MATH 562: PROBABILITY II CLASS REPORT

Optimal Brownian Inventory Control

Fall 2024

By: Shayan K Azmoodeh (shayana3@illinois.edu)

Contents

1 Introduction

The aim of this work is to present a self-contained distillation of the study of optimally controlling Brownian stochastic processes under a convex holding cost and linear adjustment cost as presented by [2] and [3] (mainly focusing on the long run average cost case). The main results are stated and explained with the goal of providing an intuitive understanding behind their mechanics and motivation. Short proofs are included only when they are deemed instructive. Such processes and the corresponding control problem have found applications in inventory control (in a warehouse, for example) as well as in finance, such as the problem of tracking a stock index whilst aiming to minimize transaction costs [1]; in fact, [1] studies a more general model allowing for processes following a geometric Brownian motion, as is the assumption in the celebrated Black-Scholes model for equity options pricing [5]. The term *stochastic control* here refers to the process of making adjustments to a random process as it propagates through time thus, in effect, controlling it's behavior so it exhibits desirable properties (in this case, minimizing a cost function).

Beyond presenting a general optimal algorithm (here referred to as a *policy*) for controlling Brownian processes, we provide an intuitive explanation behind the four-step method used to prove the optimality of this algorithm, giving a mental framework to readily adapt these steps to be applied to any similar control problem. In particular, this provides an elegant unification of the methods presented in [2] and [3]. This approach is especially of interest as it provides a simple procedure to prove optimality without the need to consider the entire space of feasible algorithms. We also include a generalization of Itô's formula for general non-continuous semimartingales and non-C^2 functions that is used to formulate the optimality verification procedure.

Ultimately, it turns out that when acting to minimize the expected long run average cost (per unit time), a banded control policy is optimal, where we make no adjustments to the base process until it hits some pre-designated bands, and then adjust it accordingly.

2 Brownian Control Models

We begin by providing an overview of the main control problem, the processes under consideration, and feasible control policies (i.e., how we can control the process). The model is presented in the context of inventory control to provide a physical and practical grounding.

2.1 The Base Process

Throughout the rest of this paper, we work on the filtered probability space $\left(\Omega, (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}, \mathbb{P}\right)$. Let $(X_t)_{t\geq 0}$ be an $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion with drift μ and variance σ^2 started from $x \in \mathbb{R}$. Then X can be written in terms of a $(\mathcal{F}_t)_{t\geq 0}$-adapted standard (i.e., drift of 0 and variance of 1) Brownian motion W_t,

$$X_t = x + \mu t + \sigma W_t, \quad t \geq 0.$$

The process X_t is referred to as the *base inventory process* or the *netput process*. This is the path the inventory follows if it is allowed to fluctuate over time with no additional external interference beyond the standard forces of supply and demand. Note that the process can be negative, in which case its magnitude represents the backlog of orders that the inventory has to satisfy. In the absence of any objectives, letting this inventory evolve as it likes is entirely acceptable. However, understandably, holding inventory is not free and incurs costs, so we also consider a holding cost for our inventory, $h : \mathbb{R} \to \mathbb{R}_+$, where $h(x)$ denotes the cost of holding an amount of inventory x. We provide more details on the assumptions we make about h to make our problem tractable below, however the idea is to control the inventory with the goal of minimizing our costs.

2.2 The Controlled Process

In order to control this base inventory process, we can apply upward and downward shifts to the inventory (i.e., increase or decrease the inventory level by some amount when we deem necessary). Such a controlling procedure is called a *policy* (or algorithm), denoted $\varphi = (Y_1, Y_2)$, where Y_1 and

Y_2 are right continuous increasing processes with left limits and $Y_1(t)$, $Y_2(t)$ represent the total upward and downward adjustments made to the inventory process at time t, respectively (note if we let $\mathbb{D}$ denote the space of all cádlág functions on $\mathbb{R}$, then have $Y_1, Y_2 \in \mathbb{D}$). We do not require full continuity as this would restrict any policy to make adjustments continuously as opposed to one-time "impulse" adjustments, which we will later see are necessary. Then at any time $t \geq 0$, $Y_1(t) - Y_2(t)$ represents the total amount that we have deviated the base inventory process process from it's "natural" path by exercising the control policy, and thus we can write,

$$Z(t) = X(t) + Y_1(t) - Y_2(t)$$

to denote the *controlled inventory process*, or simply, the inventory process. Note that Z is a semimartingale since X is a martingale and Y_1, Y_2 are finite variation processes (since they are increasing). Moreover, Z is also right-continuous with left limits (i.e., $Z \in \mathbb{D}$) just like Y_1 and Y_2.

We use the notation $Z(t-)$ for any process and time $t \geq 0$ to denote the left limit of Z at t. Note that since Z is not necessarily left-continuous everywhere, we do not always have $Z(t-) = Z(t)$. Indeed, this represents the inventory process at time t before the corresponding control adjustments at that time have been made. Similarly, $Y_i(t)$ $(i = 1, 2)$ represents the total amount of adjustments (upward/downward) made up to but not including time t. With this understanding, we denote $Z(0-) = x$ as the initial inventory level (note that we allow instantaneous adjustments to the inventory even at time $t = 0$ so it may be the case that $Z(0-) \neq Z(0)$).

2.3 Minimizing Costs

With the current setup, it is tempting to simply apply controls continuously to keep the inventory process at a minimum of the holding cost function h at all times. However, this is not realistic and, frankly, not a very interesting problem. Each adjustment exercised by the control policy incurs a cost consisting of a fixed, per-adjustment cost and a cost proportional to the magnitude of the adjustment. Let K be the fixed and k be the proportional upward adjustment costs. Similarly, let L and ℓ be the fixed and proportional downward adjustment costs. Then an upward adjustment moving the inventory process from a value a to $b > a$ incurs a cost of $K + k(b - a)$ and similarly for downward adjustments.

Let $N_i(t)$ $(i = 1, 2)$ be another process representing the number of times an adjustment has been made through Y_i (upward/downward) up to and including time t. This can be written more formally as (i.e., number of times Y_i jumps/increases),

$$N_i(t) = |\{s \in [0, t] \mid Y_i(s) > Y_i(s-)\}|.$$

Then we can write the total adjustment cost incurred when acting according to the policy φ up to time t as the sum of the total proportional cost and total fixed costs,

$$\mathrm{Adj}(\varphi, x, t) = KN_1(t) + LN_2(t) + kY_1(t) + \ell Y_2(t).$$

And the total holding cost,

$$\mathrm{Hold}(\varphi, x, t) = \int_0^t h\big(Z(t)\big)\, \mathrm{d}t\,.$$

Combining these gives the total cost incurred up to and including time t,

$$\mathrm{Cost}(\varphi, x, t) = \mathrm{Adj}(\varphi, x, t) + \mathrm{Hold}(\varphi, x, t).$$

We are interested in the expected cost of a policy φ as $t \to \infty$, which gives the natural extension as the expected long-run average cost [2],

$$\mathrm{AC}(\varphi, x) = \limsup_{t\to\infty} \frac{1}{t}\mathbb{E}_x \left[\mathrm{Cost}(\varphi, x, t)\right]. \tag{2.1}$$

where $\mathbb{E}_x$ operator denotes the expectation given that the process X starts at x ($X_0 = x$). The use of the lim sup here serves to provide a more conservative estimate of the cost and thus by minimizing it we are guaranteed to minimize the "true" cost (i.e., the limit if it exists). One can also study the expected long run discounted cost, where all costs are discounted by some rate $\beta > 0$ to account for depreciating value of assets over time ([1, 3, 4]),

$$\mathrm{DC}(\varphi, x) = \mathbb{E}_x \left[\int_0^\infty e^{-\beta t} h(Z(t))\ \mathrm{d}t + \int_0^\infty e^{-\beta t}\mathrm{Adj}(\varphi, x, t)\ \mathrm{d}t\right]. \tag{2.2}$$

The remainder of this article focuses on optimal policies minimizing the AC cost, however a brief generalizing connection will be made to the DC case as well.

We make the following assumption throughout to make the problem tractable.

Assumption 2.1. *Let $h : \mathbb{R} \to \mathbb{R}_+$ be the holding cost. We assume,*

(a) h is convex

(b) h has a minimum at some point $a \in \mathbb{R}$ with $h(a) = 0$ and h decreasing for $x < a$ and h increasing for $x > a$. Moreover we assume $h \in C^2(\mathbb{R})$ at all points except possibly at a (this allows us to include holding costs such as $h(x) = |x|$).

(c) Let $\lambda = \frac{2\mu}{\sigma^2} \neq 0$. Then $h'(x)$ has smaller order than $e^{-\lambda x}$ so in a sense $e^{-\lambda x}$ dominates $h'(x)$ so $|h(x)|/e^{\lambda y}$ doesn't grow too large or too fast. More precisely,

$$\int_{-\infty}^{a} \frac{|h'(y)|}{e^{\lambda(a-y)}}\ \mathrm{d}y < \infty \quad \text{when } \lambda > 0$$

and

$$\int_{a}^{\infty} \frac{|h'(y)|}{e^{\lambda(a-y)}}\ \mathrm{d}y < \infty \quad \text{when } \lambda < 0.$$

3 Generalized Itô's Formula

In the following, for a function $g \in \mathbb{D}$, let $\Delta g(t) = g(t) - g(t-)$ (quantify the jumps at the points of discontinuity) and let g^c denote its continuous portion,

$$g^c(t) = g(t) - \sum_{0\le s\le t} \Delta g(s), \quad t \ge 0$$

(take the current value of g and remove all discontinuous jumps to make it continuous). Moreover, for a process Z, we denote its quadratic variation by $\langle Z, Z\rangle_t$. We now briefly mention a generalization of Itô's formula whose form clearly has clear parallels to the differential equations in the next sections below and plays an important role in the proofs of existence of solutions.

Lemma 3.1 (Extant Second Derivative Meyer-Itô [6] pg. 221, Thm 71)**.** *Let Z be a semimartingale. Let $f \in C^1$ with f' absolutely continuous such that $f'(b) - f'(a) = \int_a^b f''(x)\,\mathrm{d}x$ for any $a, b \in \mathbb{R}$, with f'' locally L^1. Then,*

$$f(Z_t) = f(Z_0) + \int_{0+}^{t} f'(Z_{s-})\,\mathrm{d}Z_s + \frac{1}{2}\int_0^t f''(Z_{s-})\,\mathrm{d}\langle Z, Z\rangle_s^c + \sum_{0<s\le t}\Big(f(Z_s) - f(Z_{s-}) - f'(Z_{s-})\Delta Z_s\Big).$$

Note that when Z and f'' are continuous we have $Z_{s-} = Z_s$ and $f''(Z_{s-}) = f''(Z_s)$ thus recovering the standard Itô's formula. The formula above can be interpretted as applying the standard Itô formula to Z at the left limit of every point to account for the discontinuities, then adding back the effect of the discontinuities with the last summation term. This leads to another Itô formula applied to the inventory semimartingale process,

Lemma 3.2 ([2] lemma 3.1)**.** *Let Z_t be the semimartingale process introduced in section 2.2 controlled under the policy $\varphi = (Y_1, Y_2)$. Let $f \in C^1(\mathbb{R})$ with f' absolutely continuous such that $\forall a < b \in \mathbb{R}$, $f'(b) - f'(a) = \int_a^b f''(u)\,\mathrm{d}u$, with f'' locally L^1. Then,*

$$f(Z_t) = f(Z_0) + \int_0^t \Gamma f(Z_s)\,\mathrm{d}s + \sigma\int_0^t f'(Z_s)\,\mathrm{d}W_s + \int_0^t f'(Z_{s-})\,\mathrm{d}Y_1^c(s) - \int_0^t f'(Z_{s-})\,\mathrm{d}Y_2^c(s) + \sum_{0<s\le t}\Delta f(Z_s).$$

Here $\Gamma f = \frac{1}{2}\sigma^2 f'' + \mu f'$ is the generator of the Brownian motion X (as a Markov process) $\forall x \in \mathbb{R}$ where $f''(x)$ is defined.

This follows in a few direct steps from lemma 3.1 using the fact $\langle Z, Z\rangle_t = \langle X, X\rangle_t = \sigma^2 t$. Recall that the generator of a Markov process characterizes how fast the process evolves through time and thus $\Gamma f(x)$ measures how the expected value of f changes through time when the Brownian process takes the value x. This plays an important role in understanding the differential equations presented in the following sections to prove optimality of control policies.

4 Lower Bounding Average Cost

In this section we present a tool to obtain a lower bound on the average cost presented in (2.1) under any feasible policy regardless of the inventory process's starting point. The generality of this theorem is especially of interest as it provides a set of conditions that, if satisfied, provide a lower bound across the entire space of feasible control policies without the need to consider any specific policy. Similar statements of this theorem can be found in the literature under the name the "verification theorem" (see [4] for example), since given a policy φ and its associated $\mathrm{AC}(\varphi, x)$, it allows us to verify if it is an optimal policy (i.e., if it achieves the lower bound; of course this is only a sufficient and not a necessary condition for optimality since an obtained lower bound may not be tight).

Theorem 4.1 ([2] thm 4.1). *Let $f \in C^1(\mathbb{R})$ with f' absolutely continuous and uniformly bounded on $\mathbb{R}$, and f'' locally L^1. Further assume that f satisfies the boundary value differential inequality,*

$$\Gamma f(x) + h(x) \geq \gamma \text{ for almost all } x \in \mathbb{R} \tag{4.1}$$

$$f(y) - f(x) \leq K + k(x - y) \text{ for } y < x \tag{4.2}$$

$$f(y) - f(x) \leq L + \ell(y - x) \text{ for } y > x. \tag{4.3}$$

Then $AC(\varphi, x) \geq \gamma$ for each feasible policy φ and each initial state $x \in \mathbb{R}$.

The statement of this theorem is a bit abstract given that it admits any sufficiently smooth function that satisfies the inequalities. However, it is natural to reason that if we can find a function that makes at least some of the inequalities in (4.1)-(4.3) into equalities, then we also get an equality for the AC. This is indeed the intuition behind the methods of the next two sections.

4.1 Deriving Optimality from Lower Bounds

In both sections 5 and 6 we use the following procedure to prove optimality of a control policy:

1. We implicitly define a *relative value function* $V(x)$ of a policy as the solution to a second order differential equation (which parallels the form of (4.1)-(4.3) but with equalities in place of inequalities). $V(x)$ can be viewed as measuring the relative value of $AC(\varphi, x)$ (the cost up to a constant).

2. Show that a solution function V exists to this differential system such that $\mathrm{AC}(\varphi, x) = \gamma$ is a constant for all starting points x.

3. Show that (an extension of) this function can be admitted as the function f in theorem 4.1 and thus $\mathrm{AC}(\varphi', x) \geq \gamma$ for all starting points x and feasible policies φ'. Since we know $\mathrm{AC}(\varphi, x) = \gamma$, φ achieves the lower bound and thus is optimal.

5 Optimal Impulse Control

Here we study the control problem in the case where $K, L > 0$, i.e., every adjustment to the inventory process incurs a fixed positive cost.

5.1 Characterization of Feasible Policies

Immediately it is clear that any optimal control policy must be limited to a finite number of adjustments, $N_1(t), N_2(t) < \infty$ almost surely for all $t \geq 0$, since otherwise we would incur the fixed adjustment cost an infinite number of times giving $\mathrm{AC}(\varphi, x) = \infty$ (which is not optimal; we could do better by not imposing any control at all). Thus to describe a policy $\varphi = (Y_1, Y_2)$, it is sufficient to specify the discrete sequence of upward and downward adjustments along with when they occur. Namely, we can fully characterize the process of upward adjustment process Y_1 by the sequence $\left(T_n^1, \xi_n^1\right)_{n \in \mathbb{N}}$, where each T_n^1 is an $(\mathcal{F}_t)_{t \geq 0}$-stopping time representing the time at which the n-th upward adjustment was made (which can be defined recursively as $T_n^1 := \inf\left\{t > T_{n-1}^1 \mid \Delta Y_1(t) > 0\right\}$)

and $\xi_n^1 \in \mathcal{F}_{T_n-}$ is the amount of the n-th upward adjustment (more precisely, $\xi_n^1 = \Delta Y_1(T_n^1)$). A similar sequence $\left(T_n^2, \xi_n^2\right)_{n\in\mathbb{N}}$ can be defined to characterize Y_2. With this setup we then have,

$$Y_1(t) = \sum_{i=0}^{N_1(t)} \xi_i^1, \qquad Y_2(t) = \sum_{i=0}^{N_2(t)} \xi_i^2$$

These two sequences can be merged into a single sequence $\left(T_n, \xi_n\right)_{n\in\mathbb{N}}$ where the elements of $\left(\xi_n^2\right)_{n\in\mathbb{N}}$ are negated to differentiate between upward and downward movements to characterize the entire policy φ.

5.2 The Optimal Policy

We show that a banded control policy determined by four parameters $\varphi = (d, D, U, u)$ with $d < D < U < u$ is optimal (where the optimal parameters need to be determined). Such a policy works by waiting until the process Z_t hits u or d and then immediately adjusting the process to D or U respectively. Thus every adjustment incurs either a cost of $K + k(D - d)$ if adjusting from d to D or a cost of $L + \ell(u - U)$ if adjusting from u to U. The intuition behind such a policy is that we do not want to make too many adjustments or else we will incur too much fixed adjustment costs, so we wait until the inventory process hits some value at which the holding cost is deemed "too high" and then a large impulse adjustment is made to a lower/higher point (note we have symmetry in the holding cost about the minimum from assumption 2.1), being optimistic that from there the process will find its way to a level with lower holding cost. In this banded policy we are also providing the process a range of "cushion" on (d, D) and (U, u) where we once again give the process the opportunity to find its way to a place with lower holding cost before it hits the bands and requires adjustment, giving the controller a chance of not incurring an adjustment cost.

To show that there exists a banded policy φ that optimal, we use the "memoryless" nature of Brownian motion and thus the inventory process Z, i.e., we can reason about the process as it moves forward from some point ignorant of what happened leading up to that point. Thus we can study the behavior of Z when $Z_t = x \in [d, u]$ regardless of what happened before. The idea is that if at all possible values that Z can take, its instantaneous expected average cost is some constant γ, then it's reasonable to think that the expected average long run cost is also γ independent of where the process is started. We can directly measure the holding cost at any instant when $Z_t = x$ as $h(x)$, however, measuring the adjustment cost at an instant cannot be done as directly. We take advantage of the statement at the end of section 3 about the generator Γ to do this.

Suppose we had a function $V(\cdot)$ such that $V(a) - V(b)$ measured the difference in cost incurred as we exercise the policy φ if we start the process at a as opposed to from b. Then note that if $V(a) - V(b) > 0$, it costs more to apply the policy from the process starting at a, thus we can expect the difference in long-run cost to be the adjustment cost from a to b from whence we can run the policy to get the lower cost from b. Now, $\Gamma V(x)$ represents the rate at which V changes as the Brownian process propagates through time if it starts at x, so by the interpretation above it can be viewed as a proxy for the instantaneous adjustment cost when $Z = x$. Therefore, we can use $\Gamma V(x) + h(x)$ (adjustment cost + holding cost) as a measure of the cost of the inventory process at the instant that $Z = x$. Moreover, we can define V implicitly through boundary conditions based the behavior of the policy and the corresponding adjustment costs at the bands as shown in the theorem below.

Theorem 5.1 ([2] thm 5.1)**.** *Fix a control band policy $\varphi = (d, D, U, u)$. If $\exists \gamma \in \mathbb{R}$ and a twice continuously differentiable function $V : [d, u] \to \mathbb{R}$ satisfying*

$$\Gamma V(x) + h(x) = \gamma, \quad d \le x \le u \tag{5.1}$$

with boundary conditions

$$V(d) - V(D) = K + k(D - d) \tag{5.2}$$
$$V(u) - V(U) = L + \ell(u - U), \tag{5.3}$$

then $AC(\varphi, x) = \gamma$ independent of the starting point $x \in \mathbb{R}$.

The solution V to this boundary value differential equation problem is called the *relative value function* of the policy φ and has the desired properties stated above. Note that theorem 5.1 does not provide us the existence of a solution, but rather tells us that if a solution exists, then we can determine the average cost. The existence of a solution is provided by Proposition 1 of [2], which gives an explicit form for $V = \int_m^t g(y)\, \mathrm{d}y$ for any fixed $m \in \mathbb{R}$ in terms of d, D, U, u (we omit the form of g as it's quite messy).

A similar intuition can be applied to understanding thm 5.1 of [3], which provides a parallel result to theorem 5.1 but for the discounted cost case. In this case, equation (5.1) is replaced with,

$$\Gamma V(x) + h(x) = \beta V(x), \quad d \le x \le u.$$

The LHS can be interpretted in the same way as before and the RHS is the total cost of controlling the process starting from x discounted back to the current time. Thus, in a similar manner as above, this equation is requiring the instantaneous cost of the process when it takes a value x to be the same as the discounted cost if it's started at x and controlled by the policy.

Having established the existence of the relative value function V for a banded policy φ, we wish to now find a version of it that satisfies the assumptions of theorem 4.1 enabling us to conclude that the policy achieves lower bound and thus is optimal. The following theorem provides us with the existence of the optimal banded policy. Note that the problem at hand is a *free boundary differential equation* problem as we are finding both the solution function V and the boundary endpoints which come as part of the optimal policy as part of the solution.

Theorem 5.2 ([2] thm 5.2)**.** *Under assumption 2.1, there exists d^*, D^*, U^*, u^* and $x_1, x_2 \in \mathbb{R}$ with*

$$d^* < x_1 < D^* < U^* < x_2 < u^*$$

such that the corresponding solution $V(x) = \int_m^x g(u)\, \mathrm{d}u$ satisfies,

$$\int_{d^*}^{D^*} g(u) + k \, \mathrm{d}u = -K$$
$$\int_{U^*}^{u^*} g(u) - \ell \, \mathrm{d}u = L$$
$$g(d^*) = g(D^*) = -k$$
$$g(U^*) = g(u^*) = \ell.$$

Furthermore, g has a local minimum at $x_1 < a$ and a local minimum at $x_2 > a$. The function g is decreasing on $(-\infty, x_1)$ and increasing on (x_2, ∞).

Taking this V^* corresponding to the optimal policy φ^* given by the above theorem, The idea is to extend V from $[d, u]$ to the real line in the intuitive way, adjusting any values outside of $[d, u]$ to be within $[D, U]$ and incurring the adjustment cost,

$$f(x) = \begin{cases} K + k(D^* - x) + V(D^*) & \text{for } x < d^* \\ V(x) & x \in [d^*, u^*] \\ L + \ell(x - U^*) + V(U^*) & \text{for } x > u^*. \end{cases}$$

Let $\gamma^* = \mathrm{AC}(\varphi^*, x)$ (the starting point x does not matter as we determined under the conditions of theorem 5.1, the average cost is independent of the starting point). By the conditions on $g(x) = V'(x)$ in theorem 5.2, f satisfies the assumptions of theorem 4.1 and thus we can conclude that the average cost is lower bounded by γ^*. However, φ^* is a banded policy and achieves a cost of γ^*, thus it is optimal.

6 Optimal Singular Control

Here we study the control problem in the case where $K, L = 0$, i.e., adjustments to the inventory process incur only proportional costs.

6.1 The Optimal Policy

In this case, it turns out that the optimal control policy is a two-banded policy $\varphi = (d, u)$, which acts by applying an infinitesimally small adjustment to the inventory process if it hits d or u to keep it within $[d, u]$; under such a policy, we have that the number of upward and downward adjustments is infinite with positive probability. The intuition behind such a policy is to use bands that keep the inventory from getting to a region where the holding cost is "too high," but we no longer need the space of cushion as we did in the impulse control case, since there is no cost incurred for repeated adjustments, only the proportional cost. Thus, every time, the process reaches a region where the holding cost is undesirably large, we nudge it back to the region we want it to be in and hope that it turn itself towards the minimum of the holding cost function.

The proof of optimality of such a banded policy is nearly identical to the process followed in section 5 for the impulse control case (see section 4.1 for the general procedure), with a few slight modifications to the involved theorems to account for the reduction in policy parameters to just two from four.

7 Non-negative Inventory Processes

The methods and resulting optimal policies presented in the previous sections generalize readily to the case where we do not allow the inventory to be negative, where we require $\forall t \geq 0$, $|Z(t)| \geq 0$. It suffices to change the domain of functions appearing in the corresponding theorems in the previous section to be on $[0, \infty)$ (their proofs remain similar as well). Moreover, in the impulse control case $(K, L > 0)$ it is clear that the optimal banded policy $\varphi^* = (d^*, D^*, U^*, u^*)$ must satisfy $d^* \geq 0$ since otherwise Z_t will never hit d^* (in some cases it is optimal to have $d^* = 0$, reducing the problem of finding the optimal policy to finding just three parameters). The main technical differences arise in the proof of the existence of a solution to the free boundary problem giving the optimal banded policy parameters and corresponding value function (parallel of theorem 5.2). These differences are highlighted in greater detail in section 7 of [2], however the mechanics and motivation underlying them remain the same.

References

[1] Buckley, I., and Korn, R. Optimal index tracking under transaction costs and impulse control. *International journal of theoretical and applied Finance 1*, 03 (1998), 315–330.

[2] Dai, J. G., and Yao, D. Brownian inventory models with convex holding cost, part 1: Average-optimal controls. *Stochastic Systems 3*, 2 (Dec. 2013), 442–499.

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

[4] Feng, H., and Muthuraman, K. A computational method for stochastic impulse control problems. *Mathematics of Operations Research 35*, 4 (2010), 830–850.

[5] Le Gall, J.-F. *Brownian motion, martingales, and stochastic calculus*. Springer, 2016.

[6] Protter, P. E., and Protter, P. E. *Stochastic differential equations*. Springer, 2005.