Dynamic Sampling Applied to Problems in Optimal Control[1]

D. Assaf[2] Y. Ritov[3]
Communicated by D. A. Carlson

Abstract

Dynamic sampling utilizes the option of varying the sampling rates according to the situation of the systems, by that obtaining procedures with improved efficiencies. In this paper, the technique is applied to a typical problem in optimal control theory, that of tracking and controlling the position of an object. It is shown that the dynamic sampling results in a significantly improved procedure for this case, even when applying a suboptimal policy which can be analyzed in a closed form.

Key Words: Brownian motion, diffusion processes, observers.

---

[1]Supported by a grant from the Israel Foundation Trustees.
[2]Associate Professor, Department of Statistics, The Hebrew University, Jerusalem, Israel.
[3]Professor, Department of Statistics, The Hebrew University, Jerusalem, Israel.

1 Introduction and Summary

Consider a typical problem in control theory, such as a tracking or regulator problem. As a specific illustration we study the problem of tracking and controlling an object using a radar system. The trajectory is governed by a system of differential equations depending on the specification and the control applied. Our goal is to keep the object on path, taking the cost of the control into account. Our knowledge regarding the current state of the object is obtained by continuous tracking by the radar system.

The main difficulty of the tracking system is the presence of noise and we make the distinction between two types of noise. The first is a noise affecting the actual movement of the object and the tracking system, due to various factors such as unpredictable weather conditions or inadequate modelling. The second type is the measurement noise present in the tracking system, which results in an imprecise "reading" of the exact position. While both noises are present in the models studied in this paper, the main focus is the second type. In the dynamic sampling formulation, it is assumed that this second type of noise can be adjusted at any time according to our will.

A common model, which is used in the theoretical as well as in the applied work, is that the noise is a Gaussian white noise with fixed variance coefficient which depends on the characteristics of the tracking system, see e.g., Refs. 1–3. In a multitracking system the variance is inversely proportional to the amount of the resources devoted to the particular object. This merely reflects the fairly trivial fact that the more resources allocated — the better the performance. See Ref. 4.

The basic idea studied in this paper is the following. Rather than allocating a fixed amount of the resources to each tracked object, apply a "dynamic allocation" procedure. Using this approach, the allocation is under our control and can be changed any time according to actual need. Thus, when the location of a given object is known fairly well, most of the resources may be shifted to the tracking of the other objects.

Mathematically, we study a somewhat simplified version of this problem in which a single object is being tracked, but rather than having a fixed measurement rate, this rate is under our control and a cost (or constraint) is associated in accordance with the rate selected. Thus, in this model, the "tracking intensity" may be selected at any given instant according to all current known information at that time. In particular the intensity may depend on the current estimated state of the object and the variance of the estimate. The surprising result is that under the optimal strategy the sampling rate may be infinity. This may be not too practical, but we believe that this is a fair mathematical approximation to what can be done in practice when most of the resources are devoted to the particular unit under study.

A review of the results for the basic model (with a fixed sampling rate) is given in Section 2. The dynamic sampling formulation of the problem, as well as some basic conclusions, is presented and analyzed in Section 3. In Section 4 we propose a suboptimal sampling-and-control policy which can be analyzed in closed form. The performance of this policy is then compared to the fixed rate one in Section 5, and it is shown that the improvement obtained by dynamic sampling is indeed most significant. For ease of reference, we include a short appendix with a summary of the basic results for stochastic processes needed in the analysis.

It should be emphasized that the main feature of the dynamic sampling method is that it provides a highly improved performance without using any additional "physical" resources.

2 Basic Model

The model described may be thought of as a simplified version of tracking and controlling the position of an object. We consider a one dimensional movement of an object whose current position can be estimated. Our goal is to keep the object on track that for simplicity we take to be identically 0. Deviations from the desired position are assumed to result in a cost proportional to the square of the distance of the object from 0.

To make the description precise, we consider the following model. Let $(t_\zeta, s_\zeta, u_\zeta)$, $\zeta > 0$, be a path in R^3 describing accumulated "real time", "sampling time", and "control time". The functions t_ζ and s_ζ are non-decreasing in ζ, $(t_0, s_0, u_0) = (0, 0, 0)$, $1 \leq \dot{t}_\zeta + \dot{s}_\zeta + |\dot{u}_\zeta| < \infty$ (where dot denotes the corresponding derivative with respect to ζ) and $\lim_{\zeta\to\infty} t_\zeta = \infty$. In the standard model, t_ζ is simply time, $s_\zeta = t_\zeta$ describes sampling at a fixed rate of unity, and $\dot{u}_\zeta$ equals the control intensity. The control time is the integral of the control intensity.

Mathematically, this unusual parameterization enables us to describe policies with possible infinite sampling rate and control intensity without worrying about regularity issues. From a practical point of view, an infinite control or sampling rate will typically serve as an approximation to situations in which a very high rate is employed. To keep the corresponding cost within reason, this high rate will be used for very short time durations. The "infinite" rate is the limit as the rate goes to infinity while the time duration goes to zero.

The true position of the object is a random process described by $y_\zeta = y(t_\zeta, s_\zeta, u_\zeta)$:

$$dy(t_\zeta, s_\zeta, u_\zeta) = du_\zeta + \sigma_2 \, dB_y(t_\zeta), \qquad y(0,0,0) = 0.$$

The position y is not observed directly. Instead, we observe the process z_ζ which is given by

$$dz_\zeta = dz(t_\zeta, s_\zeta, u_\zeta) = y(t_\zeta, s_\zeta, u_\zeta) \, dt_\zeta + \sigma_1 \, dB_z(s_\zeta)$$

The processes B_y and B_z are two independent standard Brownian motions, and σ_1, σ_2 are assumed positive and known. Let $\{\mathcal{F}_\zeta\}_{\zeta>0}$ be the smallest filtration that makes the $z(\cdot)$ measurable. To be precise, we should say that the process $\{(t_\zeta, s_\zeta, u_\zeta),\ \zeta > 0\}$ is adaptive to $\{\mathcal{F}_\zeta\}$. Let

$$\begin{aligned} x_\zeta = x(t_\zeta, s_\zeta, u_\zeta) &= \mathrm{E}\left\{ y_\zeta \mid \mathcal{F}_\zeta \right\}, \qquad \zeta > 0 \\ v_\zeta = v(t_\zeta, s_\zeta, u_\zeta) &= \mathrm{E}\left\{ (y_\zeta - x_\zeta)^2 \mid \mathcal{F}_\zeta \right\}, \qquad \zeta > 0, \end{aligned}$$

be the current estimate of the object's position and the a-posteriori variance of this position. It is well known that the state trajectory, $(x_\zeta, v_\zeta)^T$ is described by the following differential equation:

$$d \begin{pmatrix} x(t_\zeta, s_\zeta, u_\zeta) \\ v(t_\zeta, s_\zeta, u_\zeta) \end{pmatrix} = \begin{pmatrix} 0 \\ \sigma_2^2 \end{pmatrix} dt_\zeta + \begin{pmatrix} (v_\zeta/\sigma_1) \, dB_x(s_\zeta) \\ -v_\zeta^2 \sigma_1^{-2} \, ds_\zeta \end{pmatrix} + \begin{pmatrix} du_\zeta \\ 0 \end{pmatrix}, \tag{1}$$

where B_x is some standard Brownian motion.

The objective considered in this paper is that of minimizing the long run average expected cost per unit time given by

$$\lim_{T\to\infty} \frac{1}{T} \mathrm{E} \left\{ \int \left[y^2(t_\zeta, s_\zeta, u_\zeta) \, dt_\zeta + c \, | \, du_\zeta | \right] \mathbf{1}(t_\zeta \leq T) \right\}, \tag{2}$$

where $\mathbf{1}$ is the indicator function. Since $\mathrm{E}\{y^2 \mid \mathcal{F}_\zeta\} = x_\zeta^2 + v_\zeta$ the objective (2) is equivalent to

$$\lim_{T\to\infty} \frac{1}{T}\mathrm{E}\left\{ \int \left[(x_\zeta^2 + v_\zeta)\, dt_\zeta + c \,|\, du_\zeta| \right] \mathbf{1}(t_\zeta \le T) \right\}. \tag{3}$$

For the basic model $t_\zeta = s_\zeta$ and the result is well known. Clearly $v_\zeta \to \sigma_1\sigma_2$. The control is of the 0-∞ type. Formally, for some $\mu > 0$ depending on c in (3), the control time is a step function with steps whenever $|x_\zeta| = \mu$ to enforce $|x_\zeta| \le \mu$. The process $\tilde{x}(\cdot)$, $\tilde{x}(t_\zeta) = x(t_\zeta, s_\zeta, u_\zeta))$ becomes a Brownian motion with variance coefficient σ_2^2 and reflecting barriers at $-\mu$ and μ. This can be observed as a limit of the the bounded control case, see Ref. 3. The resulting controlled process is with impulsive control, see Ref. 5 for a general discussion and Ref. 6 for the separation principle for impulsive control. See Ref. 7 for application of similar models.

3 Dynamic Sampling Formulation

3.1 Preliminaries

Think of σ_1^2 as the variance due to sampling error when sampling at some standard rate of unity. When sampling is done in this fixed rate, a sampling amount of Δt is accumulated for every Δt real time units, resulting in an estimator with variance $\sigma_1^2/\Delta t$ as described in the basic model. Note that in the basic model sampling time and real time are equal by definition.

Applying dynamic sampling, the following generalization is studied: Suppose that rather than sampling at a fixed rate we may vary the sampling rate as we please, as long as some constraint on the *average* sampling rate is satisfied. (Alternatively a cost on sampling may be imposed rather than constraint.)

This possibility may be particularly appealing when a large number of objects should be observed in parallel. Rather than giving each object a fixed share of the observer (either in terms of observer units or time slots) we may want to divide the observer between the objects depending on their state and the importance of an immediate updating of the position of each of them. As an example, any procedure which results in devoting to each object two time slots per cycle half of the time and no time slots for the remaining half, will keep the long run average sampling rate at the same level of the fixed rate sampling procedure which constantly devotes one time slot per object.

Dynamic sampling is the idea behind the cumbersome description of the previous model. Namely, we permit paths with $s_\zeta \ne t_\zeta$. We actually permit "infinite" sampling rate, that is, paths with intervals where s_ζ is strictly increasing while t_ζ remains constant.

3.2 Optimality Equation

It is easier to write the version with costs rather than constraints. Let c_1 be the cost of one unit of control time and c_2 the cost of one unit of sampling time. Next write the formal optimality equation with the objective of minimizing the average cost per unit time.

Begin with some state (x, v), apply the control rate $\dot{u}$ and sampling rate a for Δt units of time, and proceed optimally from the resulting state onwards. The general theory of dynamic programming and control (with average cost criteria) suggests to try and find a number l and a function $F(x, v)$ satisfying, up to $o(\Delta t)$ terms:

$$\begin{aligned} & F(x,v) + l\Delta t \\ & = \inf_{a,\dot{u}} \left\{ c_1|\dot{u}|\Delta t + c_2 a\Delta t + (x^2+v)\Delta t + \mathrm{E}\{F(x+\Delta X, v+\Delta V)\}\right\} \end{aligned} \tag{4}$$

where the third term in the RHS of (4) is the cost due to the square deviation from 0, "variation cost" henceforth, during the given time interval and the expectation in the fourth term is for the distribution of ΔX and ΔV which in turn depends on the choice of $\dot{u}$ and a. In the Taylor expansion of the last term, both $\mathrm{E}(\Delta V)^2$ and $\mathrm{E}(\Delta X \cdot \Delta V)$ are $\mathrm{o}(\Delta t)$ terms. Carrying on this expansion, yields, after cancelling $F(x, v)$ from both sides of (4) and omitting all $\mathrm{o}(\Delta t)$ terms

$$\begin{aligned} l &= \inf_{a,\dot{u}} \Bigg[c_1|\dot{u}| + c_2 a + x^2 + v + \dot{u}F_1'(x,v) \\ & \qquad + av^2F_{11}''(x,v)/2\sigma_1^2 + \left(\sigma_2^2 - av^2/\sigma_1^2\right)F_2'(x,v)\Bigg]. \end{aligned} \tag{5}$$

Rearrange (5) to obtain

$$\begin{aligned} & l - x^2 - v - \sigma_2^2 F_2'(x,v) \\ & = \inf_{\dot{u},a} \left\{ |\dot{u}|[c_1 - |F_1'(x,v)|] + a\left[c_2 + v^2\sigma_1^{-2}\left(F_{11}''(x,v)/2 - F_2'(x,v)\right)\right]\right\}. \end{aligned} \tag{6}$$

For formality, assume that we are restricted to $a, |\dot{u}| \leq B$. Since the term inside the infimum on the right hand side of (6) is linear in the sampling rate a and piece-wise linear in the control $\dot{u}$, it follows that the minimizing values are 0 or B for a and $-B$, 0, or B for $\dot{u}$, depending on the signs of the appropriate quantities. Take the limit as $B \to \infty$ to heuristically deduce that the optimal sampling and control rates are 0—∞. Hence the optimal policy consists of a partition of the half plane $\{(x, v) : -\infty < x < \infty,\ v \geq 0\}$ into at most six regions.

The 0—∞ policy of Section 2 follows immediately from (4) when we consider $F(x, v) = F(x)$ and the infimum is taken only over $\dot{u}$.

To obtain the general structure of these regions we can use the following heuristic arguments:

H1 The optimal control is symmetric in x.

H2 A control period should not be followed by a sampling period, and there should not be a period of control and sampling. Rationale: since the state moves in a deterministic way during a control period, then if there is a sampling period after the control period, it be beneficial to change the order of these two periods, it may be that the sampling period would reveal that there is no need for control. Hence there are three types of regions: $\mathcal{C}$ in which only control is applied, $\mathcal{S}$ in which we sample and $\mathcal{R}$ in which nothing is done. We conclude that if $(x, v) \in \mathcal{C}$, and $|x''| > |x| > |x'|$ then $(x', v) \in \mathcal{C} \cup \mathcal{R}$ and $(x'', v) \in \mathcal{C}$.

H3 After "doing nothing", there should be a sampling period. Hence if $(x,v) \in \mathcal{R}$ and $v' > v$ then $(x,v') \in \mathcal{R} \cup \mathcal{S}$.

H4 If $(x,v) \in \mathcal{S}$, then after the sampling period starting at (x,v) the path cannot end necessarily in $\mathcal{C}$, nor necessarily in $\mathcal{R}$. We sample to get a "practical" conclusion.

H5 If $(x,v) \in \mathcal{R}$ and $|x'| < |x|$ then $(x',v) \in \mathcal{R}$.

H6 If $(x,v) \in \mathcal{R}$ and $v' < v$ then $(x,v') \in \mathcal{R}$.

H7 If $(x,v) \in \mathcal{S}$ and $v' > v$ then $(x,v') \in \mathcal{S}$.

Careful analysis of the heuristic arguments reveals that the general structure of the optimal policy must be of the form described in Figure 1. The main points are as follows: The policy is symmetric by **H1** . By **H7** $\mathcal{S}$ is open from above, and by **H2** , $\mathcal{C} \cap \{(x,v) : \ x > 0)\}$ is open to the right. Next, by **H2** the boundary between $\mathcal{C}$ and $\mathcal{R}$ is not empty. By **H3** and **H6** the line N of Figure 1 must be vertical. By **H4** the boundary between $\mathcal{S}$ and $\mathcal{C}$ is not empty. The line M of Figure 1 should be horizontal by **H2** and **H5** . By **H4** the boundary between $\mathcal{S}$ and $\mathcal{R}$ is not empty. The line L is decreasing by **H5 .**

It is worth noting that using a policy of this type forces the (x,v) process to remain along the line L almost always on the real time scale, excluding perhaps some initial transient period. The state moves along this line in a *non*-continuous path.

4 Rectangular Policy and Its Analysis

The optimization problem described at the end of Section 3 requires explicit solutions to the equations of the lines L, M, and N. This is a two dimensional free boundary problem involving a p.d.e. Problems of this type can rarely be solved explicitly, and the present one does not seem to be an exception. In this section we focus on a class of suboptimal policies which can be analyzed in a closed form.

Definition 4.1 *The rectangular policy, with parameters ν_1, ν_2, and μ_0 (such that $0 < \nu_1 < \nu_2$, $\mu_0 > 0$), is defined by*

$$(\dot{t}_\zeta, \dot{s}_\zeta, \dot{u}_\zeta)(x,v) = \begin{cases} (1,0,0) & |x| \leq \mu_0, \quad v < \nu_2 \\ (1,(\sigma_1\sigma_2/\nu_2)^2,0) & |x| \leq \mu_0, \quad v = \nu_2 \\ (0,0,-1) & x > \mu_0, \quad v \leq \nu_1 \\ (0,0,1) & x < -\mu_0, \quad v \leq \nu_1 \\ (0,1,0) & \textit{otherwise} \end{cases}$$

In other words, we may begin with a transient period. Afterwards, as long as $|x| < \mu_0$, we sample at a constant rate to keep the a-posteriori variance equal ν_2. After "crossing" the μ_0 line we "examine" the position more carefully (in negligible real time), until either the estimate becomes less than μ_0 or the variance drops down to ν_1. In the latter case, a fast control brings the position back into the $(-\mu_0, \mu_0)$ interval. Note that after the transient period, the object is kept almost always (on the real time scale) along the boundary of the no-sample/no-control region.

Theorem 1 *Suppose we use the rectangular policy, then the long run average sampling, variation, and control per unit time are given by S/R, Q/R, and $1/R$, respectively, where:*

$$S = \frac{\sigma_1^2}{\nu_2^2}\left[2\mu_0 + \sqrt{\frac{2}{\pi}}\frac{(\nu_2-\nu_1)^{1/2}}{\nu_1/\nu_2} + \frac{\nu_2^{1/2}}{\sqrt{2\pi}}\log\left(\frac{1+\sqrt{1-\nu_1/\nu_2}}{1-\sqrt{1-\nu_1/\nu_2}}\right)\right]$$

$$Q = \frac{2}{\sigma_2^2}\left[\mu_0\left(\frac{1}{3}\mu_0^2+\nu_2\right) + \sqrt{\frac{2}{\pi}}\sqrt{\nu_2-\nu_1}\left(\frac{1}{3}\nu_1+\frac{2}{3}\nu_2+\mu_0^2\right)\right]$$

and

$$R = 2\sigma_2^{-2}\left[\mu_0 + \sqrt{2/\pi}\sqrt{\nu_2-\nu_1}\right]$$

Corollary 2 *The results for constant rate sampling are given by the theorem with $\nu_1 = \nu_2$. In particular, the case of unit sampling rate is given by $\nu_1 = \nu_2 = \sigma_1\sigma_2$.*

The rest of the paper is devoted to the discussion and proof of the theorem.

Using the rectangular policy, the state of the system (x_ζ, v_ζ) follows a renewal process (to be exact, a delayed renewal process for some initial transient period). It thus suffices to compute quantities of interest for some specific renewal cycle. We find it convenient to consider cycles beginning and ending at the top right corner (the point (μ_0, ν_2)). Symmetry of the whole process w.r.t. the vertical axis will be used in an obvious way. Furthermore, we split the above cycle into two parts and compute the expected real time R, sampling time S , total variation Q, and total control needed U for each of the two parts.

I. *First part of the cycle.* Consider the quantities starting at $(\mu_0-\varepsilon, \nu_2)$ until reaching either (μ_0, ν_2) or $(-\mu_0+\varepsilon, \nu_2)$. If (μ_0, ν_2) is reached first, then the second part of the cycle begins at that state. If $(-\mu_0+\varepsilon, \nu_2)$ is reached first — the complete cycle is terminated (and another one begins starting at the symmetric point $(-\mu_0+\varepsilon, \nu_2)$). Quantities for this first part of the cycle will be indexed by 1.

It follows from (1) that to sustain the fixed error variance ν_2 we should sample with constant sampling rate: $ds = (\sigma_1^2\sigma_2^2/\nu_2^2)\ dt$. Moreover, it follows from (1) that during that period the x coordinate is distributed as $\mu_0-\varepsilon+\sigma_2 B_1(t)$ (when the first part starts at $t=0$). Hence, we can use (14) to conclude that

$$R_1 \quad = \quad 2(\mu_0-\varepsilon)\varepsilon/\sigma_2^2.$$

Hence

$$S_1 \quad = \quad 2(\mu_0-\varepsilon)\sigma_1^2\varepsilon/\nu_2^2.$$

Since no control is applied during this stage we also have

$$U_1 = 0 \tag{7}$$

Finally we compute the expected square deviation during the first part of a cycle, Q_1.

$$Q_1 \quad = \quad \mathrm{E}\int_0^\tau (x^2(t)+\nu_2)\ dt$$

$$= \quad \mathrm{E}\int_0^\tau x^2(t)\,dt + \nu_2 R_1,$$

where τ is the real time needed for the first part of a cycle. By (17) with $b = 2(\mu_0 - \varepsilon)$ and $c = \mu_0 - \varepsilon$:

$$Q_1 = \frac{2\mu_0}{\sigma_2^2}\left(\frac{\mu_0^2}{3} + \nu_2\right)\varepsilon + \mathrm{o}\left(\varepsilon\right).$$

II. *Second part of the cycle.* Here we consider the process starting at (μ_0, ν_2) until it returns to $(\mu_0 - \varepsilon, \nu_2)$. In the most general case this consists of some sampling time followed by a control time and ending in real time (during which neither sampling nor control are applied). In the sampling period we sample until either x is equal to $\mu_0 - \varepsilon$ or v is equal to ν_1, whichever comes first. In the latter case, we use control to bring x to $\mu_0 - \varepsilon$, otherwise, no control is applied. In the final stage, we simply wait until reaching the point $(\mu_0 - \varepsilon, \nu_2)$ again.

A key observation for computing the relevant quantities, indexed by 2, for this part of the cycle is that random factors occur only during the sampling period. Let A be the amount of reduction along the vertical axis and let B be the coordinate of the horizontal axis, both upon termination of the sampling portion. We further clarify by pointing out that both A and B are random, $0 \le A \le \nu_2 - \nu_1$ with an atom at $\nu_2 - \nu_1$ and $\nu_2 - A$ is the actual value of $v(t)$ upon termination of sampling. Also $B \ge \mu_0 - \varepsilon$ with an atom at $\mu_0 - \varepsilon$.

Next, let $\tilde{x}(s)$, be the current estimate of the position and $\tilde{v}(s)$ the variance of the actual position given its estimate after sampling the amount s during the sampling period. It follows from (1) that $\dot{\tilde{v}}(s) = -\tilde{v}^2(s)/\sigma_1^2$. Since $\tilde{v}(0) = \nu_2$, we have

$$s = \sigma_1^2(\tilde{v}^{-1}(s) - \nu_2^{-1}). \tag{8}$$

Moreover, $\tilde{x}(\cdot) - \mu_0$ is distributed like

$$W\left(\sigma_1^{-2}\int_0^{\cdot} \tilde{v}^2(s)\,ds\right) = W\left(\nu_2 - \tilde{v}(\cdot)\right),$$

where $W(\cdot)$ is a standard Brownian motion, stopped at time $\nu_2 - \nu_1$ or when it hits $-\varepsilon$ for the first time. If we denote this stopping time by T, then $A \sim T$ and $B - \mu_0 + \varepsilon \sim W(T)$.

In the real time period, the conditional variance increases at a rate σ_2^2, see (1). It follows that $R_2 = \mathrm{E}(A/\sigma_2^2)$. By (18) of the Appendix:

$$R_2 = 2\varepsilon\sigma_2^{-2}\sqrt{2/\pi}(\nu_2 - \nu_1)^{1/2} + \mathrm{o}\left(\varepsilon\right), \tag{9}$$

Similarly for S_2, by (8)

$$S_2 = \sigma_1^2 \mathrm{E}\left(\frac{1}{\nu_2 - A} - \frac{1}{\nu_2}\right)$$

which after some elementary calculation and applying (20) results in

$$S_2 = \frac{\sigma_1^2}{\nu_2}\left[\sqrt{\frac{2}{\pi}}\frac{(\nu_2 - \nu_1)^{1/2}}{\nu_1} + \frac{1}{\sqrt{2\pi}}\frac{1}{\nu_2^{1/2}}\log\left(\frac{1 + \sqrt{1 - \nu_1/\nu_2}}{1 - \sqrt{1 - \nu_1/\nu_2}}\right)\right]\varepsilon + \mathrm{o}\left(\varepsilon\right)$$

For the computation of U_2, note that

$$U_2 = \mathrm{E}(B - \mu_0 + \varepsilon) = \mathrm{E}(B - \mu_0 + \varepsilon \mid B > \mu_0 - \varepsilon)\, \mathrm{P}(B > \mu_0 - \varepsilon).$$

Using (15) and (16)

$$U_2 = 2\varepsilon\Phi\left(-\varepsilon/\sqrt{\nu_2 - \nu_1}\,\right)$$

Hence, as $\varepsilon \to 0$,

$$U_2 = \varepsilon + \mathrm{o}\,(\varepsilon)\,. \tag{10}$$

Finally, to compute the expected variation note that it is only accumulated with real time, i.e., during the third phase. During this time the x coordinate is fixed at $\mu_0 - \varepsilon$, contributing $(\mu_0 - \varepsilon)^2 R_2$. The variance however changes from $\nu_2 - A$ to ν_2 at a rate σ_2^2, taking a real time of $r = r(A) = A/\sigma_2^2$. The contribution of this part is hence

$$\int_0^r (\nu_2 - A + \sigma_2^2 s)\, ds,$$

which, after simple integration and rearrangement, equals

$$\nu_2 A/\sigma_2^2 - A^2/2\sigma_2^2.$$

Thus

$$Q_2 = \nu_2 \mathrm{E}(A)\sigma_2^2 - \mathrm{E}(A^2)/2\sigma_2^2 + (\mu_0 - \varepsilon)^2 R_2.$$

Using (9), (19) and (20) of the appendix we have

$$Q_2 = 2\sqrt{\frac{2}{\pi}}\frac{\sqrt{\nu_2 - \nu_1}}{\sigma_2^2}\left(\frac{1}{3}\nu_1 + \frac{2}{3}\nu_2 + \mu_0^2\right)\varepsilon + \mathrm{o}\,(\varepsilon)\,.$$

Summary of quantities for the complete cycle By (13), the probability of reaching μ_0 before reaching $-(\mu_0 - \varepsilon)$ (in the first part of the cycle) is $2(\mu_0 - \varepsilon)/(2\mu_0 - \varepsilon) = 1 + \mathrm{O}\,(\varepsilon)$. Hence the total expected real time for a cycle is $R_1 + R_2 + \mathrm{o}\,(\varepsilon)$, similar expressions for S, U, and Q. For all practical purposes, the quantities of interest are S/R, U/R, and Q/R representing, respectively, the long run average sampling, control, and variation per unit time. Consider these ratios as $\varepsilon \to 0$, so that $R = R_1 + R_2$, $S = S_1 + S_2$, $U = U_1 + U_2$, and $Q = Q_1 + Q_2$. This completes the proof of the theorem.

5 Efficiency and Asymptotic Results

In this section we compare the efficiency of the optimal procedure for the basic model (described in Section 2) to that of the rectangular policy (which is suboptimal for the dynamic sampling formulation). Note, that the basic model (with a possible control of the fixed sampling intensity) is a particular case of the rectangular policy with $\nu_2 = \nu_1$.

In the basic model we may select the value μ_0 according to our will. In addition, we may also select the fixed sampling intensity by a proper choice of ν_2 $(=\nu_1)$, or equivalently by adjusting the value of σ_1^2. Given any choice of these two parameters, we denote the values of U/R, Q/R and $\sigma_1^{-2}S/R$ by u_c, q_c and s_c respectively.

It may easily be checked that, given any prescribed values for the quantities $U/R = u > 0$ and $Q/R = q > 0$, such that $12qu^2 > 1$, one may find μ_0 and ν_2 $(=\nu_1)$ such that the performance of the basic policy with these parameters satisfies $u_c = u$ and $q_c = q$. A similar result holds when setting prescribed values for any other two of the three quantities.

The situation for the rectangular policy is similar but requires a bit more care. The rectangular policy has three parameters, and any selection of these parameters leads to values U/R, Q/R and $\sigma_1^{-2}S/R$. For any two prescribed values of $U/R = u > 0$ and $Q/R = q > 0$ however, there are many values of μ_0, ν_1, ν_2 such that the performance of the corresponding rectangular policy matches u and q. Of these values it is evidently best to take a minimizer of the third quantity, $\sigma_1^{-2}S/R$. Again, this argument can be extended to any two of the three quantities of interest.

The notation u_r, q_r and s_r will be used to denote the performance of a rectangular policy which is the best in this sense.

A simple way to compare the two methods is to fix any two of the quantities U/R, Q/R, and $\sigma_1^{-2}S/R$ and consider the ratio of the obtained values of the remaining quantity. To ease the comparison, we set, without loss of generality, $\sigma_2 = 1$ (this is a matter of using a different time scale).

First note that for all values of ν_1 $(=\nu_2)$ and μ_0:

$$q_c = s_c^{-1/2} + 1/(12u_c^2). \tag{11}$$

Define the sampling efficiency of a rectangular policy as the ratio between the sampling rate of a constant policy (s_c) and the obtained sampling rate of the rectangular sampling policy (s_r) when the two policies have the same variation and control rates (i.e., $q_c = q_r$ and $u_c = u_r$). By (11)

$$s_r/s_c \quad = \quad s_r\left(q_r - 1/12u_r^2\right)^2$$

The sampling efficiency can be quite large. The most extreme case occurs when $\nu_2 = 1$, $\nu_1 \to 0$, $\mu_0 \to \infty$ and $\mu_0\nu_1 \to \infty$. It follows from Theorem 1 that $s_r \to 1$ while $q_r - 1/(12u_r^2) \to (\pi - 2)/\pi$. Hence the sampling efficiency is $\pi^2/(\pi-2)^2 \sim 7.5732$. That is, sampling at a constant rate consumes, in this limiting case, 7.57 times more resources than sampling using the rectangular policy with the same average variation and control. The sampling efficiency can be considerable at less extreme cases as well. For example, with $\mu_0^2 = \nu_2$, (that is the maximum tolerable estimated variation is equal to the estimated standard error) the sampling efficiency of 192% is obtained with $\nu_1/\nu_2 = 0.26$.

Similarly, we can consider the control efficiency, that is the ratio between the control rate of a constant policy (u_c) and the obtained control rate of the rectangular sampling policy (u_r) when the two policies have the same sampling and variation rates (i.e., $s_c = s_r$ and $q_c = q_r$). By (11)

$$u_r/u_c = u_r\sqrt{12(q_r - s_r^{-1/2})} \tag{12}$$

The control efficiency can be extremely high. The extreme case is achieved when $\nu_1/\nu_2 \to 1$, $\mu_0/\nu_2^{1/2} \to 0$ and $\mu_0/(\nu_2-\nu_1)^2 \to 0$. Formally, this can be checked with substituting $\mu_0 = 0$ and $\nu_2 = 1$, and then taking the limit as $\nu_1 \nearrow 1$. One may note that the control efficiency does not depend on ν_2 per se, but only through ν_1/ν_2 and $\mu_0/\nu_2^{1/2}$. Let $y = 1 - \nu_1/\nu_2$. Then

$$\begin{aligned}
s_r &= \frac{1}{2}\left(\frac{1}{1-y} + \frac{1}{2\sqrt{y}}\log\left(\frac{1+\sqrt{y}}{1-\sqrt{y}}\right) + \mathrm{o}\left(y^2\right)\right) \\
&= 1 + 2y/3 + 3y^2/5 + \mathrm{o}\left(y^2\right) \\
q_r &= 1 - y/3 + \mathrm{o}\left(y^2\right) \\
u_r &= (1 + \mathrm{o}\left(1\right))\sqrt{\pi/2y}
\end{aligned}$$

as $y \to 0$. Therefore it follows from (12) that

$$u_r/u_c = \sqrt{\pi y/5} + \mathrm{o}\left(y^{1/2}\right)$$

and the efficiency (u_c/u_r) converges to infinity as $y \to 0$. As a concrete example, if $\mu_0 = 0.1$, $\nu_1 = 0.83$ and $\nu_2 = 1$ then the efficiency is 219.5%.

The asymptotic result for the control efficiency may seem paradoxical, since $\nu_2/\nu_1 \to 1$ seems to mean that the efficiency is maximal when the rectangular policy is almost equal to a constant sample policy. The efficiency is large when a large rate of control is needed to satisfy the average variation and control demands. At the limit (when the product $q_r s_r \to 1$ which entails $\mu_0 \to 0$ and $\nu_2/\nu_1 \to 1$) the system is in one of two states, either no sampling and no control or extremely fast sampling to find if a control is needed and, if needed, in what direction. In particular, there is no constant sampling period. This adaptive second stage enables the efficient usage of the controller and the sampling.

6 Concluding Remarks

In this paper a particular simple control problem with the possibility of dynamic sampling is investigated and the form of an optimal policy is heuristically deduced. Its exact form however appears intractable and hence the rectangular policy is suggested and studied in detail. The rectangular policy closely imitates the basic characteristics of the optimal one, and its performance is shown to be extremely efficient compared to fixed rate sampling.

The rectangular policy (as well as the optimal one) utilizes control and sampling at "infinite" rates, and a precise mathematical formulation is set and analyzed for this case. From a practical point of view the use of infinity requires further clarification. Control at an infinite rate has been studied fairly widely in the literature (the resulting policy is often referred to as a control-limit policy), and we hence focus on the infinite sampling introduced in this paper. In practice sampling at infinite rate will typically be approximated by using the highest rate possible (and reasonable) under the circumstance at hand. In an example of controlling N objects using a radar, fixed rate sampling allocates, all the time, the same amount of sampling to any of the objects. The rectangular policy on the other hand suggests allocating most of the resources to one specified object for a short (and random) time, and then shifting the resources to another specified object.

It should be emphasized that the dynamic sampling procedure is suitable for those situations in which the intensity of the sampling equipment (e.g., radar) may be adjusted according to our will. We believe that the basic ideas presented and studied in this paper can be used in many relevant problems and complex situations to yield highly improved efficiencies.

7 Appendix: Results for Brownian Motion

The appendix lists several results which are used in the paper. All results are fairly standard or may easily be derived from such. Basic and convenient references are Refs. [8]– [10].

The notation used in this appendix is as follows. The standard Brownian motion is denoted by $\{B(t),\ t \geq 0\}$, $B(0) = 0$ unless otherwise stated. Let $W(\cdot) = \sigma B(\cdot)$, $T_a = \inf\{t : W(t) = a\}$ and $a \wedge b = \min\{a, b\}$. The results needed are:

For any $a, \varepsilon > 0$

$$\mathrm{P}(T_{-\varepsilon} < T_a) = \frac{a}{a+\varepsilon}. \tag{13}$$

$$\sigma^2 \mathrm{E}(T_{-\varepsilon} \wedge T_a) = \mathrm{E}\left\{W^2(T_{-\varepsilon} \wedge T_a)\right\} = a\varepsilon. \tag{14}$$

$$\mathrm{P}\left\{\inf_{0 \leq s \leq t} B(s) \leq -\varepsilon\right\} = 2\Phi\left(-\varepsilon/\sqrt{t}\right). \tag{15}$$

$$\mathrm{E}\left\{B(t) \mid \inf_{0 \leq s \leq t} B(s) > -\varepsilon\right\} = 2\varepsilon\Phi(-\varepsilon/\sqrt{t})/\left[1 - 2\Phi(-\varepsilon/\sqrt{t})\right]. \tag{16}$$

$$\begin{aligned} \mathrm{E}\left\{\int_0^{T_b \wedge T_{-\varepsilon}} (W(t)+c)^2\, dt\right\} &= \frac{1}{6\sigma^2}\left(\varepsilon\frac{(b-c)^4 + b(c+\varepsilon)^4}{b+\varepsilon} - c^4\right) \\ &= \frac{1}{6\sigma^2}\left(b^3 - 4b^2c + 6bc^2\right)\varepsilon + \mathrm{o}\left(\varepsilon\right). \end{aligned} \tag{17}$$

For the remaining results let $T = T(\varepsilon, t) = T_{-\varepsilon} \wedge t$ and $\sigma = 1$.

$$\mathrm{E}\,T = 2\sqrt{2t/\pi}\,\varepsilon + \mathrm{o}\left(\varepsilon\right). \tag{18}$$

$$\mathrm{E}\,T^2 = \frac{4}{3}\sqrt{2/\pi}\,t^{3/2}\varepsilon + \mathrm{o}\left(\varepsilon\right). \tag{19}$$

$$\mathrm{E}\left\{\frac{T}{b-T}\right\} = \sqrt{\frac{2}{\pi}}\left(\frac{t^{1/2}}{b-t} + \frac{1}{2b^{1/2}}\log\left(\frac{1+\sqrt{t/b}}{1-\sqrt{t/b}}\right)\right)\varepsilon + \mathrm{o}\left(\varepsilon\right) \tag{20}$$

References

1. GIKHMAN, I. I. and SKOROHOD, A. V. , Controlled Stochastic Processes, Springer Verlag, New York, New York, 1979.

2. KRYLOV, N. V. , Controlled Diffusion Processes, Springer Verlag, New York, New York, 1980.

3. BENEŠ, V. E. and KARATZAS, I. Optimal Stationary Linear Control of the Wiener Process, Journal of Optimization Theory and Applications, Vol, 35, pp. 611-633, 1981.

4. BAR SHALOM, Y. and FORTMAN, T. E., Tracking and Data Association, Academic Press, London, England, 1988.

5. BENSOUSSAN, A., Stochastic Control by Functional Analysis Methods, North-Holland Publications, Amsterdam, Holland, 1982.

6. Separation Principle for Impulse Control Problems , Proceedings of the American Mathematic Society, vol. 82, pp 439-445, 1982.

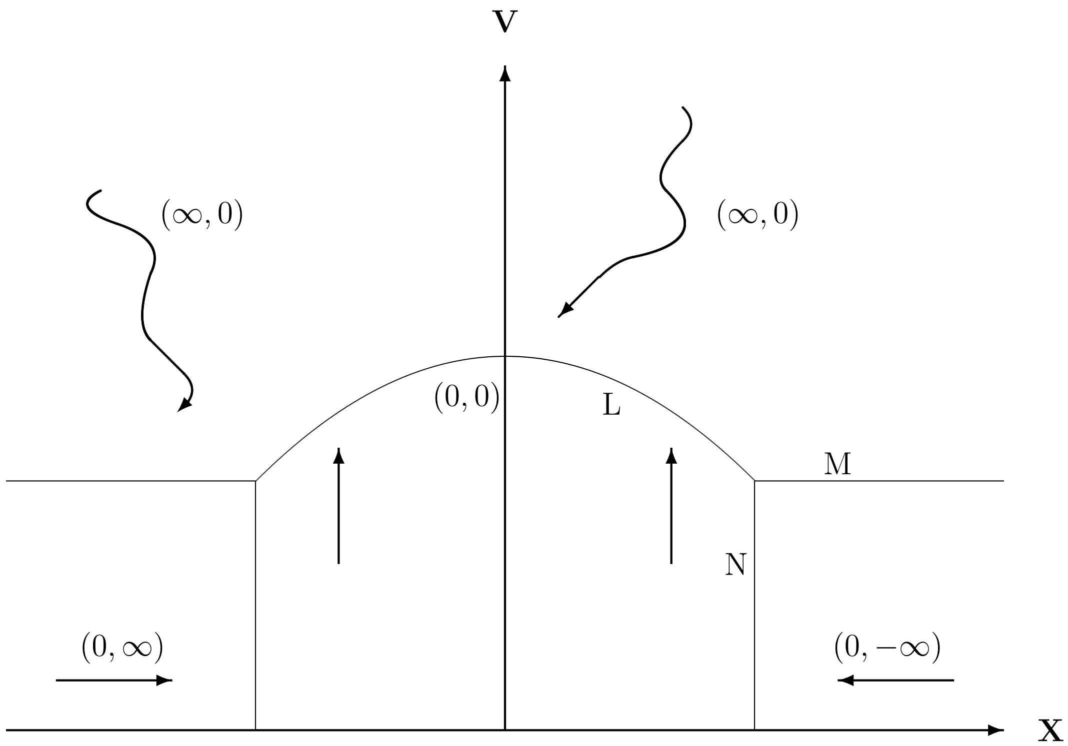
7. BENSOUSSAN, A. and TAPIERO, C. S. , Impulsive Control in Management: Prospects and Applica Journal of Optimization Theory and Applications, vol, 37, pp. 419-442, 1982.

8. ROSS, S. M. , Stochastic Process, Wiley, New York, New York, 1983.

9. KARLIN, S. and TAYLOR, H. M. , A First Course in Stochastic Processes, 2nd edition, Academic Press, New York, New York, 1975.

10. KARLIN, S. and TAYLOR, H. M. , A Second Course in Stochastic Processes, 2nd edition, Academic Press, New York, New York, 1981.

List of Figures


V
(∞, 0)
(∞, 0)
(0, 0)
L
M
N
(0, ∞)
(0, −∞)
X