https://doi.org/10.1287/opre.2019.1935

On the Optimality of Reflection Control

Jiankui Yang *
School of Science
Beijing University of Posts and Telecommunications, Beijing, China

David D. Yao†
Dept of Industrial Engineering and Operations Research
Columbia University, New York, USA

Heng-Qing Ye ‡
Faculty of Business
Hong Kong Polytechnic University, Hong Kong, China

October 2018

Abstract

We study the control of a Brownian motion with a negative drift, so as to minimize a long-run average cost objective. We show the optimality of the reflection control, which prevents the Brownian motion from dropping below a certain level by cancelling out from time to time part of the negative drift; and this optimality is established for a holding cost function that appears more general than what's allowed in prior studies. We also show the optimal reflection level can be derived as the fixed point that equates the long-run average cost to the holding cost. Furthermore, we show the *asymptotic* optimality of this reflection control when it is applied to a discrete production-inventory system driven by (delayed) renewal processes; and we do so via identifying the limiting regime of the system under diffusion scaling. In the case of controlling a birth-death model, we establish the optimality of the reflection control directly via a linear programming based approach.

Keywords: reflection control, Brownian motion, diffusion limit, production-inventory system, birth-death queue.

1 Introduction

Consider the control of a Brownian motion with a negative drift, so as to minimize a long-run average cost objective. We show the optimality of a class of so-called reflection controls, which

*Supported in part by NSFC No. 11541006.
†Supported in part by NSF Grant CMMI-1462495.
‡Supported in part by HK/RGC Grants 15508114 and T32-102/14N.

prevents the Brownian motion from dropping below some level r, by cancelling out from time to time part of the negative drift; and this optimality is established for a holding cost function $h(x)$ that is increasing in $x \geq 0$ and decreasing in $x \leq 0$, where x is the state variable. (Throughout the paper, "increasing" and "decreasing" are in the non-strict sense, meaning non-decreasing and non-increasing, respectively.) This is a natural and desirable form of a cost function, since in applications the absolute value of the state variable, $|x|$, can be interpreted as finished-goods inventory or backordered demand (depending on the sign of x), both incurring costs. Furthermore, let $C(r)$ be the long-run average cost with r as the reflection level. We show the optimal reflection level can be derived as the fixed point r^* that equates the long-run average to the holding cost, $C(r^*) = h(r^*)$.

To prove the optimality of the reflection control, we first focus on a subclass (reflection control) and obtain the optimal reflection control for this class; we then show that the optimal reflection control leads to a cost objective that is the lowest among the broader class of all admissible controls, and is hence optimal. This is reminiscent of the lower-bound approach of Harrison and Taksar [5] (also see Harrison [6] and Taksar [9]).

We establish the optimality of reflection control for a class of holding cost function that is exponentially bounded, $h(x) \leq ae^{bx}$ for $x \geq 0$ and for some $a, b > 0$; and we specify explicitly how large b can be. This class appears to be more general than what's allowed in prior studies, which mostly assume polynomially bounded holding costs.

When applied to a discrete production-inventory system, reflection control cannot be optimal in general. This perhaps explains why most studies in the literature that apply reflection control (including its two-sided version such as the (s, S) policy) to production-inventory systems model the latter as Brownian motion. Refer to [1, 3, 4, 7, 8, 10, 11], among many others.

Here we show, for a discrete production-inventory system driven by renewal counting processes, reflection control is *asymptotically* optimal in a limiting regime. Specifically, we show the diffusion limit of the system under a reflection control, with the reflection level optimized, will yield a long-run average cost that is no greater than the same cost corresponding to the diffusion limit of the system under any other admissible control.

It turns out that when the interarrival times (of demand) and service times (to produce the goods) follow exponential distributions, the results can be significantly enhanced. First, we can allow *state-dependent* arrival and service rates, as in a birth-death queue model. Second, the optimality of reflection control can be *directly* established, without the need to go through diffusion scaling and its limiting regime; hence, the optimality so established is without the qualifier "asymptotic". Our approach is linear programming, revealing the problem structure and hence, the optimality of reflection control, through examining duality and complementarity. The analysis leads to a simple algorithm that identifies the optimal reflection level and demonstrates many features of the model that are analogous to those in the Brownian setting.

The rest of the paper consists of three sections studying the reflection control, respectively, in the Brownian setting, in a discrete production-inventory system, and in a birth-death queue model as outlined above. The three sections/topics involve three different approaches: martingales and Itô's calculus for the first one, diffusion limits and heavy-traffic theory for the second, and linear programming for the third.

Thus, our contribution to the literature is a thorough understanding of the reflection control: its optimality in both the Brownian setting and the birth-death queue, and the connections and similarities between the two, their continuous and discrete natures notwithstanding; and its asymptotic optimality in a discrete, non-Markovian system driven by renewal processes, through a Brownian motion limiting regime. And in all cases, we allow an exponentially bounded cost function, which appears to be more general than what's allowed in prior studies.

2 The Brownian 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 $Z(t)$ follows the dynamics,

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

with z_0 being the initial state, and the following long-run average cost is minimized:

$$\mathrm{AC}(x, Y) = \limsup_{t \to +\infty} \mathsf{E}_x \frac{1}{t} \int_0^t h(Z(u)) \mathsf{d}u, \quad \text{where} \quad x := z_0. \tag{2}$$

Note, we have $\mathrm{AC}(x, Y) = \mathsf{E}h[Z(+\infty)]$, provided the control Y induces a steady-state distribution embodied by $Z(+\infty)$. On the other hand, $\mathrm{AC}(x, Y)$ is well defined even if the steady-state distribution does not exist.

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 to t becomes $\mu[t - U(t)]$; and, with $Y(t) = \mu U(t)$, the state process can be expressed as follows:

$$Z(t) = z_0 + (\lambda - \mu)t + \mu U(t) + \sigma B(t) = z_0 + X(t) + Y(t). \tag{3}$$

Note that $Z(t)$, when positive, represents the volume of waiting orders; when negative, its absolute value represents the volume of products waiting to supply demand (i.e., inventory). This motivates the following assumption on the cost function.

- The holding cost $h(x)$ is increasing in $x \geq 0$ and decreasing in $x \leq 0$. This implies $h(x) \geq h(0)$ for all x, and $h(0) \geq 0$ is assumed to be finite. Further assume that $h(x)$ is strictly increasing at $0+$ and strictly decreasing at $0-$. (This rules out trivial case like $h(x) = h(0)$ for all x — no charge for any volume of waiting orders or any amount of inventory.)

- As it will become evident later, for the long-run average cost objective in (2) to be finite requires h to be exponentially bounded; i.e., there exist some parameters $a, b > 0$, such that

$$h(x) \le a e^{bx}, \quad \forall x \ge 0. \tag{4}$$

Note, the results below will require specific ranges for b, but no restriction whatsoever on a. Indeed, the above requirement is equivalent to $\limsup_{x\to+\infty} h(x)/e^{bx} < +\infty$ for some $b > 0$.

There's no cost to carry out the control $Y(t)$. Yet, we do have restrictions on the control: it can only cancel out, from time to time, the negative drift of $X(t)$ (so as to prevent the state process from dropping too much). More formally, we have

- *Set of admissible controls* $\mathcal{A}$. To be admissible, a control $Y(t)$ must be non-anticipative, and increasing in $t \in [0, +\infty)$, with $Y(0) \ge 0$.

2.1 Reflection Control

Recall, a Brownian motion 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 the corresponding state process as Z_r. The control Y_r is defined by a reflection level r, meaning that it ensures $Z_r(t) \ge r$ 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:

$$Y_r(t) = \sup_{0\le u\le t} (r - z_0 - X(u))^+, \tag{5}$$

$$Z_r(t) = z_0 + X(t) + \sup_{0\le u\le t} (r - z_0 - X(u))^+. \tag{6}$$

In addition, *complementarity* holds: $[Z_r(t) - r]\mathsf{d}Y_r(t) = 0$ for all $t \ge 0$, i.e., when $Z_r(t) > r$, $Y_r(t)$ cannot increase. Furthermore, $Z_r(+\infty) - r$ follows an exponential distribution with rate $-2\theta/\sigma^2$ (recall, $\theta < 0$). Thus, under the reflection control Y_r, we have

$$\mathrm{AC}(x, Y_r) = \mathsf{E}_x h(Z_r(+\infty)) = \gamma \int_0^{+\infty} h(r + x) e^{-\gamma x} \mathsf{d}x := C(r), \quad \text{where} \quad \gamma := -2\theta/\sigma^2. \tag{7}$$

From the above, we can write $C(r) = \mathsf{E}h(r + \xi)$, where ξ follows an exponential distribution with parameter (rate) γ. (Thus, $r + \xi$ follows a shifted exponential distribution, with a density function $\gamma e^{-\gamma(x-r)}$ for $x \ge r$.)

For $r < 0$, we can derive

$$C(r) = \left(\int_0^{\infty} + \int_r^0\right) h(x)\gamma e^{-\gamma(x-r)}\mathsf{d}x = e^{\gamma r} C(0) + e^{\gamma r}\int_r^0 h(x)\gamma e^{-\gamma x}\mathsf{d}x. \tag{8}$$

Since $h(0) < h(x) \leq h(r)$ for $x \in [r, 0)$, the second term on the right side above is (strictly) bounded between $h(0)(1-e^{\gamma r})$ and $h(r)(1-e^{\gamma r})$. Thus, we have

$$e^{\gamma r} C(0) + h(0)(1-e^{\gamma r}) < C(r) < e^{\gamma r} C(0) + h(r)(1-e^{\gamma r}), \quad \forall r < 0. \tag{9}$$

To minimize $C(r)$, taking derivative and applying integration by parts, we have

$$C'(r) = \gamma[C(r) - h(r)]. \tag{10}$$

Since $C(0) = \mathsf{E}h(\xi) > h(0)$, we know the minimizer r^* must be negative, and can be obtained as follows:

$$C(r) = h(r) \quad \Rightarrow \quad r^*. \tag{11}$$

In the case of multiple solutions to the above equation, they must all yield the same $C(\cdot)$ value (and locate next to each other to form a contiguous interval); hence, any one of them can be designated as r^*; see the remarks in (ii) below.

There are two cases:

(1) There exists some $r_0 < 0$ such that $h(r_0) \geq \mathsf{E}h(\xi) = C(0)$. From the upper bound in (9), we have

$$C(r_0) < h(r_0) - e^{\gamma r}[h(r_0) - C(0)] \leq h(r_0).$$

This, along with $C(0) > h(0)$, implies there must exist a solution to the equation in (11), and $r^* \in (r_0, 0)$.

(2) On the other hand, if the equation has no solution, then it means $C(r) > h(r)$ for all $r < 0$, i.e., $C'(r) \geq 0$, or $C(r)$ increasing for all $r < 0$. In this case, setting $r^* = -\infty$ is optimal.

Several remarks are in order.

(i) From Case (1) we know, $C(0) \leq h(-\infty)$ implies r^* is finite (since $h(-\infty) \geq h(r_0)$ for any negative r_0). The contrapositive of this is: $r^* = -\infty$ implies $C(0) > h(-\infty)$. Yet, when $C(0) > h(-\infty)$, we may still have a finite r^*; refer to Example 3 below.

(ii) From the above analysis, $C(r)$ is clearly increasing in $r \geq r^*$, including when $r^* = -\infty$. When r^* (< 0) is finite, we can verify that $C(r)$ is decreasing in $r \leq r^*$. (Note this will not necessarily follow from r^* being a minimizer of $C(r)$, as it may be a local minimizer.) To do so, similar to (8), but breaking at r^* instead of 0, we have

$$C(r) = e^{-\gamma(r^* - r)} C(r^*) + e^{\gamma r} \int_r^{r^*} h(x) \gamma e^{-\gamma x} dx, \qquad r \leq r^* < 0.$$

Since $C(r^*) = h(r^*) \leq h(r)$, and $h(x) \leq h(r)$ for $x \in [r, r^*]$, we have

$$C(r) \leq e^{-\gamma(r^* - r)} h(r) + h(r)\big[1 - e^{-\gamma(r^* - r)}\big] = h(r), \qquad r \leq r^* < 0.$$

From (10), we then have $C'(r) \leq 0$, or $C(r)$ decreasing, for all $r \leq r^*$.

(iii) When r^* is finite, for $C(r^*) < \infty$, it suffices to assume that when $x \to +\infty$, $h(x)$ grows no faster than e^{bx} for some $b < \gamma$. This explains why we need the exponential boundedness condition in (4).

(iv) On the other hand, $r^* = -\infty$ means exercise no control, in which case the Brownian motion with a negative drift will go to $-\infty$. Hence, the corresponding long-run average cost in (7) should be, by definition, $C(-\infty) = h(-\infty)$, in which case we must have $h(-\infty) < C(0)$ as argued above in (i). (And $C(0) < \infty$, same as $C(r^*) < \infty$, follows from (4) for $b < \gamma$.) This also explains why there's no need for the exponential boundedness condition to stipulate any restriction on the negative side.

To summarize, we have

Proposition 1 Suppose the condition in (4) holds for some $b < \gamma := -2\theta/\sigma^2$.
(i) 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 (11) if it exists (in which case it must be negative, and in the case of multiple solutions, pick the least negative as r^*); or if the solution does not exist, then the optimal reflection level is $r^* = -\infty$.
(ii) A sufficient condition for r^* to be finite: if for some $r_0 < 0$, $h(r_0) \geq \mathsf{E}h(\xi) = C(0)$, then $r^* \in (r_0, 0)$. On the other hand, if $r^* = -\infty$, then $C(0) > h(-\infty)$.
(iii) $C(r)$ is increasing in $r \geq r^*$ (including when $r^* = -\infty$), and decreasing in $r \leq r^*$ when r^* is finite.

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 $\mathcal{A}$. We defer this to the next subsection, illustrating several examples first. In the first two examples below, $h(-\infty) = +\infty$; thus $C(0) < h(-\infty)$ holds, and a finite r^* exists as in Proposition 1(ii) above. The third example is more interesting. Not only because $h(x)$ is *concave* for $x \geq 0$, which is rarely studied in the literature, it also gives an example that even when $C(0) > h(-\infty)$, a finite r^* may still exist.

Example 1. Let $h(x) = |x|$. By (7) and (11), we have for $r \leq 0$,

$$C(r) = \frac{2e^{\gamma r}}{\gamma} - r - \frac{1}{\gamma}, \quad \text{and} \quad r^* = -\frac{\log 2}{\gamma}.$$

Example 2. Let $h(x) = \mathrm{e}^{b|x|}$ $(0 < b < \gamma)$, and we have for $r \leq 0$,

$$C(r) = \frac{\gamma}{\gamma + b} e^{-br} + \left(\frac{\gamma}{\gamma - b} - \frac{\gamma}{\gamma + b}\right)\mathrm{e}^{\gamma r} \quad \text{and} \quad r^* = \log\left(\frac{1}{2} - \frac{b}{2\gamma}\right)/(\gamma + b).$$

Example 3. Let

$$h(x) = \begin{cases} \kappa(1 - e^x), & x < 0 \\ 1 - e^{-x}, & x \geq 0 \end{cases} \tag{12}$$

where $\kappa > 0$ is a given parameter.

Making use of (8), we can derive

$$C(r) = \frac{e^{\gamma r}}{\gamma+1} + \kappa\Big[(1-e^{\gamma r}) + \frac{\gamma}{\gamma-1}(e^{\gamma r}-e^{r})\Big], \qquad r \le 0.$$

Equating the above to $h(r) = \kappa(1-e^r)$ yields

$$r^* = \frac{1}{1-\gamma}\log\big[1+\frac{\gamma-1}{\kappa(\gamma+1)}\big], \qquad \gamma \ge 1 \text{ or } \kappa > \frac{1-\gamma}{1+\gamma}. \tag{13}$$

Note, the above yields a finite $r^* < 0$ for both $\gamma < 1$ (with the necessary restriction on κ) and $\gamma > 1$, and also $r^* = -\frac{1}{2\kappa}$ for $\gamma = 1$ via l'Hôspital. In particular, when $\kappa \ge 1$, we have $h(-\infty) = \kappa > C(0) = 1/(\gamma+1)$, confirming the existence of a finite r^* as according to Proposition 1 (ii). Furthermore, (13) also includes the case $1/(1+\gamma) > \kappa$, i.e., $C(0) > \kappa = h(-\infty)$. Thus, while $C(0) < h(-\infty)$ is a sufficient condition for the existence of a finite r^*, it is not necessary.

2.2 Optimality

We now extend the optimality of $Y_{r^*}(t)$ as stated in Proposition 1 to the class of all admissible controls, $\mathcal{A}$. To facilitate exposition, below we shall focus on the case of a *finite* optimal reflection level r^*. After completing the proof (of Theorem 2 below), the degenerate case, $r^* = -\infty$ will be briefly discussed.

To start with, for any admissible control $Y(t)$, with the corresponding state being $Z(t) = X(t) + Y(t)$, consider another control, $\tilde{Y}_r(t) := Y_r(t) \wedge Y(t)$, where Y_r is a reflection control, with $r \ge 0$. Clearly, $\tilde{Y}_r(t)$ is non-anticipative, nondecreasing and right continuous with left limit, and thus is an admissible control. Write the state process under this control as

$$\tilde{Z}_r(t) := X(t) + \tilde{Y}_r(t), \quad \text{and} \quad \tilde{Z}_r(t) = Z_r(t) \wedge Z(t). \tag{14}$$

Write the cost under this control as

$$h(\tilde{Z}_r(t)) = h(\tilde{Z}_r(t))\mathbf{1}\{Z(t) < r\} + h(\tilde{Z}_r(t))\mathbf{1}\{Z(t) \ge r\}.$$

When $Z(t) < r$ $(\le Z_r(t))$, we have $\tilde{Z}_r(t) = Z(t)$, and hence,

$$h(\tilde{Z}_r(t))\mathbf{1}\{Z(t) < r\} = h(Z(t))\mathbf{1}\{Z(t) < r\}.$$

On the other hand, when $Z(t) \ge r$, we have $Z(t) \ge \tilde{Z}_r(t) \ge r(\ge 0)$. Since $h(x)$ is increasing in $x \ge 0$, we have

$$h(\tilde{Z}_r(t))\mathbf{1}\{Z(t) \ge r\} \le h(Z(t))\mathbf{1}\{Z(t) \ge r\}.$$

Thus, we know $h(\tilde{Z}_r(t)) \le h(Z(t))$, which implies $\mathrm{AC}(x, \tilde{Y}_r) \le \mathrm{AC}(x, Y)$. So, it suffices to show, for any initial state x,

$$\mathrm{AC}(x, \tilde{Y}_r) \ge \nu^* := \mathrm{AC}(x, Y^*) \ \ [= h(r^*) = C(r^*)]. \tag{15}$$

To do so, consider a smooth, auxiliary function $V(z)$, to be specified later. Applying Itô's formula to $V(\tilde{Z}_r(t))$, we have

$$\begin{aligned} V(\tilde{Z}_r(t)) &= V(x) + \int_0^t \big[\frac{\sigma^2}{2} V''(\tilde{Z}_r(u)) + \theta V'(\tilde{Z}_r(u))\big] \mathsf{d}u \\ &\quad + \sigma \int_0^t V'(\tilde{Z}_r(u)) \mathsf{d}B(u) + \int_0^t V'(\tilde{Z}_r(u)) \mathsf{d}\tilde{Y}_r(u). \end{aligned} \tag{16}$$

For reasons that will become evident shortly, we want to have

$$\frac{\sigma^2}{2} V''(z) + \theta V'(z) = \nu^* - h(z). \tag{17}$$

Taking into account $C'(z) = \gamma[C(z) - h(z)]$ following (10), and along with the γ expression in (7), the above is implied by

$$V'(z) = -\frac{1}{\theta}[C(z) - \nu^*], \tag{18}$$

which, in turn, is implied by

$$V(z) = -\frac{1}{\theta}\Big(\int_0^z C(u)\mathsf{d}u - \nu^* z \Big), \tag{19}$$

When $z < 0$, the above integral is interpreted as $\int_0^z C(u)\mathsf{d}u = -\int_z^0 C(u)\mathsf{d}u$. Clearly, $V'(z) \geq 0$, since $\theta < 0$ and $C(z) \geq \nu^* := C(r^*)$; hence, the last integral on the right side of (16) is non-negative, taking into account $\mathsf{d}\tilde{Y}_r \geq 0$. Hence,

$$V(\tilde{Z}_r(t)) \geq V(x) + \int_0^t (\nu^* - h(\tilde{Z}_r(u)))\mathsf{d}u + \sigma \int_0^t V'(\tilde{Z}_r(u))\mathsf{d}B(u). \tag{20}$$

The last term on the right side, being Itô's integral, is a local martingale; and with the square-integrability condition, to be established shortly below,

$$\mathsf{E}\left[\int_0^t \Big(V'(\tilde{Z}_r(u)) \Big)^2 \mathsf{d}u \right] = \int_0^t \mathsf{E}\Big(V'(\tilde{Z}_r(u)) \Big)^2 \mathsf{d}u < \infty, \quad \text{for} \quad \text{every } t \geq 0, \tag{21}$$

it is a martingale, implying

$$\mathsf{E} \int_0^t V'(\tilde{Z}_r(u))\mathsf{d}B(u) = 0.$$

Hence, taking expectations on both sides of (20) yields

$$\mathsf{E}V(\tilde{Z}_r(t)) \geq V(x) + \mathsf{E}\int_0^t \big[\nu^* - h(\tilde{Z}_r(u))\big]\mathsf{d}u. \tag{22}$$

Dividing both sides by t and letting $t \to \infty$, we have

$$\limsup_{t\to\infty} \frac{1}{t} \mathsf{E}V(\tilde{Z}_r(t)) \geq \nu^* - \mathrm{AC}(x, \tilde{Y}_r). \tag{23}$$

If we can further show

$$\limsup_{t\to\infty} \frac{1}{t}\mathsf{E}V(\tilde{Z}_r(t)) = 0, \tag{24}$$

then we will have established the desire inequality $\nu^* \le \mathrm{AC}(x, \tilde{Y}_r)$. The precise statement is as follows.

Theorem 2 Suppose the cost function $h(\cdot)$ satisfies the exponential boundedness condition in (4) with $b < \gamma/2$. Then, the optimal reflection control $Y_{r^*}(t)$ as specified in Proposition 1 is optimal over all admissible controls, i.e., $\mathrm{AC}(x, Y_{r^*}) \le \mathrm{AC}(x, Y)$ for any admissible control $Y \in \mathcal{A}$ and from any initial state x.

Proof. As discussed above, what remains is to show (21) and (24). First note (to become evident shortly), we will never need $h(x)$ for $x < r^*$ (< 0). So, we can consider a modified cost function, for any given $r_* < r^*$, define

$$\tilde{h}(z) := \begin{cases} h(r_*) & z < r_*, \\ h(z) & z \ge r_*, \end{cases} \tag{25}$$

Clearly, $\tilde{h}(z) \le h(z)$ for all z. Hence, replacing h by $\tilde{h}$ will reduce $\mathrm{AC}(x, \tilde{Y}_r)$, whereas ν^* will remain unchanged. So, it suffices to establish the desired inequality $\nu^* \le \mathrm{AC}(x, \tilde{Y}_r)$ under this modified cost function. Thus, to simplify notation, we shall continue to denote the cost function as h, which takes up the definition of $\tilde{h}$ in (25) throughout below.

Second, recall $Z_r(t)$ is a RBM, reflected at r. Consider another RBM, $Z_{r'}(t)$, reflected at $r' := Z_r(0) \vee r$ and started at $Z_{r'}(0) = r'$. Then, $Z_r(t) \le Z_{r'}(t)$; and $Z_{r'}(t) - r'$ is known to be equal in distribution to $\sup_{s\le t} X(s)$, the "running max". The latter is increasing (in t) to a limiting distribution that is the same as the exponential distribution of ξ. Hence,

$$\mathsf{E}e^{2bZ_r(t)} \le \mathsf{E}e^{2bZ_{r'}(t)} \le e^{2br'}\mathsf{E}e^{2b\xi} = \frac{\gamma e^{2br'}}{\gamma - 2b} < \infty. \tag{26}$$

We are now ready to prove (21). Write $C(\tilde{Z}_r(t)) = \mathsf{E}[h(\tilde{Z}_r(t)+\xi)|\tilde{Z}_r(t)]$, where recall $\tilde{Z}_r$ follows (14), and ξ follows the exponential distribution with parameter γ and is independent of $\tilde{Z}_r$ (and Z_r). From the $V'(\cdot)$ expression in (18), it suffices to argue that $\mathsf{E}C^2(\tilde{Z}_r(t))$ is bounded by a constant (i.e., independent of t). Write

$$\begin{aligned} \mathsf{E}C^2(\tilde{Z}_r(t)) &= \mathsf{E}\big[\mathsf{E}[h(\tilde{Z}_r(t)+\xi)|\tilde{Z}_r(t)]\big]^2 \le \mathsf{E}\big[\mathsf{E}[h^2(\tilde{Z}_r(t)+\xi)|\tilde{Z}_r(t)]\big] = \mathsf{E}h^2(\tilde{Z}_r(t)+\xi) \\ &= \mathsf{E}\big[h^2(\tilde{Z}_r(t)+\xi)\mathbf{1}\{\tilde{Z}_r(t)+\xi > 0\}\big] + \mathsf{E}\big[h^2(\tilde{Z}_r(t)+\xi)\mathbf{1}\{\tilde{Z}_r(t)+\xi \le 0\}\big]. \end{aligned} \tag{27}$$

Note the second expectation on the right side is dominated by $h^2(r_*)$, following the (modified) cost function definition in (25); and the first expectation is dominated by $\mathsf{E}h^2(Z_r(t)+\xi)$, taking into account $\tilde{Z}_r(t) \le Z_r(t)$ and $h(z)$ increasing in $z > 0$; furthermore,

$$\mathsf{E}h^2(Z_r(t)+\xi) \le a^2\mathsf{E}e^{2b(Z_r(t)+\xi)} < \infty, \tag{28}$$

where the first inequality makes use of $h(x) \leq ae^{bx}$ for $x \geq 0$ — tha assumption in (4), along with the fact that $Z_r(t) \geq r \geq 0$; the second inequality follows from (26). Thus, both terms on the right side of (27) can be bounded by constants; and hence, (21) is proven.

To show (24), it suffices to show $\sup_{t\geq 0} \mathsf{E}V(Z_r(t)) < \infty$ taking into account $\tilde{Z}_r(t) \leq Z_r(t)$ and $V(z)$ increasing. From (19), there are two terms involved:

$$\mathsf{E}V(Z_r(t)) = -\frac{1}{\theta}\left(\mathsf{E}\int_0^{Z_r(t)} C(u)\mathsf{d}u - \nu^*\mathsf{E}Z_r(t)\right).$$

For the second term, we have $Z_r(t) \geq 0$ (as $r \geq 0$) and therefore $\frac{\nu^*}{\theta}\mathsf{E}Z_r(t) \leq 0$. So, we only need to show the first term is finite, i.e.,

$$\sup_{t\geq 0} \mathsf{E}\int_0^{Z_r(t)} C(u)\mathsf{d}u < \infty. \tag{29}$$

To this end, we have

$$\mathsf{E}\int_0^{Z_r(t)} C(u)\mathsf{d}u = \mathsf{E}\int_0^{Z_r(t)} \mathsf{E}h(u+\xi)\mathsf{d}u \leq a\mathsf{E}\int_0^{Z_r(t)} \mathsf{E}e^{b(u+\xi)}\mathsf{d}u = \frac{a}{b}\mathsf{E}e^{b\xi}\mathsf{E}(e^{bZ_r(t)} - 1),$$

where the first equality takes into account $C(u) = \mathsf{E}h(u+\xi)$ and the inequality makes use of (4). Then, the bound in (29) follows from the property in (26), which is clearly valid with $2b$ being replaced by b. □

Remarks.

- The case of $r^* = -\infty$ is readily accommodated in the above proof. First, in this case (refer to (15)), $\nu^* = C(-\infty) = h(-\infty) < C(0) < \infty$, as argued in the remark (iv) preceding Proposition 1. Second, in this case no need to modify the holding cost function h, i.e., $\tilde{h}(z) = h(z)$ in (25). All other arguments in the proof remain intact.
- Note the range for b in Theorem 2 is different from the one in Proposition 1; specifically, for the reflection control Y_{r^*} to be optimal among all reflection controls, $b < \gamma$ is sufficient; for Y_{r^*} to be optimal among all admissible controls, the range is shrunk by half, to $b < \gamma/2$.

3 Reflection Control of a Discrete Production-Inventory System

Next, we consider a discrete version of the production-inventory model outlined in §2, i.e., with both demand and production following (delayed) renewal counting processes. We want to show that applying reflection control to this discrete system is *asymptotically* optimal. Specifically, considering a sequence of such systems indexed by n under the usual diffusion scaling, we show the limiting regime of the systems under the reflection control will yield a long-run average cost that is minimal among all admissible controls.

Let's start by describing the dynamics of the discrete system. Denote by u_1 the residual arrival time initially. Let $\{u_i, i = 2, 3, \cdots\}$ denote the interarrival times of the orders (demand), an i.i.d. sequence with $\mathsf{E}(u_2) = 1/\lambda$ and the squared coefficient of variation c_a^2. Denote by v_1 the residual processing time initially. Let $\{v_i, i = 2, 3, \cdots\}$ denote the required processing times of the orders, another i.i.d. sequence with $\mathsf{E}(v_2) = 1/\mu$ and the squared coefficient of variation c_s^2. Note, the assumed finite second moments of the primitives implies these processes are uniformly integrable, which is sufficient for the weak convergence below.

Assume the two sequences, $\{u_i\}$ and $\{v_i\}$, are independent; and let $A(t)$ and $S(t)$ denote the corresponding (delayed) counting processes:

$$A(t) = \max\{i : \sum_{j=1}^{i} u_j \le t\}, \quad S(t) = \max\{i : \sum_{j=1}^{i} v_j \le t\}.$$

Let $T(t)$ denote the cumulative amount of time production is active (with processing orders) up to time t. 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. That is, $Q(t)$ is the net demand at time t, the discrete counter-part of $Z(t)$ in the Brownian model. Then, the dynamics of the system can be written as follows:

$$Q(t) = Q(0) + A(t) - S(T(t)), \quad t \ge 0. \tag{30}$$

For the above system, reflection control means, whenever the level of inventory reaches a certain level, $Q(t) = r$, for some integer r, production will be stopped; i.e., $T(t) = \int_0^t \mathbf{1}\{Q(s) > r\}ds$. Specifically, the reflection control is $Y(t) := \mu(t - T(t))$. Note, under the reflection control, $Q(t) - r$ coincides with the state process of a single-server queue.

Next, consider a sequence of systems as described above, indexed by a superscript "(n)". Let $\rho^{(n)} < 1$ be a sequence of scaling parameters (to be further specified below). Let $u_i^{(n)} = u_i/\rho^{(n)}$ be the i-th ($i = 2, 3, \dots$) interarrival time to the n-th system, and accordingly, let $\lambda^{(n)} = \rho^{(n)}\lambda$. Note this scaling will not affect the squared coefficient of variation c_a^2. No change in all other primitives: the service times $v_i^{(n)} = v_i$ ($i = 2, 3, \cdots$) stay fixed (i.e., remain the same among all systems in the sequence), and hence, so do μ and c_s^2.

To carry out the analysis below, we need to assume the so-called "heavy traffic" condition. First, assume $\lambda = \mu$. (Note, since $\rho^{(n)} < 1$, we have $\lambda^{(n)} < \mu$, i.e., every system in the sequence is still stable.) Second, assume the scaling satisfies the following limit:

$$\theta^{(n)} := \sqrt{n}(\lambda^{(n)} - \mu) \to \theta < 0. \tag{31}$$

Clearly, the above implies $\lambda^{(n)} \to \mu$, and hence $\rho^{(n)} \to 1$, both from below. (Again, this ensures the stability of every system in the sequence.) Thus, when n is large, the n-th system is heavily utilized, with the production capacity near saturation.

The so-called diffusion scaling is to scale time by n and space by $1/\sqrt{n}$ in all processes involved (along with proper centering):

$$\left(\hat{Q}^{(n)}(t),\, \hat{Y}^{(n)}(t)\right) := \frac{1}{\sqrt{n}}\left(Q^{(n)}(nt),\, \mu(nt - T^{(n)}(nt))\right),$$

$$\left(\hat{A}^{(n)}(t),\, \hat{S}^{(n)}(t)\right) := \frac{1}{\sqrt{n}}\left(A^{(n)}(nt) - \lambda^{(n)}nt,\, S^{(n)}(nt) - \mu nt\right),$$

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

$$\hat{Q}^{(n)}(t) = \hat{Q}^{(n)}(0) + \hat{X}^{(n)}(t) + \hat{Y}^{(n)}(t), \tag{32}$$

with

$$\hat{X}^{(n)}(t) = \hat{A}^{(n)}(t) - \hat{S}^{(n)}(\bar{T}^{(n)}(t)) + \theta^{(n)}t \quad \text{and} \quad \bar{T}^{(n)}(t) = \frac{1}{n}T^{(n)}(nt).$$

To apply reflection control to the n-th (discrete) system, let $r^{(n)} := \lfloor \sqrt{n}r \rfloor$ be the reflection level, i.e., the server will stop producing when $Q^{(n)}(t)$ reaches $r^{(n)}$; and use a subscript r to emphasize the reflection control. Then,

$$Q_r^{(n)}(t) - r^{(n)} \geq 0 \quad \text{and} \quad \big(Q_r^{(n)}(t) - r^{(n)}\big)\mathsf{d}Y_r^{(n)}(t) = 0. \tag{33}$$

From the diffusion limit of a single-server queue (e.g., [2], Chapter 6), we have, under the heavy-traffic condition in (31) and along with $(\hat{Q}^{(n)}(0), u_1^{(n)}/\sqrt{n}, v_1^{(n)}/\sqrt{n}) \Rightarrow (z_0, 0, 0)$ (with $\Rightarrow$ denoting weak convergence):

$$\hat{X}^{(n)}(t) \Rightarrow X(t) := \theta t + \sigma B(t), \qquad \hat{Y}_r^{(n)}(t) \Rightarrow Y_r(t) := \sup_{0 \leq s \leq t} (r - z_0 - X(s))^+; \tag{34}$$

where θ is the constant in (31), $\sigma^2 = \lambda c_a^2 + \mu c_s^2$, $B(t)$ is a standard Brownian motion as usual, and Y_r is the reflection control in (5). Hence,

$$\hat{Q}_r^{(n)}(t) \Rightarrow Z_r(t) := z_0 + X(t) + Y_r(t). \tag{35}$$

Next, apply any admissible control to the n-th system under the above time-space scaling, and drop the subscript r. The control $\hat{Y}^{(n)}(t)$, being non-anticipative and increasing in $t \in [0, +\infty)$ with $\hat{Y}^{(n)}(0) \geq 0$, must have a weak limit, possibly along some subsequence of $\{n\}$. Consequently, for any subsequence of $\{n\}$ there is a further subsequence, also denoted as $\{n\}$, such that

$$(\hat{Q}^{(n)}(t), \hat{X}^{(n)}(t), \hat{Y}^{(n)}(t)) \Rightarrow (Z(t), X(t), Y(t)),$$

where the limit satisfies $Z(t) = z_0 + X(t) + Y(t)$, with $X(t)$ being the same Brownian motion as in (34), and $Y(t)$ another admissible control. Whereas $Z(t)$ may or may not have a stationary distribution, the long-run average cost as defined in (2) will exist (or go to $+\infty$). Following Theorem 2, this long-run average cost must be higher than the one associated with the reflection control Y_r specified above (provided the level of reflection is optimized):

$$\mathrm{AC}(z_0, Y) \geq \mathrm{AC}(z_0, Y_{r^*}) = C(r^*). \tag{36}$$

To summarize, we have

Theorem 3 Suppose the exponential boundedness condition in (4) is satisfied with $b < \gamma/2$. Then, applying reflection control to the n-th system, with $\sqrt{n}r^*$ being the reflection level and r^* specified in Proposition 1, is asymptotically optimal in the sense of (36); i.e., its diffusion limit yields a long-run average cost that is no greater than the long-run average cost of the diffusion limit of the same system under any other admissible control.

4 Reflection Control of the Birth-Death Queue

The last theorem can be significantly enhanced, when the interarrival times and processing times follow exponential distributions. First, we can allow *state-dependent* arrival and service rates, as in a birth-death queue model. Second, the optimality of the reflection control can be *directly* established, without the need to go through diffusion scaling and the Brownian limit. The optimality so established is hence without the qualifier "asymptotic".

Let λ_n and μ_n denote the state-dependent birth/arrival and death/service rates in a birth-death queue, and allow the state (n) to take on all integer values including negative ones. Suppose $\lambda_n > 0$ for all n, and let the death rates $\mu_n \geq 0$ be decision variables.

Suppose there is a given $\mu > 0$, which is the maximal possible service rate (capacity):

$$\mu_n \leq \mu, \qquad \forall n; \tag{37}$$

and

$$\sup_n \lambda_n := \lambda < \mu, \quad \text{or} \quad \rho := \frac{\lambda}{\mu} < 1. \tag{38}$$

This guarantees the existence of a steady-state distribution (for the state process), which is denoted $\pi := (\pi_n)$.

In each state n, there is a per time-unit holding cost $h_n \geq 0$, and (as in the Brownian model) assume h_n is increasing in $n \geq 0$ and decreasing in $n < 0$, with $h_0 < h_1$ and $h_0 < h_{-1}$ (so as to rule out trivial cases).

We want to minimize the long-run average cost:

$$\min_{(\mu_n)} \sum_n h_n \pi_n, \quad \text{s.t.} \quad \lambda_n \pi_n = \mu_{n+1}\pi_{n+1},\ 0 \leq \mu_n \leq \mu; \quad \forall n. \tag{39}$$

Note the above problem formulation, in particular in making (μ_n) as decision variables, allows an admissible control class that is as broad as non-anticipative. Here, any history-dependent control (of the service) amounts to dependence on current state only, because of the Markovian nature of the system, in particular the exponential sojourn times in each state.

Following the approach in [12], we can rewrite the above optimization problem as follows, with $\pi := (\pi_n)$ as decision variables:

$$\min_{\pi} \sum_n h_n \pi_n, \quad \text{s.t.} \quad \mu\pi_{n+1} - \lambda_n \pi_n \geq 0,\ \pi_n \geq 0, \quad \forall n; \quad \sum_n \pi_n \geq 1. \tag{40}$$

(Note that the last constraint ≥ 1 is effectively $= 1$, since if $\pi = (\pi_n)$ is a solution that satisfies $\sum_n \pi_n > 1$, then $\pi_n / \sum_n \pi_n$ is certainly a better solution — yielding a smaller objective value, since $h_n \geq 0$ for all n.)

The above is a linear programming (LP) problem; and its first constraint clearly implies the following: If $\pi_m > 0$ for some m, then $\pi_n > 0$ for all $n \geq m$ (since $\lambda_n > 0$ as assumed); if $\pi_m = 0$ for some m, then $\pi_n = 0$ for all $n \leq m$. This, in turn, implies that the optimal $\mu = (\mu_n)$ is such that there exists some r, such that $\mu_n = 0$ (and $\pi_n = 0$) for all $n \leq r - 1$, and $\mu_n > 0$ (and $\pi_n > 0$) for all $n \geq r$. Furthermore, the optimal r cannot be positive. For if $r > 0$ and $\pi_r > 0$ whereas $\pi_{r-1} = 0$, then we can reduce π_r by an amount $\delta > 0$, sufficiently small so as not to violate the constraints, and increase π_{r-1} to δ. This will decrease the objective value in (40) since $h_r \geq h_{r-1}$ for $r > 0$.

The next question is, whether for all $n \geq r$, the server should serve at full capacity, i.e., whether it's optimal to have $\mu_n = \mu$ for all $n \geq r$. To address this question, we examine the dual of the above LP, with the range of n restricted to $n \geq r$. Let y_n be the dual variable associated with the first set of constraints in (40), and let z be the dual variable associated with the constraint $\sum_n \pi_n \geq 1$. Then, the dual LP is:

$$\max z \quad \text{s.t.} \quad z - \lambda_r y_r \leq h_r; \quad z + \mu y_{n-1} - \lambda_n y_n \leq h_n,\ n \geq r+1; \quad z \geq 0,\ y_n \geq 0,\ \forall n. \tag{41}$$

Complementarity indicates that all the above inequalities, except the non-negativity on z and y_n, should hold as equalities, since $\pi_n > 0$ for all $n \geq r$ in the primal LP. This, in turn, will imply that $y_n > 0$ for all $n \geq 0$. (Note $r \leq 0$ as argued above.) To see this, consider two consecutive constraints,

$$z + \mu y_{n-1} - \lambda_n y_n = h_n, \quad z + \mu y_n - \lambda_{n+1} y_{n+1} = h_{n+1}.$$

If $y_n = 0$, then the above will lead to

$$-\lambda_{n+1} y_{n+1} - \mu y_{n-1} = h_{n+1} - h_n \geq 0, \quad n \geq 0.$$

Since λ_{n+1} and μ are both positive, and y_{n-1} and y_{n+1} both non-negative, the above means we must have $y_{n-1} = y_{n+1} = 0$; and hence, $h_{n+1} = h_n$. This, in turn, will lead to $h_n = h_0$ for all $n \geq 0$, a trivial case that has been ruled out.

The fact that $y_n > 0$ for all $n \geq 0$ means the constraint in the primal LP is binding, $\lambda_{n-1}\pi_{n-1} = \mu\pi_n$ for all $n \geq 0$, i.e., it is optimal to have $\mu_n = \lambda_{n-1}\pi_{n-1}/\pi_n = \mu$ for all $n \geq 0$.

Yet, the above argument does not apply to the negative n's, where h_n is *decreasing*. So, what remains is to pin down the optimal rates in the negative states $j = -1, \ldots, r$. To this end, write the long-run average cost as follows:

$$\pi_0 \sum_{j=-1}^{r} \frac{h_j}{R_j} + \pi_0 \sum_{n \geq 0} h_n \bar{R}_n,$$

where

$$R_j := \frac{\lambda_{-1} \cdots \lambda_j}{\mu_0 \cdots \mu_{j+1}}, \quad j = -1, \ldots r; \quad \bar{R}_0 := 1, \quad \bar{R}_n := \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu^n}; \quad n = 1, 2, \ldots;$$

and

$$\pi_0 := \Big(S_0 + \sum_{j=-1}^{r} \frac{1}{R_j}\Big)^{-1},$$

with

$$S_0 := \sum_{n\geq 0} \bar{R}_n \leq \frac{1}{1-\rho} < \infty,$$

where the inequalities follow from $\bar{R}_n \leq \rho^n$ for $n \geq 0$, with $\rho := \lambda/\mu < 1$; refer to (38).

To determine the optimal rates, write $x_j := 1/R_j$, $x := (x_j)_{j=-1}^r$; and denote $H_0 := \sum_{n\geq 0} h_n \bar{R}_n$. Then, we have

$$\min_x \frac{H_0 + \sum_{j=-1}^r h_j x_j}{S_0 + \sum_{j=-1}^r x_j} \quad \text{s.t.} \quad 0 \leq x_j \leq \frac{1}{\bar{R}_j}, \quad j = -1, \dots, r. \tag{42}$$

Writing the objective function above as $C(x) := f(x)/g(x)$, and taking derivatives, we have

$$\frac{\partial C}{\partial x_j} = \frac{1}{g}\big[h_j - C(x)\big].$$

Thus, if $h_j < C$, then increasing x_j will reduce C. This, along with h_j decreasing in j, leads to a simple algorithm, which also reveals more structure of the optimal solution.

Start with $j = -1$ and compare h_j with $C = H_0/S_0$. If $h_j \geq C$, then any positive x_j will increase C, so we must have $x_j = 0$. On the other hand, if $h_j < C$, then we can increase x_j, and thereby reducing C, since

$$h_j < \frac{H_0 + h_j x_j}{S_0 + x_j} < \frac{H_0}{S_0}.$$

So, we can increase x_j to its upper bound, $1/\bar{R}_j$, and update C to

$$C = \frac{H_0 + (h_j/\bar{R}_j)}{S_0 + (1/\bar{R}_j)},$$

and compare it against h_{j-1}, and repeat the above.

So here is the algorithm to solve the minimization problem in (42):

- *Step 0.* Set $j = -1$, $x_j = 0$; set $S = S_0$ and $H = H_0$; set $C = H/S$.
- *Step 1.* Stop, if $h_j \geq C$; otherwise, continue.
- *Step 2.* Set $x_j = 1/\bar{R}_j$, and update:

  $$H \leftarrow H + (h_j/\bar{R}_j), \quad S \leftarrow S + (1/\bar{R}_j), \quad C = H/S, \quad j \leftarrow j - 1;$$

  goto Step 1.

Thus, for all $j = -1, \ldots, r$, we have $h_j < C$; and we stop at r such that $h_{r-1} \geq C$ (this is analogous to $C(r) = h(r)$ in the Brownian motion setting), and this is the optimal r (which could be $-\infty$, as in the Brownian case). Furthermore, for all $j = -1, \ldots, r$, we have $x_j = 1/\bar{R}_j$; that is, it is optimal to set $\mu_j = \mu$, just like what's optimal for μ_n with $n \geq 0$.

In retrospect, it's quite clear that we must have $y_r > 0$ (which then implies $\pi_r > 0$; and hence, $\pi_n > 0$ for all $n \geq r$). For if $y_r = 0$, we would have $z \leq h_r$ — refer to the first constraint in (41), and this would contradict $C > h_r$ (refer to the above algorithm), since $z = C$ at optimality.

Finally, observe that for C to be finite, we need $H_0 < \infty$; refer to the objective function in (42). To guarantee this, we need the holding cost $\{h_n\}$ to satisfy $H_0 := \sum_{n\geq 0} h_n \rho^n < \infty$, which is analogous to the exponential boundedness condition on $h(x)$ in (4). Specifically, the condition can be expressed as:

$$\exists b < -\ln(\rho), \quad \exists n_0 > 0, \qquad \text{s.t.} \quad h_n \leq e^{bn}, \quad \forall n \geq n_0; \tag{43}$$

so that $\rho e^b < 1$, and hence, $\sum_{n\geq 0} h_n \rho^n \leq \sum_{n\geq 0} (\rho e^b)^n < \infty$. Similarly, the remarks preceding Proposition 1 regarding the optimal reflection level and the corresponding holding cost in the Brownian setting all have analogous counterparts here. In particular, if it is optimal to have $r = -\infty$, then we must have $h_{-\infty} (:= C_{-\infty}) < H_0/S_0 (:= C_0)$, as evident from the analysis above.

In summary, we have

Theorem 4 Suppose the holding cost $\{h_n\}$ satisfies the exponential boundedness condition in (43). Then, the reflection control is optimal for the birth-death queue model; specifically, set the service rate $\mu = 0$ for $n < r < 0$, and $\mu_n = \mu$ for $n \geq r$; with the optimal r identified by the above algorithm.

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] He, S., Yao, D.C. and H. Zhang, Optimal ordering policy for inventory models with quantity-dependent fixed costs. *Mathematics of Operations Research*, **42** (2017), 979-1006.

[8] 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.

[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.

[12] Yao, D.D. and Schechner, Z., Decentralized Control of Service Rates in a Closed Jackson Network. *IEEE Transactions on Automatic Control*, **34** (1989), 236-240.