+A^2)E\{x^2_{N-3}|y_0^{N-3}\} + A^2\sigma^2_{N-2|N-2} \\ &\quad + A^2\sigma^2_{N-1|N-1} + \sum_{n=0}^{2}\sigma^2_w\end{aligned}$$

If we do apply control at stage $(2,3)$, we must choose $u^{(1)}_{(2,3)}$ so as to minimize

$$\begin{aligned}J^{(1)}_{(2,3)} &= E\{x^2_{N-3}|y_0^{N-3}\} + K_{(0,1)}A^2\sigma^2_{N-2|N-2} \\ &\quad + K_{(0,1)}\sigma^2_w + \sum_{n=0}^{1}\sigma^2_w \\ &\quad + \min_{u^{(1)}_{(2,3)}}\Big\{E\Big\{(Ax_{N-3}+u^{(1)}_{(2,3)})^2|y_0^{N-3}\Big\} \\ &\qquad + \int_{|\hat{x}_{N-2|N-2}|\le\tau^+_{(1,2)}}\Delta_{(1,2)} \\ &\qquad\qquad \times f^{(1)}_{\hat{x}_{N-2|N-2}|y_0^{N-3}} d\hat{x}_{N-2|N-2}\Big\}\end{aligned}$$

where $f^{(1)}_{\hat{x}_{N-2|N-2}|y_0^{N-3}}$ is the conditional density function of $\hat{x}_{N-2|N-2}$ given the available information. We have

$$E\{\hat{x}_{N-2|N-2}|y_0^{N-3}\} = A\hat{x}_{N-3|N-3} + u^{(1)}_{(2,3)}$$

since $E\{y_{N-2}|y_0^{N-3}\} = A\hat{x}_{N-3|N-3} + u^{(1)}_{(2,3)}$. Also, we can write

$$\begin{aligned}E\{y^2_{N-2}|y_0^{N-3}\} &= A^2\sigma^2_{N-3|N-3} + A^2\hat{x}^2_{N-3|N-3} + (u^{(1)}_{(2,3)})^2 \\ &\quad + \sigma^2_w + \sigma^2_v\end{aligned}$$

Now, the conditional variance can be calculated as

$$\begin{aligned}&E\{\left(\hat{x}_{N-2|N-2} - E\{\hat{x}_{N-2|N-2}|y_0^{N-3}\}\right)^2|y_0^{N-3}\} \\ &= E\{(\hat{x}_{N-2|N-2} - A\hat{x}_{N-3|N-3} - u^{(1)}_{(2,3)})^2|y_0^{N-3}\}\end{aligned}$$

Writing out $\hat{x}_{N-2|N-2}$ from the Kalman filter recursion, we see that the variance is the same as before:

$$E\{(\hat{x}_{N-2|N-2} - A\hat{x}_{N-3|N-3} - u^{(1)}_{(2,3)})^2|y_0^{N-3}\} = \frac{(\sigma^2_{N-2|N-3})^2}{\sigma^2_{N-2|N-3}+\sigma^2_v}$$

Thus, $f^{(1)}_{\hat{x}_{N-2|N-2}|y_0^{N-3}} \sim N(A\hat{x}_{N-3|N-3} + u^{(1)}_{(2,3)}, \frac{(\sigma^2_{N-2|N-3})^2}{\sigma^2_{N-2|N-3}+\sigma^2_v})$. The optimal control $u^{(1)}_{(2,3)}$ can now be determined from the following minimization problem:

$$\begin{aligned}&\min_{u^{(1)}_{(2,3)}}\Big\{E\{(Ax_{N-3}+u^{(1)}_{(2,3)})^2|y_0^{N-3}\} \\ &\quad + \int_{|\hat{x}_{N-2|N-2}|\le\tau^+_{(1,2)}}\Delta_{(1,2)}\, f^{(1)}_{\hat{x}_{N-2|N-2}|y_0^{N-3}} d\hat{x}_{N-2|N-2}\Big\}\end{aligned}$$

Recall from Proposition 1 that the above integral is minimized when $\hat{x}_{N-2|N-2}=0$, which is equivalent to saying $A\hat{x}_{N-3|N-3} + u^{(1)}_{(2,3)} = 0$. Also note that the quadratic term inside the minimization is minimized at the exact same point. Hence, we conclude that the optimum control at stage $(2,3)$, when we choose to control, is $u^{(1)}_{(2,3)} = -A\hat{x}_{N-3|N-3}$. Plugging this into the cost-to-go function, we obtain $J^{(1)}_{(2,3)}$ as

$$\begin{aligned}J^{(1)}_{(2,3)} &= E\{x^2_{N-3}|y_0^{N-3}\} + A^2\sigma^2_{N-3|N-3} + K_{(0,1)}A^2\sigma^2_{N-2|N-2} \\ &\quad + K_{(0,1)}\sigma^2_w + \sum_{n=0}^{1}\sigma^2_w + A^2\Psi_{(1,3)}(0)\end{aligned}$$

Now, to determine whether we should control or not at stage $(2,3)$, we compare the cost-to-go functions of the two alternatives, $J^{(0)}_{(2,3)}$ and $J^{(1)}_{(2,3)}$. For this purpose, let us define $\Delta_{(2,3)} := J^{(0)}_{(2,3)} - J^{(1)}_{(2,3)}$. Substituting the expressions for $J^{(0)}_{(2,3)}$ and $J^{(1)}_{(2,3)}$ yields

$$\Delta_{(2,3)} = A^2(\hat{x}^2_{N-3|N-3} + \sigma^2_{N-1|N-1} - \sigma^2_{N-1|N-2} - \Psi_{(1,3)}(0))$$

From Proposition 1, we know that the function $\Psi_{(1,3)}(\cdot)$ is minimized at 0. We also know that the minimum value $\Psi_{(1,3)}$ takes at this point is negative. Now we go one step further, and show that this value is lower bounded by $\frac{1}{A^2}\Delta_{(1,2)}(0)$. To see this, recall that $\Delta_{(1,2)}(\hat{x}_{N-2|N-2})$ is minimized at $\hat{x}_{N-2|N-2} = 0$, and $\Delta_{(1,2)}(\hat{x}_{N-2|N-2}) \leq 0$, if $|\hat{x}_{N-2|N-2}| \leq \tau^+_{(1,2)}$. Therefore, we have

$$\begin{aligned}\Psi_{(1,3)}(0) &= \frac{1}{A^2}\int_{|\hat{x}_{N-2|N-2}|\leq\tau^+_{(1,2)}} \Delta_{(1,2)} \\ &\qquad \times f^{(0)}_{\hat{x}_{N-2|N-2}|y_0^{N-3}} d\hat{x}_{N-2|N-2} \\ &> \frac{1}{A^2}\Delta_{(1,2)}(0) = \sigma^2_{N-1|N-1} - \sigma^2_{N-1|N-2}\end{aligned}$$

since the integral of a probability measure is bounded above by 1. Using this fact, we see that as a function of $\hat{x}_{N-3|N-3}$, $\Delta_{(2,3)}(\hat{x}_{N-3|N-3})$ achieves its minimum at $\hat{x}_{N-3|N-3} = 0$ with

$$\begin{aligned}\Delta^{\min}_{(2,3)}(\hat{x}_{N-3|N-3}) &= \Delta_{(2,3)}(0) \\ &= A^2(\sigma^2_{N-1|N-1} - \sigma^2_{N-1|N-2} - \Psi_{(1,3)}(0)) \\ &< 0\end{aligned}$$

Since $\Delta_{(2,3)}(\hat{x}_{N-3|N-3})$ is an even function of $\hat{x}_{N-3|N-3}$, the equation $\Delta_{(2,3)}(\hat{x}_{N-3|N-3}) = 0$ has two distinct real roots

$$\tau^+_{(2,3)} = -\tau^-_{(2,3)} = \sqrt{\Psi_{(1,3)}(0) + \sigma^2_{N-1|N-2} - \sigma^2_{N-1|N-1}}$$

with the property that $\Delta_{(2,3)}(\hat{x}_{N-3|N-3}) < 0$ if $|\hat{x}_{N-3|N-3}| < \tau^+_{(2,3)}$, and $\Delta_{(2,3)}(\hat{x}_{N-3|N-3}) \geq 0$, otherwise. The numbers $\tau^-_{(2,3)} < \tau^+_{(2,3)}$ are the thresholds on the current estimate of the state, $\hat{x}_{N-3|N-3}$, and at stage $(2,3)$ we do not apply control if $|\hat{x}_{N-3|N-3}| \leq \tau^+_{(2,3)}$.

Finally, if $s = 3$, $(3,3)$ can only lead to $(2,2)$. Since $J_{(2,2)}$ is quadratic, the optimum choice for the control at stage $(3,3)$ is $u_{(3,3)} = -A\hat{x}_{N-3|N-3}$, which results in the cost-to-go

$$\begin{aligned}J_{(3,3)} &= E\{x^2_{N-3}|y_0^{N-3}\} + A^2\sigma^2_{N-3|N-3} + A^2\sigma^2_{N-2|N-2} \\ &\quad + A^2\sigma^2_{N-1|N-1} + \sum_{n=0}^{2}\sigma^2_w\end{aligned}$$

Proceeding in the same manner with $t = 4, 5, \ldots, N$, by induction we see that the optimal control policy is a threshold policy on the best estimate of the plant state which can be recursively generated by a Kalman filter. Furthermore, the threshold at time k is a function of four variables:

1) Length of the decision horizon: N
2) Number of decision instances left: t_k
3) Number of control actions left: s_k
4) Error covariance: $\sigma^2_{k|k-1}$, which can be calculated recursively starting with $\sigma^2_{0|-1} = E\{(x_0 - E\{x_0\})^2\}$, and iterating

$$\begin{aligned}\sigma^2_{k+1|k} &= A^2\sigma^2_{k|k} + \sigma^2_w \\ \sigma^2_{k|k} &= \sigma^2_{k|k-1} - \frac{(\sigma^2_{k|k-1})^2}{\sigma^2_{k|k-1} + \sigma^2_v}\end{aligned}$$

Note that, for a given N, the thresholds can be calculated *offline* using the procedure described in this section. The online computation, on the other hand, requires implementing a Kalman filter with the initial condition $\hat{x}_{0|-1} = E\{x_0\}$. Thus, starting with $\hat{x}_{0|-1} = E\{x_0\}$, $s_0 = M$, $t_0 = N$, the optimal control policy can be implemented by the following algorithm:

For each k in $0 \leq k \leq N-1$ do the following:

1) Look up the threshold $\tau^+_{(s_k,t_k)}$ corresponding to the current stage from the table.
2) Observe y_k and update the state estimate to $\hat{x}_{k|k}$ using the Kalman filter recursion:

$$\begin{aligned}\hat{x}_{k|k} &= \hat{x}_{k|k-1} + u_{k-1} + \frac{\sigma^2_{k|k-1}}{\sigma^2_{k|k-1} + \sigma^2_v}(y_k - \hat{x}_{k|k-1} - u_{k-1}) \\ \hat{x}_{k|k-1} &= A\hat{x}_{k-1|k-1}\end{aligned}$$

3) Apply the control policy

$$u_k = u_{(s_k,t_k)} = \begin{cases} 0 & \text{if } |\hat{x}_{k|k}| \leq \tau^+_{(s_k,t_k)} \\ -A\hat{x}_{k|k} & \text{if } |\hat{x}_{k|k}| > \tau^+_{(s_k,t_k)} \end{cases}$$

4) Update

$$\begin{aligned}s_{k+1} &= s_k - \mathscr{I}_{\{|\hat{x}_{k|k}|\leq\tau_{(s_k,t_k)}\}^c} \\ t_{k+1} &= t_k - 1\end{aligned}$$

where $\mathscr{I}_S$ denotes the indicator function of the set S.

We finally give the iterations that can be used to calculate the thresholds $\tau^+_{(s,t)}$ for a given decision horizon $N \geq 1$, and an arbitrary pair of integers (s,t) such that $1 \leq s \leq t \leq N$. Note that the cost-to-go functions $J^{(0)}_{(s,t)}$ and $J^{(1)}_{(s,t)}$ can be written as

$$\begin{aligned}J^{(0)}_{(s,t)} &= (1+A^2)E\{x^2_{N-t}|y_0^{N-t}\} + \sigma^2_w + \Lambda_{(s,t-1)} \\ &\quad + \int_{|\hat{x}_{N-t+1|N-t+1}|\leq\tau^+_{(s,t-1)}} \Delta_{(s,t-1)} \\ &\qquad \times f^{(0)}_{\hat{x}_{N-t+1|N-t+1}|y_0^{N-t}} d\hat{x}_{N-t+1|N-t+1} \\ J^{(1)}_{(s,t)} &= E\{x^2_{N-t}|y_0^{N-t}\} + \Lambda_{(s,t)}\end{aligned}$$

where $f^{(0)}_{\hat{x}_{N-t+1|N-t+1}|y_0^{N-t}} \sim N(A\hat{x}_{N-t|N-t}, \frac{(\sigma^2_{N-t+1|N-t})^2}{\sigma^2_{N-t+1|N-t}+\sigma^2_v})$, and for $1 < s < t \leq N$, $\Lambda_{(s,t)}$ is defined by recursion

$$\begin{aligned}\Lambda_{(s,t)} &= \Lambda_{(s-1,t-1)} + A^2\sigma^2_{N-t|N-t} + \sigma^2_w \\ &\quad + \int_{|\hat{x}_{N-t+1|N-t+1}|\leq\tau^+_{(s-1,t-1)}} \Delta_{(s-1,t-1)} \\ &\qquad \times f^{(1)}_{\hat{x}_{N-t+1|N-t+1}|y_0^{N-t}} d\hat{x}_{N-t+1|N-t+1}\end{aligned}$$

where $f^{(1)}_{\hat{x}_{N-t+1|N-t+1}|y_0^{N-t}} \sim N(0, \frac{(\sigma^2_{N-t+1|N-t})^2}{\sigma^2_{N-t+1|N-t}+\sigma^2_v})$. For $s=1$, $\Lambda_{(s,t)}$'s are given by

$$\Lambda_{(1,t)} = K_{(0,t-1)}A^2\sigma^2_{N-t|N-t} + \sum_{n=0}^{t-1} K_{(0,n)}\sigma^2_w, \quad 1 \le t \le N$$

Recall that $K_{(0,t)}$'s were defined by the recursion

$$K_{(0,t)} = 1 + A^2 K_{(0,t-1)}, \quad 1 \le t \le N-1$$

with $K_{(0,0)} = 1$. Let $\Delta_{(s,t)}$ be defined by $\Delta_{(s,t)} := J^{(0)}_{(s,t)} - J^{(1)}_{(s,t)}$. Then, for $1 \le s < t \le N$, we have

$$\begin{aligned}\Delta_{(s,t)}(\hat{x}_{N-t|N-t}) \;=\;& A^2(\hat{x}_{N-t|N-t})^2 + A^2\sigma^2_{N-t|N-t} + \sigma^2_w \\ & + \Lambda_{(s,t-1)} - \Lambda_{(s,t)} \\ & + \int_{|\hat{x}_{N-t+1|N-t+1}| \le \tau^+_{(s,t-1)}} \Delta_{(s,t-1)} \\ & \quad \times f^{(0)}_{\hat{x}_{N-t+1|N-t+1}|y_0^{N-t}} d\hat{x}_{N-t+1|N-t+1}\end{aligned}$$

Note that $\Lambda_{(s,t)}$ is a sequence of real numbers, whereas $\Delta_{(s,t)}(\hat{x}_{N-t|N-t})$ is a sequence of real-valued functions. We are also given the boundary conditions:

$$\begin{aligned}\Lambda_{(t,t)} &= \Lambda_{(t-1,t-1)} + A^2\sigma^2_{N-t|N-t} + \sigma^2_w, 1 \le t \le N \\ \Delta_{(t,t)}(\hat{x}_{N-t|N-t}) &= A^2(\hat{x}_{N-t|N-t})^2, 1 \le t \le N\end{aligned}$$

with $\Lambda_{(0,0)} = 0$. We also have $\tau^+_{(t,t)} = 0, \quad 1 \le t \le N$, and for $1 \le s < t \le N$, the thresholds, $\tau^+_{(s,t)}$, are given by the positive solution of the nonlinear equation $\Delta_{(s,t)}(\tau^+_{(s,t)}) = 0$. In order to show that such a solution exists, we have the following result which can be proven by induction [4].

Theorem 1: Let $N \ge 2$ be given. For $1 \le s < t \le N$, the sequence of functions $\Delta_{(s,t)}(u)$ are even, differentiable with a unique critical point at $u=0$, i.e., $\left.\frac{\partial \Delta_{(s,t)}(u)}{\partial u}\right|_{u=0} = 0$. Furthermore, we have $\frac{\partial \Delta_{(s,t)}(u)}{\partial u} > 0$, if $u > 0$, and $\frac{\partial \Delta_{(s,t)}(u)}{\partial u} < 0$, if $u < 0$. Thus, $\Delta_{(s,t)}(u)$ achieves its global minimum at the critical point $u=0$. Also, the minimum value of $\Delta_{(s,t)}(u)$ achieved at $u=0$ is nonpositive, i.e., $\Delta_{(s,t)}(0) \le 0$.

In the offline determination of the thresholds, $\tau^+_{(s,t)}$, we start with $s=1$, and increase t from $t=1$ to $t=N$ and determine $\tau^+_{(1,t)}$. Next, we increment s by 1, to $s=2$, and increase t from $t=2$ to $t=N$, and determine $\tau^+_{(2,t)}$. We repeat the procedure until $s=N$, at which point we stop since we already know that $\tau^+_{(N,N)} = 0$. This procedure enables us to determine $\tau^+_{(s,t)}$ for all pairs of (s,t) such that $1 \le s \le t \le N$.

IV. Numerical Solutions

In this section, we use numerical integration and Matlab to compute the recursions of $\Delta_{(s,t)}$ and $\Lambda_{(s,t)}$, with the aim of determining the sequence of thresholds $\{\tau^+_{(s,t)}\}$ for a given N. After determining the thresholds, we implement the optimal controller with $1 \le M \le N$ controls for a given plant. As we vary M, we look at how the average optimal N-stage sample-path cost, $J^*_{(M,N)}$, changes. In particular, let the decision horizon N be of length $N=20$, plant be marginally stable with $A=1$, and the noise variances be given as $\sigma^2_w = \sigma^2_v = 1$.

TABLE I
Comparison of optimal average sample-path costs for $1 \le M \le N$, $N = 20$.

M	$J^*_{(M,N)}$	%
1	96.4266	203.9327
2	68.1907	114.9343
3	47.2060	48.7914
4	44.0160	38.7366
5	39.8642	25.6503
6	37.1557	17.1132
7	35.6168	12.2627
8	34.1551	7.6555
9	33.6935	6.2005
10	33.6913	6.1936
11	33.2445	4.7853
12	32.9262	3.7820
13	32.8267	3.4684
14	32.4639	2.3249
15	32.1082	1.2037
16	31.9824	0.8072
17	31.8822	0.4914
18	31.8417	0.3637
19	31.7337	0.0233
20	31.7263	0

The initial state is taken to be zero-mean, with variance $\sigma^2_{x_0} = 1$. After determining the thresholds, we run several instances of the plant process applying the optimal control policy with $M \in [1,20]$ controls, and calculate the sample-path average of the performance criterion, $J_{(M,N)}$. We create 100 instances of the plant process, and for each instance we calculate the 20-stage sample-path cost as the value of M is varied from 1 to 20. To investigate the dependence of the average optimal sample-path cost, $J^*_{(M,N)}$, to M, we average the sample path costs over 100 instances of the plant process for each M. The results are tabulated in Table I, where the "%" column indicates the percentage difference between the sample-path average cost of the control policy with limited $M < N$ controls, and unlimited $M = N = 20$ controls. As can bee seen from Table I, the percentage improvement in the average cost in going from M to $M+1$ decreases as M approaches $N=20$. Therefore, if control is expensive, we may for instance settle for $M=9$ controls, instead of $M=20$, which gives a cost within approximately 6% of the unlimited number of controls case. Note that the improvement in the cost in going from a particular number of controls to another one is a function of the plant parameters.

References

[1] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 1995.

[2] "Industrial wireless technology for the 21st century," based on views of the industrial wireless community, in collaboration with U.S. Department of Energy (DOE), 2002.

[3] Y. B. Shalom and T. Edison, "Dual effect, certainty equivalence, and separation in stochastic control," *IEEE Transactions on Automatic Control*, vol. 19, no. 5. pp. 494–500, October 1974.

[4] O. C. Imer, "Optimal estimation and control under communication network constraints," *Ph.D. Dissertation*, University of Illinois, 2005.

[5] O. C. Imer and T. Başar, "Optimal estimation with limited measurements," in *Proc. IEEE CDC 2005*, December 2005.