On the Optimality of Reflection Control, with Production-Inventory Applications

[Extended Abstract]

Jiankui Yang *

David D. Yao †

Heng-Qing Ye ‡

ABSTRACT

We study the control of a Brownian motion (BM) with a negative drift, so as to minimize a long-run average cost objective. We show the optimality of a class of reflection controls that prevent the BM from dropping below some negative level r, by cancelling out from time to time part of the negative drift; and this optimality is established for any holding cost function $h(x)$ that is increasing in $|x|$. Furthermore, we show the optimal reflection level can be derived as the fixed point that equates the long-run average cost to the holding cost. We also show the asymptotic optimality of this reflection control when it is applied to production-inventory systems driven by discrete counting processes.

1. INTRODUCTION

Consider the control of a Brownian motion (BM) with a negative drift, so as to minimize a long-run average cost objective. We show the optimality of a class of reflection controls that prevent the Brownian motion from dropping below some negative level r, by cancelling out from time to time part of the negative drift; and this optimality is established for any holding cost function $h(x)$ that is increasing (i.e., non-decreasing) in $|x|$, where x is the state variable. This is a natural and desirable form of a cost function, since in applications, the absolute value of the state variable can be interpreted as finished-goods inventory or backordered demand (depending on the sign of x), both incurring costs. To the best of our knowledge, this is the most general form of the cost function for which the optimality of the reflection control is known. (Existing studies in the literature often require second-order properties such as convexity.) Furthermore, let $C(r)$ be the long-run average cost under the reflection control with the level r. We show the optimal reflection level r^* can be derived as the fixed point that equates the long-run average to the holding cost, $C(r^*) = h(r^*)$.

To prove the optimality of the reflection control, we follow the lower-bound method in Harrison and Taksar [5]. (Also see Harrison [6] and Taksar [9].) Focusing on a sub-class of the class of admissible controls (with the sub-class including the reflection control), we first find a lower bound on the cost objective, and then show this lower bound can be attained by a reflection control with a proper reflection level, which is thus optimal. To connect the result to a discrete production-inventory system driven by counting processes, we use the standard diffusion-limit approach (e.g., Reiman [8]), and establish the asymptotic optimality of the reflection control.

Many related studies in the literature that use BM in production-inventory systems focus on two-sided controls such as the (s, S) policy, whereas the reflection control we focus on here is one-sided. Refer to [1, 3, 4, 7, 10, 11], among many others. However, these papers all need to assume piecewise linear or convex/quasi convex cost functions with polynomially-bounded growth; whereas we only need a cost function $h(x)$ that is increasing in $|x|$, and we can allow it to have exponentially-bounded growth.

2. THE CONTROL PROBLEM

Given a Brownian motion with a *negative* drift, $X(t) := \theta t + \sigma B(t)$, where $\theta < 0$ and $\sigma > 0$ are given constants and $B(t)$ denotes the standard Brownian motion, we want to find a control, denoted $\{Y(t), t \in [0, +\infty)\}$, such that the state process

$$Z(t) = z_0 + X(t) + Y(t), \quad t \in [0, +\infty), \tag{1}$$

with z_0 being the initial state, will approach a stationary limit $Z(+\infty)$ that minimizes a cost objective $\mathsf{E}h[Z(+\infty)]$. In general, the expected long-run average cost is given by

$$\mathrm{AC}(x, Y) = \limsup_{t\to+\infty} \mathsf{E}_x \frac{1}{t}\int_0^t h(Z(u))\mathsf{d}u. \tag{2}$$

And we have $AC(z_0, Y) = \mathsf{E}h[Z(+\infty)]$, provided the control Y induces a steady-state distribution for $Z(t)$.

- *Assumptions on the cost function.* Here the cost function $h(x)$ is assumed to be continuous, and increasing in $x \geq 0$ and decreasing in $x \leq 0$. This implies $h(x) \geq h(0)$ for all x, and $h(0)$ is assumed to be finite. In both directions, x tends to $+\infty$ or $-\infty$, $h(x)$ goes to $+\infty$. Of course, we also need $\mathsf{E}h[Z(+\infty)] < +\infty$, and this requires h to be exponentially bounded; see (7) below.

Note the negative drift of $X(t)$ will drive the state process to $-\infty$ without any control, and hence achieve an objective

*School of Science, Beijing University of Posts and Telemmunications, Beijing; yangjk@bupt.edu.cn.

†Supported in part by NSF Grant CMMI-1462495. Dept of IEOR, Columbia University, New York; yao@columbia.edu.

‡Supported in part by HK/RGC Grant PolyU 155081/14B. Dept of Logistics and Maritime Studies, Hong Kong Polytechnic University, Hong Kong; lgtyehq@polyu.edu.hk.

value that is at one of the two largest extremes of h, i.e., $h(-\infty)$. Thus, the control $Y(t)$ is trying to cancel out, from time to time, this negative drift; and in this sense, $Y(t)$ is a *cumulative* effort up to t.

- *Admissible controls.* Let $\mathcal{A}$ denote the set of admissible controls. To be admissible a control must be non-anticipative and satisfy the following requirements: $Y(t)$ is increasing in $t \in [0, +\infty)$, with $Y(0) = 0$.

To motivate, consider a production-inventory system that supplies demand. Suppose demand rate is λ and production rate is μ. Let the state at time t be the *net* demand in the system, i.e., waiting orders minus produced quantities (both are cumulative up to t). Then, without any control, this net demand is $(\lambda - \mu)t + \sigma B(t)$, where $\sigma B(t)$ models the volatility (Gaussian noise) associated with demand (or, with both demand and production). Assume $\lambda < \mu$; hence, $\theta := \lambda - \mu < 0$, and denote this net demand as $X(t)$. Here, the control is to insert idle time into production; so denote the cumulative idle time up to t as $U(t)$. Then, production up t becomes $\mu[t - U(t)]$; and, with $Y(t) = \mu U(t)$, the state process can be expressed as follows:

$$\begin{aligned} Z(t) &= z_0 + (\lambda - \mu)t + \mu U(t) + \sigma B(t) \\ &:= z_0 + X(t) + Y(t). \end{aligned} \qquad (3)$$

Note that $Z(t)$, when positive, represents the volume of waiting orders; when $Z(t)$ is negative, its absolute value represents the volume of products waiting to supply demand (i.e., inventory). This also motivates why the cost function $h(x)$ is increasing in $x \geq 0$ and decreasing in $x \leq 0$ (the more negative x is, the higher the cost).

3. REFLECTION CONTROL

Recall, a Brownian motion (starting from 0) with a negative drift will have a stationary limit if it is *reflected* at some pre-specified value. Hence, we first focus on a sub-class of admissible controls, called "reflection controls," $\mathcal{A}^* \subset \mathcal{A}$; and denote a control in this class as $Y_r \in \mathcal{A}^*$, and denote the corresponding state process as Z_r. The control Y_r is defined by a reflection level r, meaning it ensures $Z(t) \geq r$ for all t.

Then, $Z_r(t) - r$ is a standard reflected Brownian motion (RBM); refer to [2] Section 6.2). It is known that Y_r and Z_r can be explicitly expressed as functions of X, the primitive (Brownian motion with drift), as follows:

$$Y_r(t) = \sup_{0 \leq u \leq t} (r - z_0 - X(u))^+, \qquad (4)$$

$$Z_r(t) = z_0 + X(t) + \sup_{0 \leq u \leq t} (r - z_0 - X(u))^+. \qquad (5)$$

In addition, *complementarity* holds: $[Z_r(t) - r]dY_r(t) = 0$ for all t, i.e., when $Z_r(t) > r$, $Y_r(t)$ cannot increase. Furthermore, the steady-state distribution of $Z_r(+\infty) - r$ follows an exponential distribution with rate $\gamma := -2\theta/\sigma^2$ (recall, $\theta < 0$).

Thus, under the reflection control Y_r, we have

$$\mathsf{E}h(Z_r(+\infty)) = \gamma \int_0^{+\infty} h(r + x)e^{-\gamma x}dx := C(r). \qquad (6)$$

Note that the left-hand-side is indeed $AC(z_0, Y_r)$.

Next, we want to find the r value that minimizes $\mathsf{E}h(Z_r(+\infty))$. But first note that for this expectation to be finite, we need the function $h(x)$ to satisfy the following condition: there exist positive numbers a and $b < \gamma/2$ (γ is specified above), such that

$$h(x) \sim o(ae^{bx}), \quad \exists a > 0, \, 0 < b < \gamma/2. \qquad (7)$$

The $C(r)$ expression in (6) confirms that the optimal reflection level, if exists, must be negative, since $C(r)$ is increasing in $r > 0$. Taking derivative on $C(r)$, and applying the variable change, we have

$$C'(r) = \gamma[C(r) - h(r)].$$

Hence, the optimal r can be obtained from

$$h(r) = C(r) \qquad (8)$$

The optimal solution r must exist and be strictly negative. To see this, first observe from (6) that $C(0) > h(0)$. So, the equation in (8) must have a finite and strictly negative solution (denoted as $r^* < 0$), unless $C(r) > h(r)$ for all $r < 0$. But then, this means $C'(r) > 0$, i.e., $C(r)$ is increasing in $r < 0$, which, via (6), contradicts the fact that $h(r)$ is increasing to $+\infty$ as $r \to -\infty$. Furthermore, taking into account $C(r^*) = h(r^*)$, it is direct to verify $C(r) \geq C(r^*)$ for $r \leq r^*$. To summarize, we have

PROPOSITION 1. *The reflection control Y_{r^*} is optimal among all controls in the sub-class $\mathcal{A}^*$, with the optimal reflection level r^* being the solution to $C(r) = h(r)$.*

What remains is to argue that the reflection control Y_{r^*} is not only optimal within the sub-class $\mathcal{A}^*$ of all reflection controls but also optimal over all admissible controls in $\mathcal{A}$.

To this end, for any admissible control $Y(t)$, consider another control, $\tilde{Y}_r(t) := Y_r(t) \wedge Y(t)$. It is then readily verified that (a) $\tilde{Y}_r(t)$ is an admissible control, and (b) $\tilde{Y}_r(t)$ yields a lower cost objective than $Y(t)$, where the reflection level $r \geq 0$ is fixed arbitrarily. Thus, it suffices to show (with details spelled out in the full paper):

$$\mathrm{AC}(x, \tilde{Y}_r) \geq \mathrm{AC}(x, Y_{r^*})[= C(r^*)]. \qquad (9)$$

Consequently, we have the following theorem.

THEOREM 2. *The reflection control Y_{r^*} specified in Proposition 1 is optimal over all controls in the admissible class $\mathcal{A}$, i.e., $AC(x, Y_{r^*}) \leq AC(x, Y)$ for any initial state x and any $Y \in \mathcal{A}$.*

4. ASYMPTOTIC OPTIMALITY

Consider a discrete version of the production-inventory model outlined in §2, i.e., with both demand and production processes being renewal counting processes. Let $\{u_i, \, i = 1, 2, \cdots\}$ denote the inter-arrival times of the orders (demand), an i.i.d. sequence with $\mathsf{E}u_1 = 1/\lambda$ and the squared coefficient of variation c_e^2. Let $\{v_i, \, i = 1, 2, \cdots\}$ denote the required processing times of the orders, another i.i.d. sequence with $\mathsf{E}v_1 = 1/\mu$ and the squared coefficient of variation c_s^2. Assume the two sequences, $\{u_i, i = 1, 2, \cdots\}$ and $\{v_i, i = 1, 2, \cdots\}$, are independent; and let $E(t)$ and $S(t)$ denote the corresponding counting processes.

Let $T(t)$ denote the cumulative amount of time production is active (with processing orders) up to time t. Let $U(t) = t - T(t)$ be the cumulative inactive (idle) time. Let $Q(t)$ denote the state of the system at time t, the difference between the number of orders that have arrived and the number of completed products by time t. Then, the dynamics of the system can be written as follows:

$$Q(t) = Q(0) + E(t) - S(T(t)), \quad t \geq 0. \tag{10}$$

For the above system, reflection control means, whenever the level of inventory reaches a certain level, $Q(t) = r$, for some negative (integer) r, production will be stopped; i.e., $T(t) = \int_0^t \mathbf{1}[Q(s) > r]ds$.

We want to show that applying reflection control to the above discrete production-inventory system is *asymptotically* optimal in a precise sense to be spelled below. Consider a sequence of systems as described above, indexed by a superscript "(n)", with the n-th system having arrival rate $\lambda^{(n)}$, while the service rate μ stays fixed. Assume the following limit.

$$\sqrt{n}(\lambda^{(n)} - \mu) \to \theta < 0. \tag{11}$$

When $n \to \infty$, the above implies $\lambda^{(n)} \to \mu$, from *below*. Thus, when n is large, the above alludes to a heavily utilized system, with production capacity (μ) near saturation. Accordingly, we scale time t by n and space by $1/\sqrt{n}$ in all processes involved (along with proper centering):

$$\begin{aligned} \hat{E}^{(n)}(t) &:= \frac{1}{\sqrt{n}}(E^{(n)}(nt) - \lambda^{(n)}nt), \\ \hat{S}^{(n)}(t) &:= \frac{1}{\sqrt{n}}(S^{(n)}(nt) - \mu nt), \end{aligned}$$

and

$$\hat{U}^{(n)}(t) := \frac{1}{\sqrt{n}}U^{(n)}(nt), \quad \hat{Q}^{(n)}(t) := \frac{1}{\sqrt{n}}Q^{(n)}(nt).$$

Then, the dynamics of the n-th system can be written as,

$$\hat{Q}^{(n)}(nt) = \hat{Q}^{(n)}(0) + \hat{X}^{(n)}(t) + \hat{Y}^{(n)}(t), \tag{12}$$

with

$$\begin{aligned} \hat{X}^{(n)}(t) &= \hat{E}^{(n)}(t) - \hat{S}^{(n)}(\bar{T}^{(n)}(t)) + \sqrt{n}(\lambda^{(n)} - \mu)t, \\ \hat{Y}^{(n)}(t) &= \mu\hat{U}^{(n)}(t), \quad \bar{T}^{(n)}(t) = \frac{1}{n}T^{(n)}(nt). \end{aligned}$$

Applying the standard approach to the diffusion limit of a single-server queue under heavy traffic, we have the following proposition.

Proposition 3. *Under the condition in (11), along with $\hat{Q}^{(n)}(0) \Rightarrow z_0$, and applying reflection control to the n-th system with $\sqrt{n}r$ being the reflection level and r any given negative integer, we have, as $n \to \infty$, the following weak convergence (denoted $\Rightarrow$):*

$$\hat{X}^{(n)}(t) \Rightarrow X(t) := \sigma B(t) + \theta t, \quad \hat{Y}^{(n)}(t) \Rightarrow Y_r(t); \tag{13}$$

where $\sigma^2 = \lambda c_e^2 + \mu c_s^2$, θ is the constant in (11), and Y_r is the reflection control in (4). Hence,

$$\hat{Q}^{(n)}(t) \Rightarrow Z_r(t) := z_0 + X(t) + Y_r(t). \tag{14}$$

Next, consider any admissible control applied to the n-th system, under the above time-space scaling. Admissibility means $\hat{Y}^{(n)}(t)$ must be increasing in $t \in [0, +\infty)$, with $\hat{Y}^{(n)}(0) = 0$, and non-anticipative. Similar to the diffusion limits in the above proposition, we can show that along some subsequence of n, $\hat{Y}^{(n)}(t)$ will converge to a weak limit, denoted $Y(t)$, thereby taking the corresponding $\hat{Q}^{(n)}(t)$ to a weak limit as well, denoted $Z(t)$, and with

$$Z(t) := z_0 + X(t) + Y(t),$$

in parallel to the (Z_r, Y_r) relation in (14). Note, in particular, here $X(t)$ remains the same as (13), as it involves primitive data only.

Then, from Theorem 2 and Proposition 3, we have

Theorem 4. *Applying reflection control to the n-th system as described above, with $\sqrt{n}r^*$ being the reflection level and r^* specified in Proposition 1, is asymptotically optimal in the sense that its diffusion limit (as $n \to \infty$) yields a long-run average cost $AC(x, Y_{r^*})$ that is no greater than the long-run average cost of the diffusion limit of the same system under any other admissible control.*

5. REFERENCES

[1] Bensoussan, A., Liu, R.H. and S.P. Sethi, Optimality of an (s, S) policy with compound Poisson and diffusion demands: A quasi-variational inequalities approach. *SIAM Journal on Control and Optimization*, **44** (2005), 1650-1676.

[2] Chen, H. and D.D. Yao, *Fundamentals of Queueing Networks: Performance, Asymptotics, and Optimization.* Springer-Verlag, New York, 2001.

[3] Dai, J.G. and D.C. Yao, Brownian inventory models with convex holding cost, part 1: Average-optimal controls. *Stochastic Systems*, **3** (2013), 442-499.

[4] Dai, J.G. and D.C. Yao, Brownian inventory models with convex holding cost, part 2: Discount-optimal controls. *Stochastic Systems*, **3** (2013), 500-573.

[5] Harrison, J.M. and M.I. Taksar, Instantaneous control of Brownian motion. *Mathematics of Operations Research*, **8** (1983), 439-453.

[6] Harrison, J.M., Sellke, T.M. and A.J. Taylor, Impulse control of Brownian motion. *Mathematics of Operations Research*, **8** (1983), 454-466.

[7] Ormeci, M., Dai, J.G. and J. Vande Vate, Impulse control of Brownian motion: The constrained average cost case. *Operations Research*, **56** (2008), 618-629.

[8] Reiman, M.I., Open queueing networks in heavy traffic. *Mathematics of Operations Research*, **9** (1984), 441-458.

[9] Taksar, M.I., Average optimal singular control and a related stopping problem. *Mathematics of Operations Research*, **10** (1985), 63-81.

[10] Wu, J. and X. Chao, Optimal control of a Brownian production/inventory system with average cost criterion. *Mathematics of Operations Research*, **39** (2014), 163-189.

[11] Yao, D.C., Chao, X. and J. Wu, Optimal control policy for a Brownian inventory system with concave ordering cost. *Journal of Applied Probability*, **52** (2015), 909-925.