OPERATIONS RESEARCH
Vol. 56, No. 3, May–June 2008, pp. 618–629
ISSN 0030-364X | EISSN 1526-5463 | 08 | 5603 | 0618
DOI 10.1287/opre.1060.0380

Impulse Control of Brownian Motion: The Constrained Average Cost Case

Melda Ormeci, J. G. Dai, John Vande Vate
School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332
{melda.ormeci@isye.gatech.edu, dai@isye.gatech.edu, john.vandevate@isye.gatech.edu}

When a manufacturer places repeated orders with a supplier to meet changing production requirements, he faces the challenge of finding the right balance between holding costs and the operational costs involved in adjusting the shipment sizes. We consider an inventory whose content fluctuates as a Brownian motion in the absence of control. At any moment, a controller can adjust the inventory level by any positive or negative quantity, but incurs both a fixed cost and a cost proportional to the magnitude of the adjustment. The inventory level must be nonnegative at all times and continuously incurs a linear holding cost. The objective is to minimize long-run average cost. We show that control band policies are optimal for the average cost Brownian control problem and explicitly calculate the parameters of the optimal control band policy. This form of policy is described by three parameters $\{q, Q, S\}$, $0 < q \leqslant Q < S$. When the inventory falls to zero (rises to S), the controller expedites (curtails) shipments to return it to q (Q). Employing apparently new techniques based on methods of Lagrangian relaxation, we show that this type of policy is optimal even with constraints on the size of adjustments and on the maximum inventory level. We also extend these results to the discounted cost problem.

Subject classifications: inventory/production: uncertainty, stochastic; probability: stochastic model applications.
Area of review: Stochastic Models.
History: Received October 2005; revision received April 2006; accepted May 2006.

1. Introduction and Motivation

Consider a manufacturer that places repeated orders for a part with a supplier to meet changing production requirements. Because of the costs involved in idling manufacturing capacity, backordering or stockouts of the part are not acceptable. Thus, if inventory of this part falls to precariously low levels, the manufacturer may expedite shipments or take other actions to increase it. On the other hand, space and capital constraints limit the inventory the manufacturer is willing to hold. When inventory grows too large, the manufacturer may take actions to reduce it. The manufacturer's challenge is to minimize the space and capital costs associated with holding inventory and the operational costs involved in adjusting supply.

This problem is common in the automobile industry, where the costs of idling production at an assembly plant can exceed \$1,000 per minute. Although an assembly plant typically produces vehicles at a remarkably constant rate, the composition of those vehicles can vary widely either in terms of the options they require or, as manufacturers move to more flexible lines, in terms of the mix of models produced. It is not unusual in the industry to see usage of a part vary by over 70% from one day to the next. Electronically transmitted releases against a standing purchase order have essentially eliminated ordering costs and careful packaging, loading, and transportation planning have squeezed planned transportation costs to the last penny. But the increasing number of models and options have increased the variability in usage, while growing reliance on suppliers in lower-cost countries such as Mexico and China have compounded the complexity of supply. Thus, we focus on the balance between the capital and space costs of carrying inventory and the unplanned costs such as expediting and curtailing shipments incurred in controlling inventory levels.

We model this problem as a Brownian control problem and seek a policy that minimizes the long-run average cost. We model the netput process or the difference between supply and demand for the part in the absence of any control as a Brownian Motion with drift μ and variance σ^2. Inventory incurs linear holding costs continuously and must remain nonnegative at all times. The manufacturer may, at any time, adjust the inventory level by, for example, expediting or curtailing shipments, but incurs both a fixed cost for making the adjustment and a variable cost that is proportional to the magnitude of the adjustment. The fixed cost and the unit variable cost depend on whether the adjustment increases or decreases inventory because these involve different kinds of interventions.

We address the average cost problem rather than the more traditional discounted cost problem for several reasons: First, although the notion of a discount factor may be natural and intuitive in many applications including finance, it is generally alien to material planners and the challenges of motivating it and eliciting a value for it outweigh the potential benefits. Second, researchers have traditionally pursued the discounted problem because that version of the dynamic programming operator exhibits favorable contraction properties that facilitate analysis. Although the discounted cost version of our Brownian control problem has been studied in the literature (Harrison et al. 1983), the average cost version has not.

A common method for solving long-run average cost problems is to utilize the discounted cost problem and take the limit as the discount rate goes to zero; see, for example, Feinberg and Kella (2002), Hordijk and van der Duyn Schouten (1986), Robin (1981), and Sulem (1986). In this paper, we address the average cost problem directly without the traditional reliance on a limit of the discounted cost version. This approach is more direct, more elegant, and opens the possibility of tackling average cost problems with more general holding cost structures because it does not require explicit solutions to the discounted cost problem. We show that control band policies are optimal for the average cost Brownian control problem. This form of policy is described by three parameters $\{q, Q, S\}$, $0 < q \leqslant Q < S$. When the inventory falls to zero, the manufacturer expedites a shipment to return it to q. When the inventory rises to S, the maximum allowed, the manufacturer curtails shipments, reducing the balance to Q. The simplicity of this policy greatly facilitates its application in industrial settings such as automobile assembly, with thousands of parts to manage.

We extend the Brownian control problem by introducing constraints that reflect more of the realities of the inventory management problem. First, we impose an upper bound on the inventory level to reflect physical limits on total available inventory space or financial limits on the inventory budget. In addition, noting that the magnitude of adjustments may be limited, for example, by the nominal shipment quantity or the capacity of the transportation mode, we introduce bounds on the magnitude of each control. We prove that control band policies are optimal for the average cost Brownian control problem even with these constraints on the maximum inventory level and on the magnitudes of adjustments to the inventory. In each case, we provide optimality conditions that allow explicit calculation of the optimal control parameters. In fact, employing apparently new applications of Lagrangian relaxation techniques (Fisher 1981), we show how to reduce the constrained problem to a version of the original unconstrained problem and, in the process, provide methods for computing the optimal control band policy in the presence of the constraints. This approach extends to constrained versions of the discounted cost problem and we state the analogous results in that setting.

In their paper, Harrison et al. (1983) addressed a problem in finance called the stochastic cash management problem in which a certain amount of income or revenue is automatically channelled into a cash fund from which operating disbursements are paid. If the balance in the cash fund grows too large, the controller may invest the excess. If it becomes too small, he may sell off investments to replenish it. The challenge is to minimize the discounted opportunity costs associated with holding cash in the fund and the transaction costs involved in buying and selling investments. They showed that a control band policy is optimal for the stochastic cash management problem and provided methods for computing the optimal control band policy. Although they studied exactly our problem but with discounted costs rather than average costs, Taksar (1985) studied the average cost problem but with a different cost structure on the controls: He minimized the average holding and control costs when there are *no* fixed costs for control and singular control is employed. He showed that the optimal policy, characterized by two constants $a < b$, keeps the process inside $[a, b]$ with minimal effort. Constantinides (1976) studied a similar cash management problem that allowed both positive and negative cash balance. Although he looked at the average cost problem, he assumed the optimal policy to be of a simple form and proceeded to find the optimal parameters of this policy. On the other hand, Richard (1977) looked at a diffusion process with fixed plus proportional adjustments costs and general holding costs, and showed that the optimal policy is one of impulse control in both the finite- and infinite-horizon discounted cases, without addressing the existence of such a control.

Our model differs from the ones in these works in many ways. While Harrison et al. (1983) and Richard (1977) allow adjustments of any magnitude, we introduce constraints on the magnitude of adjustments to the inventory level. While Constantinides (1976) did not have any constraints on the inventory level, even allowing negative values, Harrison et al. (1983) required that inventory remain nonnegative at all times, and Taksar (1985) allowed the holding cost to be infinite outside a range, which amounts to constraints on the inventory level. On the other hand, we introduce a constraint on the maximum inventory level while keeping inventory nonnegative. We develop a method built on ideas from Lagrangian relaxation techniques to handle these additional constraints. The method is quite general and we extend it to analogously constrained versions of the discounted cost problem. Finally, Harrison et al. (1983) considered the Brownian control problem in a financial setting where discounting costs over time is natural and appropriate. We consider the problem in an industrial setting where the long-run average cost is more natural and accepted.

The rest of this paper is organized as follows. In §2, we describe the average cost Brownian control problem and its policy space. The main results of this paper, the optimality of control band policies and optimality conditions that permit ready computation of the optimal policy parameters, are stated here. Sections 3 and 4 set up the preliminaries for the solution of the problem. In §3, we introduce a lower bound for the optimal cost, and in §4, we define a relative value function for control band policies with average cost criteria and show that the average cost can be calculated through this function. In §5, we first consider the bounded inventory average cost Brownian control problem in which M, the maximum inventory level allowed, is finite. We prove that a control band policy is optimal for the bounded inventory average cost Brownian control problem and derive explicit equations used in calculating the optimal control parameters. As a special case, we characterize an optimal solution for the unconstrained average cost problem with no bounds on the maximum inventory level. We also demonstrate the optimal policy for the discounted cost setting with finite M. In §6, we introduce constraints on the magnitude of the adjustments. In particular, employing Lagrangian relaxation techniques, we reduce the constrained problem to a version of the original unconstrained problem. Once again, we solve the constrained problem in the average cost setting and characterize the optimal policies for the discounted cost setting as well. We conclude the paper in §7 by looking at possible extensions and by stating the optimal policy for the bounded inventory constrained average cost Brownian control problem, which simultaneously imposes bounds on the controls and a finite upper limit on inventory.

2. Impulse Control of Brownian Motion

In this paper, we use the following notation and assumptions. Let Ω be the space of all continuous functions ω: $[0,\infty) \to \mathbb{R}$, the real line. For $t \geqslant 0$, let X_t: $\Omega \to \mathbb{R}$ be the coordinate projection map $X_t(\omega) = \omega(t)$. Then, $X = (X_t,\ t \geqslant 0)$ is the canonical process on Ω. Let $\mathcal{F} = \sigma(X_t,\ t \geqslant 0)$ denote the smallest σ-field such that X_t is $\mathcal{F}$-measurable for each $t \geqslant 0$, and similarly let $\mathcal{F}_t = \sigma(X_s,\ 0 \leqslant s \leqslant t)$ for $t \geqslant 0$. When we mention adapted processes and stopping times hereafter, the underlying filtration is understood to be $\{\mathcal{F}_t,\ t \geqslant 0\}$. Finally, for each $x \in \mathbb{R}$, let $\mathbb{P}_x$ be the unique probability measure on $(\Omega, \mathcal{F})$ such that X is a Brownian motion with drift μ, variance σ^2, and starting state x under $\mathbb{P}_x$. Let $\mathbb{E}_x$ be the associated expectation operator.

We are to control a Brownian motion $X = \{X_t,\ t \geqslant 0\}$ with mean μ, variance σ^2, and starting state x. Upward or downward adjustments, ξ_n, are exerted at discrete times, T_n, so that the resulting inventory process represented by $Z = \{Z_t,\ t \geqslant 0\}$ remains within $[0, M]$, where M is a possibly infinite bound on inventory. We adopt the convention that the sample path of Z is right continuous on $[0,\infty)$ having left limits in $(0,\infty)$. (The time parameter of a process may be written either as a subscript or as an argument, depending on which is more convenient.)

A policy $\varphi = \{(T_n, \xi_n),\ n \geqslant 0\}$ consists of stopping times $\{T_0, T_1, \ldots\}$ at which control is exerted and random variables $\{\xi_0, \xi_1, \ldots\}$ representing the magnitude and direction of each control. So, T_n is the time at which we make the $(n+1)$st adjustment to inventory and ξ_n describes the magnitude and direction of that adjustment. We only consider policies that are nonanticipating, i.e., each adjustment ξ_n must be $\mathcal{F}_{T_n^-}$ measurable, where, for a stopping time τ, $\mathcal{F}_{\tau^-}$ is defined as in Definition I.1.11 of Jacod and Shiryaev (2003).

When a policy increases inventory by $\xi > 0$, it incurs cost $K + k\xi$ representing the fixed costs $K > 0$ of changing the inventory and the variable costs $k\xi \geqslant 0$ that grow in proportion to the size of the adjustment. When a policy reduces inventory, i.e., when it adjusts inventory by $\xi < 0$, it incurs cost $L - l\xi$, where $L > 0$ is the fixed cost for reducing inventory and $-l\xi \geqslant 0$ is the variable cost. Finally, we assume that inventory incurs a positive holding cost of $h > 0$ per unit per unit of time.

We consider the average cost Brownian control problem, which is to find a nonanticipating policy $\varphi = \{(T_n, \xi_n),\ n \geqslant 0\}$ that minimizes.

$$\mathrm{AC}(x,\varphi) = \limsup_{n\to\infty} \mathbb{E}_x\left[\frac{1}{T_n}\left(\int_0^{T_n} hZ_t\,dt + \sum_{i=1}^{n}((K+k\xi_i)1_{\{\xi_i>0\}} + (L+l|\xi_i|)1_{\{\xi_i<0\}})\right)\right], \tag{1}$$

the expected long-run average cost starting at a given initial point $x \in \mathbb{R}_+ = [0,\infty)$. Setting

$$\phi(\xi) = \begin{cases} K + k\xi & \text{if } \xi > 0, \\ 0 & \text{if } \xi = 0, \\ L - l\xi & \text{if } \xi < 0, \end{cases} \tag{2}$$

the control problem (1) can be written compactly as

$$\mathrm{AC}(x,\varphi) = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_x\left[\int_0^T hZ_t\,dt + \sum_{i=1}^{N(T)} \phi(\xi_i)\right], \tag{3}$$

where, for each time $t \geqslant 0$, $N(t) = \sup\{n \geqslant 0\colon\ T_n \leqslant t\}$ denotes the number of jumps by time t.

We introduce a possibly infinite upper bound M on the inventory level and restrict our attention to the policy space $\mathcal{P}$, which is the set of all nonanticipating policies satisfying

$$\mathbb{P}_x(0 \leqslant Z_t \leqslant M \text{ for all } t > 0) = 1 \quad \text{for all } x \in \mathbb{R}_+. \tag{4}$$

When the upper bound M on inventory is finite, we refer to the problem as the bounded inventory average cost Brownian control problem. When M is infinite, adding the constraints

$$-d \leqslant \xi_i \leqslant u \quad \text{for each } i = 1, 2, \ldots$$

to the average cost Brownian control problem gives rise to the constrained average cost Brownian control problem.

Harrison et al. (1983) proved that a simple form of policy, called a control band policy, is optimal for the discounted cost problem. A control band policy is defined by three parameters $\{q, Q, S\}$, $0 < q \leqslant Q < S$ (they define control band policies with strict inequalities between the parameters, $0 < q < Q < S$; however, we allow $0 < q \leqslant Q < S$). When inventory falls to zero, the policy exerts a control to bring it up to level q. When the inventory rises to S, the maximum allowed, the policy exerts a control to reduce it by $s=S-Q$ and bring it to level Q. If the initial inventory level lies outside the range $[0,S]$, the policy exerts a one-time control ξ_0 to bring it to the closer of level q or level Q. Thus, the control band policy $\{q,Q,S\}$ is defined by $\{(T_n,\xi_n),\ n\geqslant 0\}$, where $T_0=0$,

$$\xi_0=\begin{cases} q-X_0 & \text{if } X_0\leqslant 0,\\ 0 & \text{if } 0<X_0<S,\\ Q-X_0 & \text{if } X_0\geqslant S,\end{cases}$$

thereafter, $\{T_n,\ n>0\}$ are the hitting times for $\{0,S\}$, i.e., $T_n=\{t>T_{n-1}\colon Z_t=0 \text{ or } Z_t=S\}$ for all $n=1,2,\ldots$ and

$$\xi_n=\begin{cases} q & \text{if } Z(T_n^-)=0,\\ -s & \text{if } Z(T_n^-)=S.\end{cases}$$

When φ is a control band policy with parameters $\{q,Q,S\}$, the average cost does not depend on the initial state x, and hence we also use $\mathrm{AC}(\varphi)$ or $\mathrm{AC}(q,Q,S)$ to denote its average cost.

Now we state the main results, Theorems 1 and 2, of this paper. Theorem 1 says that a control band policy is optimal for the bounded inventory average cost Brownian control problem. Theorem 2 says that a control band policy is also optimal for the constrained average cost Brownian control problem. Both Theorems 1 and 2 provide explicit formulas for calculating the optimal control parameters.

Theorem 1. *The bounded inventory average cost Brownian control problem admits an optimal policy that is a control band policy. Furthermore, if $\mu\neq 0$, the parameters $\{q^*,Q^*,S^*\}$ of the optimal control band policy φ^* are defined by the unique nonnegative values of λ, s, Δ and $Q\geqslant\Delta$ satisfying*

$$L=-h\left(\frac{s^2(1+e^{\beta s})}{2\mu(1-e^{\beta s})}+\frac{s}{\beta\mu}\right)+\lambda\left(\frac{s}{1-e^{-\beta s}}-\frac{1}{\beta}\right), \tag{5}$$

$$k+l=-\frac{h\Delta}{\mu}-\frac{hs(1-e^{\beta\Delta})}{\mu(1-e^{-\beta s})}+\lambda\left(\frac{e^{\beta\Delta}-1}{1-e^{-\beta s}}\right), \tag{6}$$

$$K=\frac{h(Q-\Delta)se^{\beta s}}{\mu(1-e^{\beta s})}+\frac{h(\Delta^2-Q^2)}{2\mu}+\frac{hs(e^{\beta Q}-e^{\beta\Delta})}{\mu\beta(1-e^{-\beta s})}-(l+k)(Q-\Delta)+\lambda\left(\frac{e^{\beta Q}-e^{\beta\Delta}}{\beta(1-e^{-\beta s})}-\frac{(Q-\Delta)}{1-e^{-\beta s}}\right), \tag{7}$$

such that

$$\lambda(S-M)=0, \tag{8}$$

$$S\leqslant M, \tag{9}$$

where $\beta=2\mu/\sigma^2$, $q\equiv Q-\Delta$, and $S\equiv Q+s$.

If $\mu=0$, the parameters of the optimal control band policy φ^* are defined by the unique nonnegative values of λ, s, Δ, and Q satisfying (8), (9) *and*

$$L=\frac{hs^3}{6\sigma^2}+\frac{\lambda s}{2}, \tag{10}$$

$$l+k=\frac{h(\Delta^2+\Delta s)}{\sigma^2}+\lambda\frac{\Delta}{s}, \tag{11}$$

$$K=\frac{Q^3h}{3\sigma^2}+\frac{Q^2hs}{2\sigma^2}-\frac{\Delta^2hs}{2\sigma^2}-\frac{\Delta^3h}{3\sigma^2}+(\Delta-Q)(l+k)+\lambda\left(\frac{Q^2-\Delta^2}{2s}\right), \tag{12}$$

where $q\equiv Q-\Delta$ and $S\equiv Q+s$.

For each fixed $\lambda\geqslant 0$, (5) alone determines $s\equiv S-Q$, after which (6) determines $\Delta\equiv Q-q$, and then (7) determines q. The value of λ that also satisfies (8) and (9) gives the optimal control policy.

Theorem 2. *The constrained average cost Brownian control problem admits an optimal policy that is a control band policy. Furthermore, if $\mu\neq 0$, the parameters $\{q^*(d,u),Q^*(d,u),S^*(d,u)\}$ of the optimal control band policy φ^* are defined by the unique nonnegative values of λ, η, s, Δ, and $Q\geqslant\Delta$ satisfying*

$$L-\lambda d=-h\left(\frac{s^2(1+e^{\beta s})}{2\mu(1-e^{\beta s})}+\frac{s}{\beta\mu}\right), \tag{13}$$

$$k+\eta+l+\lambda=-\frac{h\Delta}{\mu}-\frac{hs(1-e^{\beta\Delta})}{\mu(1-e^{-\beta s})}, \tag{14}$$

$$K-\eta u=\frac{h(Q-\Delta)se^{\beta s}}{\mu(1-e^{\beta s})}+\frac{h(\Delta^2-Q^2)}{2\mu}+\frac{hs(e^{\beta Q}-e^{\beta\Delta})}{\mu\beta(1-e^{-\beta s})}-(l+\lambda+k+\eta)(Q-\Delta), \tag{15}$$

such that

$$\lambda(d-s)=0, \tag{16}$$

$$\eta(u-Q+\Delta)=0, \tag{17}$$

$$s\leqslant d, \tag{18}$$

$$Q-\Delta\leqslant u, \tag{19}$$

where $\beta=2\mu/\sigma^2$, $q\equiv Q-\Delta$, and $S\equiv Q+s$. If $\mu=0$, the parameters of the optimal control band policy φ^* are defined by the unique nonnegative values of λ, η, s, Δ, and $Q\geqslant\Delta$ satisfying (16)–(19) *and*

$$L-\lambda d=\frac{hs^3}{6\sigma^2}, \tag{20}$$

$$l+\lambda+k+\eta=\frac{h(\Delta^2+\Delta s)}{\sigma^2}, \tag{21}$$

$$K-\eta u=\frac{Q^3h}{3\sigma^2}+\frac{Q^2hs}{2\sigma^2}-\frac{\Delta^2hs}{2\sigma^2}-\frac{\Delta^3h}{3\sigma^2}+(\Delta-Q)(l+\lambda+k+\eta), \tag{22}$$

where $q\equiv Q-\Delta$ and $S\equiv Q+s$.

Note that when $d = \infty$, we take $\lambda = 0$ as the unique solution to (16), and when $u = \infty$, we take $\eta = 0$ as the unique solution to (17).

Control band policies are also optimal for the discounted Brownian control problem, with constraints on the inventory space and magnitude of adjustments; see Theorem 3 in §5 and Theorem 4 in §6.

The rest of this paper is devoted to the proofs of Theorems 1 and 2. In §3, we establish a lower bound on the average cost over all feasible policies. In §4, we define a relative value function for each control band policy with average cost criteria and show that the average cost can be calculated through this function. In §5, we complete the proof of Theorem 1 by showing that the average cost of a control band policy with a particular choice of control parameters achieves the lower bound. We prove Theorem 2 in §6 by employing Lagrangian relaxation techniques. We reduce the constrained problem to a version of the original unconstrained problem.

3. Lower Bound

In this section, we show how to construct a lower bound on the average cost over all feasible policies. Then, in §4 we define the relative value functions for control band policies and show how to compute their average costs. In §5, we construct a particular control band policy $\{q^*, Q^*, S^*\}$ and show that its value function provides a lower bound, and thus establishing the optimality of the control band policy.

Proposition 1. *Suppose that $f\colon [0, M] \to \mathbb{R}$ is continuously differentiable, has a bounded derivative, and has a continuous second derivative at all but a finite number of points. Then, for each time $T > 0$, initial state $x \in \mathbb{R}_+$, and policy $\{(T_n, \xi_n),\ n \geqslant 0\} \in \mathcal{P}$,*

$$\mathbb{E}_x[f(Z_T)] = \mathbb{E}_x[f(Z_0)] + \mathbb{E}_x\left[\int_0^T \Gamma f(Z_t)\,dt\right] + \mathbb{E}_x\left[\sum_{n=1}^{N(T)} \theta_n\right], \tag{23}$$

where

$$\theta_n = f(Z(T_n)) - f(Z(T_n^-)) \quad \textit{for } n = 1, 2, \ldots$$

and

$$\Gamma f = \frac{1}{2}\sigma^2 f'' + \mu f'. \tag{24}$$

Proof. The proof follows from an application of Ito's formula and is similar to the proof of (2.16) in Harrison et al. (1983). (Ito's formula for semimartingales can be found, for example, in Theorem I.4.57 of Jacod and Shiryaev 2003.) □

The following proposition shows that each function satisfying certain conditions provides a lower bound on the optimal average cost.

Proposition 2. *Suppose that $f\colon [0, M] \to \mathbb{R}$ satisfies all the hypotheses of Proposition 1 plus*

$$\Gamma f(x) - hx - \Gamma f(0) \leqslant 0 \quad \textit{for almost all } 0 \leqslant x \leqslant M, \tag{25}$$

$$f(x) - f(y) \leqslant K + k(x - y) \quad \textit{for } 0 \leqslant y < x \leqslant M, \tag{26}$$

$$f(x) - f(y) \leqslant L - l(x - y) \quad \textit{for } 0 \leqslant x < y \leqslant M. \tag{27}$$

Then, $\mathrm{AC}(x, \varphi) \geqslant -\Gamma f(0)$ *for each policy $\varphi \in \mathcal{P}$ and each initial state $x \in \mathbb{R}_+$.*

Proof. Recall the definition of θ_n in (24) and $\phi(\xi_n)$ in (2). Note that for each n, $\theta_n \leqslant \phi(\xi_n)$ by conditions (26) and (27). It follows from (23) and (25) that

$$\mathbb{E}_x[f(Z_T)] \leqslant \mathbb{E}_x[f(Z_0)] + \mathbb{E}_x\left[\int_0^T hZ_t\,dt\right] + \Gamma f(0)T + \mathbb{E}_x\left[\sum_{n=1}^{N(T)} \phi(\xi_n)\right]. \tag{28}$$

Dividing both sides of (28) by T and letting $T \to \infty$ gives

$$-\Gamma f(0) + \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_x[f(Z_T)] \leqslant \mathrm{AC}(x, \varphi). \tag{29}$$

If $\limsup_{T\to\infty}(1/T)\mathbb{E}_x[f(Z_T)] \geqslant 0$, (29) yields $-\Gamma f(0) \leqslant \mathrm{AC}(x, \varphi)$, proving the proposition. Now suppose that $\limsup_{T\to\infty}(1/T)\mathbb{E}_x[f(Z_T)] < 0$. We show that this implies

$$\mathrm{AC}(x, \varphi) = \infty, \tag{30}$$

which again yields $-\Gamma f(0) \leqslant \mathrm{AC}(x, \varphi)$, as desired. To prove (30), set $a = \limsup_{T\to\infty}(1/T)\mathbb{E}_x[f(Z_T)]$. Because $a < 0$ by assumption, it follows that there exists a constant $t^* > 0$ such that

$$\frac{1}{T}\mathbb{E}_x[f(Z_T)] < a/2 \quad \text{for } T > t^*,$$

and so $\mathbb{E}_x[f(Z_T)] < Ta/2$ for $T > t^*$. Because f has bounded derivatives, it is Lipschitz continuous. Thus, there exists a constant $c > 0$ such that

$$f(Z_0) - f(Z_T) \leqslant |f(Z_T) - f(Z_0)| \leqslant c|Z_T - Z_0| \leqslant c(Z_T + Z_0) \tag{31}$$

for $T \geqslant 0$. Taking expectations on both sides of (31), we see that

$$f(x) - \mathbb{E}_x[f(Z_T)] \leqslant c(\mathbb{E}_x[Z_T] + x)$$

for all $T \geqslant 0$. Therefore,

$$\mathbb{E}_x[Z_t] \geqslant \frac{1}{c}[f(x) + t|a|/2] - x = c_1 t + c_2 \quad \text{for all } t \geqslant t^*,$$

where $c_1 = |a|/(2c)$ and $c_2 = f(x)/c - x$. It follows that

$$\begin{aligned}\mathrm{AC}(x,\varphi) &\geqslant \limsup_{T\to\infty} \mathbb{E}_x\left[\frac{1}{T}\int_0^T hZ_t\,dt\right]\\ &= h\limsup_{T\to\infty}\frac{1}{T}\int_0^T \mathbb{E}_x[Z_t]\,dt\\ &\geqslant h\limsup_{T\to\infty}\left[\frac{1}{T}\int_{t^*}^T c_1 t\,dt\right]+hc_2\\ &\geqslant h\limsup_{T\to\infty}\left[\frac{c_1(T^2-t^{*2})}{2T}\right]+hc_2=\infty,\end{aligned}$$

proving (30). □

Remark. In the discounted cost problem, Harrison et al. (1983) obtained a bound similar to the one in Proposition 2. In proving their bound, they require their policies to satisfy the conditions

$$\mathbb{P}_x(Z_t \geqslant 0 \text{ for all } t>0)=1 \quad \text{for all } x\in\mathbb{R}_+ \quad \text{and} \tag{32}$$

$$\mathbb{E}_x \sum_{i=0}^{\infty} e^{-\gamma T_i}(1+|\xi_i|)<\infty \quad \text{for all } x\in\mathbb{R}_+, \tag{33}$$

where γ is the discount rate, and, as before, $\mathbb{R}_+ := [0,\infty)$. They used condition (33) to ensure that when f has bounded derivative, $\mathbb{E}_x[e^{-\gamma T} f(Z_T)] \to 0$ as $T\to\infty$.

A natural analog of (33) for the average cost problem is

$$\limsup_{n\to\infty} \mathbb{E}_x \frac{1}{T_n}\sum_{i=0}^{n}(1+|\xi_i|)<\infty \quad \text{for all } x\in\mathbb{R}_+. \tag{34}$$

One suspects that condition (34) should analogously lead to

$$\mathbb{E}_x[f(Z_T)/T]\to 0 \quad \text{as } T\to\infty \tag{35}$$

as long as the corresponding average cost is finite. Surprisingly, one can construct counterexamples such that (35) does not hold even though condition (34) holds and the corresponding average cost is finite; see §A of the online supplement for a counterexample. We are able to obtain the lower bound in Proposition 2 without condition (34) on the policies.

An electronic companion to this paper is available as part of the online version that can be found at http://or.journal.informs.org/.

4. Control Band Policies

In this section, we show that for a given control band policy $\varphi = \{q, Q, S\}$, the associated long-run average cost $\mathrm{AC}(x,\varphi)$ is independent of the initial state x. Furthermore, the average cost can be computed through a relative value function, which we define below.

Proposition 3 below shows that there is a constant g and a function V that satisfy the ordinary differential equation (ODE), known as the Poisson equation,

$$\Gamma V(x) - hx + g = 0, \quad 0\leqslant x\leqslant S, \tag{36}$$

and the boundary conditions

$$V(0)=V(q)-K-kq, \tag{37}$$

$$V(S)=V(Q)-L-ls. \tag{38}$$

The constant g is unique and the function V is unique up to an additive constant. With a slight abuse of terminology, any such function V is called the *relative value function* associated with the control band policy φ. The significance of the relative value function and the constant g is that they provide the long-run average cost $\mathrm{AC}(x,\varphi)$ of the control band policy through the formula $\mathrm{AC}(x,\varphi) = g = -\Gamma V(0)$.

Proposition 3. *Let the parameters of the control band policy $\varphi=\{q,Q,S\}$ be fixed.*

(a) *There is a function $V\colon [0,S]\to\mathbb{R}$ that is twice continuously differentiable on $[0,S]$ and satisfies (36)–(38).*

(b) *Such a function is unique up to a constant.*

(c) *The constant g is unique. The average cost of the control band policy $\{q,Q,S\}$ is independent of the starting point and is given by $g=-\Gamma V(0)$.*

Proof. We address only the case $\mu\neq 0$. The case $\mu=0$ is analogous. The general solution to the ODE (36) is

$$V(x) = Ax + Be^{-(2\mu/\sigma^2)x} + \frac{h}{2\mu}x^2 + E$$

for some constants A, B, and E. The boundary conditions (37) and (38) determine the values of A and B uniquely. Thus, we have proved both (a) and (b). Because g is a constant, it follows from (36) that $g=-\Gamma V(0)$, and thus g is unique.

To complete the proof of (c), consider the control band policy $\varphi=\{q,Q,S\}=\{(T_n,\xi_n),\ n\geqslant 0\}$. Because V is twice continuously differentiable and has bounded derivative on $[0,S]$, we have by Proposition 1 that

$$\mathbb{E}_x[V(Z_T)] = \mathbb{E}_x[V(Z_0)] + \mathbb{E}_x\left[\int_0^T \Gamma V(Z_t)\,dt\right] + \mathbb{E}_x\left[\sum_{n=1}^{N(T)}\theta_n\right],$$

where $\theta_n = V(Z(T_n)) - V(Z(T_n^-))$ for $n=1,2,\ldots$. Because V satisfies (37) and (38), $\theta_n = V(Z(T_n)) - V(Z(T_n^-)) = \phi(\xi_n)$ for $n=1,2\ldots$. Therefore, because V and g satisfy (36),

$$\begin{aligned}&\mathbb{E}_x[V(Z_T)] - \mathbb{E}_x[V(Z_0)] + gT\\ &\quad= \mathbb{E}_x\left[\int_0^T \Gamma V(Z_t)+g\,dt\right]+\mathbb{E}_x\left[\sum_{n=1}^{N(T)}\theta_n\right]\\ &\quad= \mathbb{E}_x\left[\int_0^T hZ_t\,dt\right]+\mathbb{E}_x\left[\sum_{n=1}^{N(T)}\phi(\xi_i)\right].\end{aligned}$$

Finally, dividing both sides by T, taking the limit as $T\to\infty$, and observing that

$$\lim_{T\to\infty}\frac{1}{T}\mathbb{E}_x[V(Z_T)]=0,$$

we have that $\mathrm{AC}(x,\varphi)=g$. □

REMARK. From the proof of Proposition 3, the average cost of the control band policy φ has the formula $\mathrm{AC}(x,\varphi) = g = -\Gamma V(0) = -(\sigma^2 h/(2\mu) + \mu A)$. A detailed computation in terms of A shows that the latter expression is equal to

$$\mathrm{AC}(x,\varphi) = \begin{cases} h\left(\dfrac{(S^2-Q^2)C}{2(sC-qD)} - \dfrac{q^2 D}{2(sC-qD)} - \dfrac{\sigma^2}{2\mu}\right) \\ \quad + \dfrac{((L+ls)C+(K+kq)D)\mu}{sC-qD}, & \mu \neq 0, \\ h\dfrac{3S^2-3Ss+s^2-q^2}{3(2S-s-q)} \\ \quad + \dfrac{((K/q)+k)\sigma_2}{2S-s-q} + \dfrac{((L/s)+l)\sigma_2}{2S-s-q}, & \mu = 0, \end{cases} \tag{39}$$

where

$$C = e^{-\beta q} - 1,$$

$$D = e^{-\beta S} - e^{-\beta Q}.$$

§B of the online supplement provides an alternative derivation of (39) for the average cost $\mathrm{AC}(x,\varphi)$ of a control band policy φ. The derivation is based on a basic adjoint relationship (see Harrison and Williams 1987a, b), which may be of independent interest.

Note that the relative value function of any control band policy satisfies conditions (36)–(38), which are related to conditions (25)–(27) used in Proposition 2 to construct a bound. In §5, we construct a control band policy, $\{q^*, Q^*, S^*\}$, whose relative value function can be extended to $[0, M]$. For the extended function to satisfy conditions (25)–(27) for all $x \in [0, M]$, the parameters $\{q^*, Q^*, S^*\}$ must be the unique values specified in Theorem 1.

5. Optimal Policy Parameters

One of the main results of this paper, stated in Theorem 1, is to prove that a control band policy is optimal for the bounded inventory average cost Brownian control problem and to provide optimality conditions that permit ready computation of the optimal policy parameters. In the remainder of this section, we prove Theorem 1. The outline of the proof is as follows.

In Proposition 3, we have proved that for any control band policy $\varphi = \{q, Q, S\}$, its average cost is given by $-\Gamma V(0)$, where V, defined on $[0, S]$, is the relative value function. We are to find a particular choice of parameter set $\{q^*, Q^*, S^*\}$ such that the corresponding relative value function V can be extended on $[0, M]$ and the extended function satisfies the conditions of Proposition 2. Thus, $-\Gamma V(0)$, in addition to being the average cost of the control band policy $\{q^*, Q^*, S^*\}$, is also a lower bound on the average cost of the bounded Brownian control problem. Therefore, the control band policy $\{q^*, Q^*, S^*\}$ is optimal.

Recall that for a given set of parameters $\{q, Q, S\}$, the corresponding relative value function satisfies (36)–(38). To search for the optimal parameter set $\{q^*, Q^*, S^*\}$, we impose the following conditions on $\{q, Q, S\}$ and V:

$$V'(q) = k, \tag{40}$$

$$V'(Q) = -l, \tag{41}$$

$$V'(S) = -l - \lambda, \tag{42}$$

$$\lambda(S - M) = 0, \quad \text{and} \tag{43}$$

$$S \leqslant M, \tag{44}$$

where $\lambda \geqslant 0$. Lemma 1 below shows that the parameter set $\{q^*, Q^*, S^*\}$ satisfying (40)–(44) exists. In the proof of Theorem 1, presented immediately following Lemma 2, the corresponding relative value function will be extended to $[0, M]$. Condition (40) is to ensure that the extended function satisfies inequality (26) of Proposition 2, and condition (41) is to ensure that the extended function satisfies inequality (27) of Proposition 2. Condition (42) is to ensure that the derivative of the extended function is also continuous at S^*.

LEMMA 1. (a) *There exists a unique nonnegative solution* s^*, Δ^*, Q^*, λ^* *satisfying* (5)–(9).

(b) *For the parameter set* $\{s^*, \Delta^*, Q^*, \lambda^*\}$, *the corresponding relative value function satisfies* (40)–(44).

PROOF. We demonstrate the proof for $\mu \neq 0$. When $\mu = 0$, the arguments are analogous.

The proof of part (a) is given in §C of the online supplement.

Now we prove part (b). Set $S^* = Q^* + s^*$ and $q^* = Q^* - \Delta^*$. We now show that the relative value function V corresponding to the control band policy $\{q^*, Q^*, S^*\}$ satisfies (40)–(44).

First, note that the function

$$\begin{aligned} f(x) &= -\frac{hs^*(1-e^{-(2\mu/\sigma^2)x})}{\mu(1-e^{-(2\mu/\sigma^2)s^*})} + \frac{hx}{\mu} - l - \lambda^* \frac{(1-e^{-(2\mu/\sigma^2)x})}{1-e^{-(2\mu/\sigma^2)s^*}} \\ &= -\left[\frac{hs^*}{\mu} + \lambda^*\right] \frac{(1-e^{-(2\mu/\sigma^2)x})}{(1-e^{-(2\mu/\sigma^2)s^*})} + \frac{hx}{\mu} - l \end{aligned} \tag{45}$$

is the unique solution to the ODE: $\Gamma f(x) - h = 0$ for $-Q^* \leqslant x \leqslant s^*$ and satisfies boundary conditions $f(0) = -l$ and $f(s^*) = -l - \lambda^*$. Let

$$\pi(x) = f(x - Q^*), \quad 0 \leqslant x \leqslant S^*.$$

It follows that π is the unique solution to the ODE

$$\Gamma \pi(x) - h = 0, \quad 0 \leqslant x \leqslant S^*, \tag{46}$$

satisfying the boundary conditions

$$\pi(Q^*) = -l \quad \text{and} \tag{47}$$

$$\pi(S^*) = -l - \lambda. \tag{48}$$

(When $\mu = 0$, $\pi(x) = (hx^2/\sigma^2) + Bx + C$ for some constants B and C.)

For $x \in [0, S^*]$, let

$$V(x) = \int_0^x \pi(u)\,du. \tag{49}$$

We claim that $V(x)$, defined on $[0, S^*]$, is a relative value function of the control band policy $\{q^*, Q^*, S^*\}$. To see this, we first note that $(\Gamma V(x) - hx)' = 0$ on $[0, S^*]$, and thus $\Gamma V(x) - hx$ is a constant on $[0, S^*]$. Denoting the constant by $-g$, we have that $V(x)$, together with the constant g, satisfies the Poisson equation (36). We next prove that $V(x)$ satisfies boundary conditions (37) and (38). Boundary condition (38) can be written in terms of π as

$$\begin{aligned} L &= V(Q^*) - V(S^*) - l(S^* - Q^*) \\ &= -\int_{Q^*}^{S^*} (\pi(x) + l)\,dx = -\int_0^{s^*} (f(x) + l)\,dx \\ &= -h\left(\frac{s^{*2}(1 + e^{\beta s^*})}{2\mu(1 - e^{\beta s^*})} + \frac{s^*}{\beta\mu}\right) + \lambda^*\left(\frac{s^*}{1 - e^{-\beta s^*}} - \frac{1}{\beta}\right), \end{aligned} \tag{50}$$

which holds because of (5). Similarly, boundary condition (37) can be written in terms of π as

$$\begin{aligned} K &= V(q^*) - V(0) - kq^* \\ &= \int_0^{q^*} (\pi(x) - k)\,dx = \int_{-Q^*}^{-\Delta^*} (f(x) - k)\,dx \\ &= \frac{h(Q^* - \Delta^*)s^* e^{\beta s^*}}{\mu(1 - e^{\beta s^*})} + \frac{h(\Delta^{*2} - Q^{*2})}{2\mu} \\ &\quad + \frac{hs^*(e^{\beta Q^*} - e^{\beta\Delta^*})}{\mu\beta(1 - e^{-\beta s^*})} - (l + k)(Q^* - \Delta^*) \\ &\quad + \lambda^*\left(\frac{e^{\beta Q^*} - e^{\beta\Delta^*}}{\beta(1 - e^{-\beta s^*})} - \frac{(Q^* - \Delta^*)}{1 - e^{-\beta s^*}}\right), \end{aligned} \tag{51}$$

which holds because of (7). Therefore, $V(x)$ is the relative value function of the control band policy $\{q^*, Q^*, S^*\}$.

Clearly, $V(x)$ satisfies conditions (41) and (42). To complete the proof of the lemma, it remains to prove that $V(x)$ satisfies condition (40). To see this, condition (40) requires

$$\begin{aligned} k &= \pi(q^*) = f(q^* - Q^*) = f(-\Delta^*) \\ &= -\frac{hs^*(1 - e^{(2\mu/\sigma^2)\Delta^*})}{\mu(1 - e^{-(2\mu/\sigma^2)s^*})} \\ &\quad - \frac{h\Delta^*}{\mu} - l - \lambda^* \frac{(1 - e^{(2\mu/\sigma^2)\Delta^*})}{1 - e^{-(2\mu/\sigma^2)s^*}}, \end{aligned} \tag{52}$$

which is equivalent to (6). Thus, the function $V(x)$ is the relative value function satisfying (40)–(44). □

The following properties of π are useful in the proof of Theorem 1.

Lemma 2. *Let π: $[0, S^*] \to \mathbb{R}$ be the unique solution to the ODE (46) satisfying (47) and (48) for the optimal parameters $\{s^*, \Delta^*, Q^*, \lambda^*\}$ satisfying (5)–(9). Extend $\pi(x)$ to $[S^*, M)$ via $\pi(x) = -l$ for $x \in [S^*, M)$. Then,*

(a) *For $x \in [0, q^*]$, $\pi(x) \geqslant k$, and for $x \in [q^*, M]$, $\pi(x) < k$.*

(b) *For $x \in [0, Q^*]$, $\pi(x) \geqslant -l$, and for $x \in [Q^*, M]$, $\pi(x) \leqslant -l$.*

Proof. Recall that $\pi(x) = f(x - Q^*)$ for all $x \in [0, S^*]$, where f is defined in (45). When $\mu \geqslant 0$, it is clear from (45) that π is strictly convex. The result follows from the convexity of π and conditions (47)–(48) and (52).

When $\mu < 0$, we have two cases to consider. If $(hs^*/\mu) + \lambda^* < 0$, π is again strictly convex, hence the same arguments apply. If $(hs^*/\mu) + \lambda^* > 0$, π is strictly concave and decreasing. To see this, note that

$$f'(x) = -\left[\frac{hs^*}{\mu} + \lambda^*\right] \frac{2\mu}{\sigma^2} \frac{e^{-(2\mu/\sigma^2)x}}{1 - e^{-(2\mu/\sigma^2)s^*}} + \frac{h}{\mu} < 0.$$

In this case, the result follows from (47)–(48) and (52). □

Proof of Theorem 1. Let $s^* > 0$, $\Delta^* \geqslant 0$ and $Q^* > \Delta^*$ and $\lambda^* \geqslant 0$ be the unique solution in Lemma 1. We now show that the control band policy $\{q^*, Q^*, S^*\}$ is optimal for the bounded inventory problem, where $q^* = Q(\lambda^*) - \Delta(\lambda^*)$, $Q^* = Q(\lambda^*)$, and $S^* = Q(\lambda^*) + s(\lambda^*) \leqslant M$.

Recall that in the proof of Lemma 1, the relative value function of the control band policy $\{q^*, Q^*, S^*\}$ can be expressed as $V(x) = \int_0^x \pi(y)\,dy$ for $x \in [0, S^*]$, where $\pi(x) = f(x - Q)$ is given by (45).

For $S^* < x \leqslant M$, we extend V and π as

$$V(x) = V(Q^*) - L - l(x - Q^*), \qquad \pi(x) = -l. \tag{53}$$

Thus, the extended function V, still denoted by V, satisfies

$$V(x) = \int_0^x \pi(y)\,dy, \quad x \in [0, M). \tag{54}$$

We now show that the extended function satisfies all the conditions of Proposition 2.

First, condition (38) implies that V is continuous at S^*, thus continuous on $[0, M]$. Next, we show that V has continuous derivatives in $(0, M)$. If $S^* = M$, then $V'(x) = \pi(x)$ for $x \in (0, M)$, and thus $V'(x)$ is continuous in $(0, M)$. Now, assume that $S^* < M$. By (43), $\lambda^* = 0$. Therefore, condition (42) implies that the left-side derivative of V at S^* is $-l$, which is equal to the right-side derivative obtained from (53). Clearly, the second derivative of V is continuous on $[0, S^*)$ and on (S^*, M).

We now check that V satisfies condition (25) of Proposition 2. By construction,

$$\Gamma V(x) - hx - \Gamma V(0) = 0 \quad \text{for } x \in [0, S^*]. \tag{55}$$

We show that V satisfies

$$\Gamma V - hx - \Gamma V(0) \leqslant 0 \quad \text{for } S^* < x \leqslant M. \tag{56}$$

It is enough to consider the case when $S^* < M$, and hence $\lambda^* = 0$.

From (55),

$$0 = V''(S^*-) + \mu V'(S^*-) - hS^* - \Gamma V(0)$$
$$= V''(S^*-) - \mu l - hS^* - \Gamma V(0).$$

Also, $V''(S^*-) = \pi'(S^*-) = f'(s^*) \geqslant 0$. The latter inequality follows from the fact that $f(x)$ is strictly convex with its minimum in $(0, s^*)$ when $\lambda^* = 0$.

It follows that

$$-\mu l - hx - \Gamma V(0) \leqslant -\mu l - hS^* - \Gamma V(0) \leqslant 0$$

for all $S^* < x \leqslant M$, which proves (56).

We demonstrate that (26) holds. The arguments for (27) are analogous and we leave them to the reader.

Recalling that when $0 \leqslant y < x \leqslant M$, $V(x) - V(y) = \int_y^x \pi(z)\,dz$, we apply the observations of Lemma 2 to the following cases:

Case 1. $q^* \leqslant y < x \leqslant M$. In this case,

$$V(x) - V(y) = \int_y^x \pi(z)\,dz \leqslant k(x-y) \leqslant K + k(x-y).$$

Case 2. $0 \leqslant y \leqslant q^* < x \leqslant M$. In this case,

$$V(y) - V(0) = \int_0^y \pi(z)\,dz \geqslant ky \quad \text{and} \tag{57}$$
$$V(x) - V(0) = V(q^*) - V(0) + V(x) - V(q^*)$$
$$= K + kq^* + \int_{q^*}^x \pi(z)\,dz \leqslant K + kx$$

from which it follows that $V(x) - V(y) \leqslant K + k(x-y)$ as desired.

Case 3. $0 \leqslant y < x \leqslant q^*$. In this case, we still have (57) and

$$V(x) - V(0) = V(q^*) - V(0) - \int_x^{q^*} \pi(z)\,dz = K + kq^*$$
$$- \int_x^{q^*} \pi(z)\,dz \leqslant K + kx$$

from which (26) again follows.

Thus, by Proposition 2, $\mathrm{AC}(x, \varphi) \geqslant -\Gamma V(0)$ for each policy $\varphi \in \mathcal{P}$ and each initial state $x \in \mathbb{R}_+$, where $-\Gamma V(0)$ is the average cost of policy $\{q^*, Q^*, S^*\}$. It remains only to show that this same inequality holds for $x < 0$, which we leave as an exercise. □

Note that when M is infinite, the bounded inventory average cost Brownian control problem becomes the unconstrained average cost Brownian control problem. In this case, it is clear that an optimal policy is the control band policy with parameters $\{q, Q, S\}$ determined by the solution to (5)–(7) (or (10)–(12)) with $\lambda = 0$.

Corollary 1. *A control band policy is optimal for the unconstrained average cost Brownian control problem. The parameters of this optimal policy are the unique solution* $\{q^*, Q^*, S^*\}$ *to Equations* (5)–(7) (*or* (10)–(12)) *with* $\lambda = 0$.

These results extend to the bounded inventory discounted cost problem that imposes a bound on the maximum inventory level in the discounted problem described by Harrison et al. (1983). We state the result without proof. Here, $\gamma > 0$ is the discount rate.

Theorem 3. *The bounded inventory discounted cost Brownian control problem admits an optimal policy that is a control band policy. Furthermore, the parameters* $\{q^*, Q^*, S^*\}$ *of the optimal control band policy* φ^* *are defined by the unique nonnegative values of* λ, s, Δ, *and* $Q \geqslant \Delta$ *satisfying*

$$L = rs + \frac{r(1-e^{-\rho s})(1-e^{\alpha s})}{(e^{\alpha s} - e^{-\rho s})}\left(\frac{1}{\alpha} + \frac{1}{\rho}\right)$$
$$+ \lambda\left(\frac{r}{e^{\alpha s} - e^{-\rho s}}\left(\frac{e^{\alpha s}-1}{\alpha} + \frac{e^{-\rho s}-1}{\rho}\right)\right), \tag{58}$$

$$c = r\left(\frac{1-e^{-\rho s}}{e^{\alpha s} - e^{-\rho s}}e^{-\alpha\Delta} + \frac{e^{\alpha s}-1}{e^{\alpha s} - e^{-\rho s}}e^{\rho\Delta}\right)$$
$$+ \lambda\left(\frac{r}{e^{\alpha s} - e^{-\rho s}}(e^{\rho\Delta} - e^{-\alpha\Delta})\right), \tag{59}$$

$$K = r\left[\frac{(1-e^{-\rho s})(e^{-\alpha\Delta} - e^{-\alpha Q})}{\alpha(e^{\alpha s} - e^{-\rho s})} + \frac{(e^{\alpha s}-1)(e^{\rho Q} - e^{\rho\Delta})}{\rho(e^{\alpha s} - e^{-\rho s})}\right]$$
$$- c(Q-\Delta) + \lambda\left[\frac{r(e^{-\alpha Q} - e^{-\alpha\Delta})}{\alpha(e^{\alpha s} - e^{-\rho s})} + \frac{r(e^{\rho Q} - e^{\rho\Delta})}{\beta(e^{\alpha s} - e^{-\rho s})}\right], \tag{60}$$

$$0 = \lambda(S - M), \tag{61}$$

$$S \leqslant M, \tag{62}$$

where,

$$r = h/\gamma - l, \tag{63}$$

$$c = h/\gamma + k > r, \tag{64}$$

$$\alpha = [(\mu^2 + 2\gamma\sigma^2)^{1/2} - \mu]/\sigma^2 > 0, \tag{65}$$

$$\rho = [(\mu^2 + 2\gamma\sigma^2)^{1/2} + \mu]/\sigma^2 > 0, \tag{66}$$

$S \equiv Q + s$ *and* $q \equiv Q - \Delta$.

6. Constrained Policies

In this section, we add the constraints $-d \leqslant \xi_i \leqslant u$ on the magnitude of adjustments to the inventory to the average cost Brownian control problem.

One of the main contributions of this paper is a new technique based on methods of Lagrangian relaxation that reduce the constrained problem to a version of the original unconstrained problem and, in the process, provide methods for computing the optimal control band policy in the presence of the constraints.

Theorem 2 shows that a control band policy is optimal for the constrained average cost Brownian control problem and provides optimality conditions that permit ready computation of the optimal policy parameters.

Proof. We prove Theorem 2 for the special case where $u=\infty$ and leave the rest as an exercise to the reader. In this case, the constrained problem becomes

$$\begin{aligned}\mathrm{AC}^*(d,\infty) &= \inf_{\varphi}\limsup_{n\to\infty}\mathbb{E}_x\left[\frac{1}{T_n}\left(\int_0^{T_n} hZ_t\,dt+\sum_{i=1}^n((K+k\xi_i)1_{\{\xi_i>0\}}\right.\right.\\ &\qquad \left.\left.+(L-l\xi_i)1_{\{\xi_i<0\}})\right)\right] \qquad (67)\\ \text{s.t.}\quad & -d\leqslant \xi_i,\\ & \varphi\in\mathcal{P}.\end{aligned}$$

Note that as in the unconstrained problem, $\mathrm{AC}^*(d,\infty)$ does not depend on the initial state x, so we omit the initial state from our notation. The notation $\mathrm{AC}^*(d,\infty)$ indicates that $u=\infty$ and we are only imposing a bound on the magnitude of reductions. The notation $AC^*(\infty,\infty)$ refers to the unconstrained problem.

We proved in Corollary 1 that a control band policy is an optimal policy for the unconstrained problem

$$\mathrm{AC}^*=\mathrm{AC}^*(\infty,\infty)=\min_{\varphi}\ \mathrm{AC}(\varphi)\quad \text{s.t.}\ \varphi\in\mathcal{P}. \qquad (68)$$

We show that a control band policy solves the constrained problem (67) and, in fact, reduces this constrained problem to a version of the unconstrained problem (68).

Let $\{q^*,Q^*,S^*\}$ be the optimal solution to the unconstrained problem. There are two possible cases:

Case 1. $s^*=S^*-Q^*\leqslant d$. In this case, the control band policy that is optimal for the unconstrained problem is also optimal for the constrained problem.

Case 2. $s^*=S^*-Q^*>d$.

We prove that also in Case 2 a control band policy is an optimal policy for the constrained problem. To prove this result, we use Lagrangian relaxation, a classic method for constrained optimization that moves the constraint to the objective and assigns it a price λ. The resulting unconstrained problem provides a bound on the objective of the original constrained problem. We find a control band policy that achieves this bound thereby proving its optimality.

We introduce a Lagrange multiplier $\lambda\geqslant 0$ and move the constraint $-d\leqslant\xi_i$ to the cost function. For each scalar $\lambda\geqslant 0$ and policy $\varphi\in\mathcal{P}$, we define the Lagrangian function

$$\begin{aligned}\mathcal{L}(\varphi;\lambda)&=\limsup_{n\to\infty}\mathbb{E}_x\left[\frac{1}{T_n}\left(\int_0^{T_n} hZ_t\,dt+\sum_{i=1}^n((K+k\xi_i)1_{\{\xi_i>0\}}\right.\right.\\ &\qquad\left.\left.+(L-l\xi_i)1_{\{\xi_i<0\}}-\lambda(d+\xi_i)1_{\{\xi_i<0\}})\right)\right].\end{aligned}$$

Because we show that the optimal policy does not depend on the initial state x, we omit the initial state from the notation of the Lagrangian function. For fixed $\lambda\geqslant 0$, the Lagrangian primal is

$$\mathcal{L}(\lambda)=\min\ \mathcal{L}(\varphi;\lambda)\quad \text{s.t.}\ \varphi\in\mathcal{P}. \qquad (69)$$

Note that for each $\lambda\geqslant 0$, $\mathcal{L}(\varphi;\lambda)\leqslant \mathrm{AC}(\varphi)$ for each feasible control policy φ, and so $\mathcal{L}(\lambda)\leqslant\mathrm{AC}^*(d,\infty)$.

The Lagrangian problem (69) can be expressed as a version of the unconstrained problem with modified costs for reducing inventory. In particular,

$$\begin{aligned}\mathcal{L}(\lambda)&=\inf_{\varphi}\limsup_{n\to\infty}\mathbb{E}_x\left[\frac{1}{T_n}\left(\int_0^{T_n} hZ_t\,dt+\sum_{i=1}^n((K+k\xi_i)1_{\{\xi_i>0\}}\right.\right.\\ &\qquad\left.\left.+(L-l\xi_i)1_{\{\xi_i<0\}}-\lambda(d+\xi_i)1_{\{\xi_i<0\}})\right)\right]\\ &\text{s.t.}\quad \varphi\in\mathcal{P}\\ &=\inf_{\varphi}\limsup_{n\to\infty}\mathbb{E}_x\left[\frac{1}{T_n}\left(\int_0^{T} hZ_t\,dt+\sum_{i=1}^n((K+k\xi_i)1_{\{\xi_i>0\}}\right.\right.\\ &\qquad\left.\left.+(L-\lambda d-(l+\lambda)\xi_i)1_{\{\xi_i<0\}})\right)\right]\\ &\text{s.t.}\quad \varphi\in\mathcal{P}.\end{aligned}$$

Hence, for each $0\leqslant\lambda<L/d$, the Lagrangian problem admits a control band policy $\varphi_\lambda=\{q_\lambda,Q_\lambda,S_\lambda\}\in\mathcal{P}$ as an optimal policy, and so $\mathcal{L}(\varphi_\lambda;\lambda)=\mathcal{L}(\lambda)$ for each $0\leqslant\lambda<L/d$. Now consider the dual problem

$$\mathcal{L}=\max_{\lambda}\ \mathcal{L}(\lambda)\quad\text{s.t.}\ \lambda\geqslant 0. \qquad (70)$$

Note that $\mathcal{L}\leqslant\mathrm{AC}^*(d,\infty)$. If we can find a multiplier $0\leqslant\lambda^*\leqslant L/d$ such that the control band policy $\varphi_{\lambda^*}=\{q_{\lambda^*}, Q_{\lambda^*},S_{\lambda^*}\}$ satisfies

$$s_{\lambda^*}=S_{\lambda^*}-Q_{\lambda^*}\leqslant d,\ \text{i.e., } \xi_i=-s_{\lambda^*}\geqslant -d \quad \text{whenever } \xi_i<0 \qquad (71)$$

and

$$\lambda^*(d-s_{\lambda^*})=0, \qquad (72)$$

then

$$\begin{aligned}\mathrm{AC}(\varphi_{\lambda^*})&=\mathcal{L}(\varphi_{\lambda^*};\lambda^*)=\mathcal{L}(\lambda^*)\\ &\leqslant\mathcal{L}\leqslant\mathrm{AC}(d,\infty)\leqslant\mathrm{AC}(\varphi_{\lambda^*}),\end{aligned}$$

proving that λ^* is an optimal solution to the dual problem (70) and φ_{λ^*} is an optimal policy for the constrained problem. When the unconstrained problem yields $s^*=S^*-Q^*\leqslant d$, $\lambda^*=0$ and the proof is complete. When this is not the case, existence of a Lagrange multiplier $0 \leqslant \lambda^* < L/d$ satisfying (71) and (72) for each $d > 0$ can also be shown. See §D of the online supplement for details of the proof. Thus, we have proved Theorem 2 for the special case where $u = \infty$. □

In the discounted cost problem when the magnitude of adjustments is bounded, using similar arguments it is possible to prove the optimality of the control band policies. We briefly describe the mechanism behind the proof for the special case $u = \infty$, but omit the details. Following the notation of Harrison et al. (1983) in which the problem is formulated as finding a policy that maximizes the reward, we let $R(s) = -L + rs$ denote the return achieved each time the upper boundary is hit, where r is defined in (63). We introduce a Lagrange multiplier $\lambda \geqslant 0$ and move the constraint $s \leqslant d$ to the cost function. Thus, whenever the upper bound is hit, the system incurs a reward $R(s, \lambda) = -L + rs + \lambda(d - s)$. Note that because $\lambda \geqslant 0$ and $d - s \geqslant 0$, $R(s, \lambda) \geqslant R(s)$ for all feasible s, and so the Lagrangian problem provides an upper bound on the objective.

We may rewrite $R(s, \lambda)$ as

$$R(s, \lambda) = -(L - \lambda d) + (r - \lambda)s.$$

So, $R(s, \lambda)$ is equivalent to the original discounted cost problem with parameters $L - \lambda d$ and $r - \lambda$. Hence, the optimal solution is again a control band policy. It is easy to show the existence of $\lambda < L/d$ so that the solution of the Lagrangian relaxation yields $s \leqslant d$ and the optimal solution is given by Theorem 4.

Theorem 4. *The constrained discounted cost Brownian control problem with $\xi_i \geqslant -d$ admits an optimal policy that is a control band policy. Furthermore, the parameters $\{q^*, Q^*, S^*\}$ of the optimal control band policy φ^* are defined by the unique nonnegative values of $\lambda \leqslant L/d$, s, Δ, and $Q \geqslant \Delta$ satisfying*

$$L - \lambda d = (r-\lambda)s + \frac{(r-\lambda)(1-e^{-\rho s})(1-e^{\alpha s})}{(e^{\alpha s}-e^{-\rho s})}\left(\frac{1}{\alpha}+\frac{1}{\rho}\right), \quad (73)$$

$$c = (r-\lambda)\left(\frac{1-e^{-\rho s}}{e^{\alpha s}-e^{-\rho s}}e^{-\alpha\Delta} + \frac{e^{\alpha s}-1}{e^{\alpha s}-e^{-\rho s}}e^{\rho\Delta}\right), \quad (74)$$

$$K = (r-\lambda)\left[\frac{(1-e^{-\rho s})(e^{-\alpha\Delta}-e^{-\alpha Q})}{\alpha(e^{\alpha s}-e^{-\rho s})} + \frac{(e^{\alpha s}-1)(e^{\rho Q}-e^{\rho\Delta})}{\rho(e^{\alpha s}-e^{-\rho s})}\right] - c(Q-\Delta), \quad (75)$$

$$0 = \lambda(s-d), \quad (76)$$

$$s \leqslant d, \quad (77)$$

where, r, c, α, and ρ are given in (63)–(66), $S \equiv Q + s$, and $q \equiv Q - \Delta$.

7. Concluding Remarks

In this paper, we showed that an optimal policy for the average cost Brownian control problem is a control band policy and demonstrated how to calculate the optimal control parameters explicitly. We also considered the bounded inventory average cost Brownian control problem in which the total available inventory space is bounded, and the constrained average cost Brownian control problem in which the magnitude of each adjustment is bounded. The bounded inventory constrained average cost Brownian control problem combines these two problems and simultaneously imposes an upper bound d and a lower bound u on the controls, and an upper limit M on inventory. In this general setting, one can show that a control band policy is still optimal and its parameters, $\{q, Q, S\}$, when $\mu \neq 0$, can be determined from the unique nonnegative values of λ, η, τ, s, Δ, and $Q \geqslant \Delta$ satisfying

$$L - \tau d = -h\left(\frac{s^2(1+e^{\beta s})}{2\mu(1-e^{\beta s})} + \frac{s}{\beta\mu}\right) + \lambda\left(\frac{s}{1-e^{-\beta s}} - \frac{1}{\beta}\right), \quad (78)$$

$$k+\eta+l+\tau = -\frac{h\Delta}{\mu} - \frac{hs(1-e^{\beta\Delta})}{\mu(1-e^{-\beta s})} + \lambda\left(\frac{e^{\beta\Delta}-1}{1-e^{-\beta s}}\right), \quad (79)$$

$$K - \eta u = \frac{h(Q-\Delta)se^{\beta s}}{\mu(1-e^{\beta s})} + \frac{h(\Delta^2-Q^2)}{2\mu} + \frac{hs(e^{\beta Q}-e^{\beta\Delta})}{\mu\beta(1-e^{-\beta s})} - (l+\tau+k+\eta)(Q-\Delta) + \lambda\left(\frac{e^{\beta Q}-e^{\beta\Delta}}{\beta(1-e^{-\beta s})} - \frac{(Q-\Delta)}{1-e^{-\beta s}}\right), \quad (80)$$

such that

$$\lambda(S-M) = 0, \quad S \leqslant M,$$

$$\tau(d-s) = 0, \quad s \leqslant d,$$

$$\eta(u-Q+\Delta) = 0, \quad Q-\Delta \leqslant u,$$

where $\beta = 2\mu/\sigma^2$, $q \equiv Q - \Delta$, and $S \equiv Q + s$. (When $\mu = 0$, (78)–(80) are modified accordingly.)

This paper focused on an inventory control problem whose netput process follows a Brownian motion that has continuous sample paths. However, in most applications the netput process is a pure jump process; for example, a downward jump signifies the fulfillment of a customer order. One hopes to identify a class of inventory systems whose netput processes can be discontinuous such that our optimal policy to the average cost Brownian control problem provides some key insights to these systems. In the manufacturing setting, the justification of such procedure is often carried out through heavy traffic limit theorems; see, for example, Krichagina et al. (1994). Plambeck (2005) proved a heavy traffic limit theorem for an assemble-to-order system with capacitated component production and fixed transport costs. The limit theorem enables her to find an asymptotically optimal control for the system.

An important contribution of this paper is to develop a method based on Lagrangian relaxation to solve constrained stochastic problems. Lagrangian relaxation methods have been widely used in deterministic optimization problems, both to solve constrained problems optimally and to obtain lower bounds on the optimal solution where it can not be solved to optimality. In this paper, we showed that Lagrangian relaxation techniques can be adapted to solve stochastic control problems as well. This approach makes it possible to study a whole new venue of problems. Ormeci (2006) explores this approach in more detail and describes additional applications of it.

8. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://or.journal.informs.org/.

Acknowledgments

This research was supported in part by National Science Foundation grants DMI-0300599 and CMMI-0727400, by an IBM Faculty Award, Georgia Institute of Technology's Executive Master's in International Logistics (EMIL) program, and a gift from BMW Manufacturing, Inc.

References

Constantinides, G. M. 1976. Stochastic cash management with fixed and proportional transaction costs. *Management Sci.* **22** 1320–1331.

Feinberg, E. A., O. Kella. 2002. Optimality of D-policies for an $M/G/1$ queue with a removable server. *Queueing Systems: Theory Appl.* **42** 355–376.

Fisher, M. L. 1981. The Lagrangean relaxation method for solving integer programming problems. *Management Sci.* **27** 1–18.

Harrison, J. M., R. J. Williams. 1987a. Brownian models of open queueing networks with homogeneous customer populations. *Stochastics* **22** 77–115.

Harrison, J. M., R. J. Williams. 1987b. Multidimensional reflected Brownian motions having exponential stationary distributions. *Ann. Probab.* **15** 115–137.

Harrison, J. M., T. M. Sellke, A. J. Taylor. 1983. Impulse control of Brownian motion. *Math. Oper. Res.* **8** 454–466.

Hordijk, A., F. A. van der Duyn Schouten. 1986. On the optimality of (s, S) policies in continuous review inventory models. *SIAM J. Appl. Math.* **46** 912–928.

Jacod, J., A. N. Shiryaev. 2003. *Limit Theorems for Stochastic Processes*, 2nd ed. Springer, Berlin.

Krichagina, E. V., S. X. C. Lou, M. I. Taksar. 1994. Double band policy for stochastic manufacturing systems in heavy traffic. *Math. Oper. Res.* **19** 560–595.

Ormeci, M. 2006. Inventory control in a build-to-order environment. Ph.D. thesis, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta.

Plambeck, E. L. 2005. Asymptotically optimal control for an assemble-to-order system with capacitated component production and fixed transport costs. Technical report, Graduate School of Business, Stanford University, Stanford, CA.

Richard, S. F. 1977. Optimal impulse control of a diffusion process with both fixed and proportional costs of control. *SIAM J. Control Optim.* **15** 79–91.

Robin, M. 1981. On some impulse control problems with long run average cost. *SIAM J. Control Optim.* **19** 333–358.

Sulem, A. 1986. A solvable one-dimensional model of a diffusion inventory system. *Math. Oper. Res.* **11** 125–133.

Taksar, M. I. 1985. Average optimal singular control and a related stopping problem. *Math. Oper. Res.* **10** 63–81.