Linear Programming Approach to the Optimal Stopping of Singular Stochastic Processes

Kurt Helmes[1], and Richard H. Stockbridge[2]

Abstract. Optimal stopping of stochastic processes having both absolutely continuous and singular behavior (with respect to time) can be equivalently formulated as an infinite-dimensional linear program over a collection of measures. These measures represent the occupation measures of the process (up to a stopping time) with respect to "regular time" and "singular time" and the distribution of the process when it is stopped. Such measures corresponding to the process and stopping time are characterized by an adjoint equation involving the absolutely continuous and singular generators of the process. This general linear programming formulation is shown to be numerically tractable through three examples, each of which seeks to determine the stopping rule for a perpetual lookback put option using different dynamics for the asset price. Exact solutions are determined in the cases that the asset price are given by a drifted Brownian motion and a geometric Brownian motion. Numerical results for the more realistic model of a regime switching geometric Brownian motion are also presented, demonstrating that the linear programming methodology is numerically tractable for models whose theoretical solutions are very difficult to obtain.

Key Words. optimal stopping, singular processes, linear programming, regime switching, lookback option, martingale problem, occupation measure.

MSC 2000 Classifications. 60G40, 60J55, 90C05.

1 Introduction

Many processes of interest in applications (see, for example, the survey paper by Shreve [19]) can be modelled as solutions to a stochastic differential equation of the form

$$dX(t) = b(X(t), u(t))dt + \sigma(X(t), u(t))dW(t) + m(X(t-), u(t-))d\xi(t) \tag{1}$$

where X is the state process with $E = \mathbb{R}^d$, u is a control process with values in U_0, ξ is a nondecreasing process arising either from the boundary behavior of X (e.g., the local time on the boundary for a reflecting diffusion) or from a singular control, and W is a Brownian motion. Processes in which the set of times of increase of ξ has Lebesgue measure 0 are called singular stochastic processes. This paper restricts its attention to optimal stopping problems involving uncontrolled singular processes.

The purpose of the paper is to extend the equivalence given in [2] of a linear programming formulation for optimal stopping problems for stochastic processes whose behavior occurs absolutely continuously in time so that it applies to stochastic processes which also have singular behavior in time. Key to the extension is the formulation of singular stochastic

[1]Institute for Operations Research, Humboldt University of Berlin, helmes@wiwi.hu-berlin.de.

[2]Department of Mathematical Sciences, University of Wisconsin Milwaukee, stockbri@uwm.edu, This research has been supported in part by the U.S. National Security Agency under Grant Agreement Number H98230-05-1-0062. The United States Government is authorized to reproduce and distribute reprints notwithstanding any copyright notation herein.

processes via a singular martingale problem for its generators and the existence of a solution to the martingale problem corresponding to measures which satisfy an adjoint relation involving the generators applied to test functions.

The resulting linear programs are typically infinite-dimensional whose variables are the occupation measures of the processes according to both regular time and singular time and the distribution of the state at the stopping time. The paper also investigates the numerical implementation of the LP formulation. Naturally, it is necessary to approximate the infinite-dimensional LP by a finite-dimensional LP. This paper demonstrates one such approximation approach and demonstrates that it provides accurate solutions.

The use of occupation measures to solve optimization problems began with Young [22] in the context of calculus of variations. The present work is an outgrowth of linear programming formulations for stochastic control that was initiated by Manne [15] in discrete time and which has been developed for Markov decision problems (see e.g. [9, 10, 11]). Linear programming for continuous time stochastic control has been established under very general conditions in [1, 13, 20, 21].

The processes in this paper are *not* controlled. However, the results cited above are applicable to these processes by taking the control space to consist of a single element.

To illustrate the accuracy of the numerics and to demonstrate the power of the method, this paper considers an optimal stopping problem of the form

$$E\left[e^{-\lambda\tau}\left(Y(\tau)-X(\tau)\right)\right] \tag{2}$$

in which the state process X denotes the price of an asset, Y denotes the running maximum process of the asset price, $\lambda > 0$ gives the discount rate and τ denotes the option holder's exercise time. The quantity used for illustration is a perpetual lookback put option and the goal is to determine an optimal stopping rule. This problem is very similar to the Russian option studied by Shepp and Shiryaev [18], Graverson and Peskir [5] and Peskir [17].

The paper considers three models for the asset price X, each having starting value x_0: a drifted Brownian motion process with drift parameter a, the slightly more realistic model of geometric Brownian motion and an extension of this model to include regime switching. These models have been chosen since we are able to determine exact solutions to the optimal stopping problem for the first two models, and therefore can evaluate the accuracy of the numerical solutions, whereas the last example illustrates that linear programming provides a viable solution technique for more complex models for which exact solutions are not known.

The next section formulates the singular martingale problem, establishes the key existence result and demonstrates equivalence of the linear programming formulation with the original stochastic process formulation for the optimal stopping problem. The following sections provide the numerical illustrations.

2 Stochastic Formulation and LP Reformulation

2.1 Formulation of martingale problem

For a complete, separable, metric space S, we define $M(S)$ to be the space of Borel measurable functions on S, $B(S)$ to be the space of bounded, measurable functions on S, $C(S)$

to be the space of continuous functions on S, $\overline{C}(S)$ to be the space of bounded, continuous functions on S, $\mathcal{M}(S)$ to be the space of finite Borel measures on S, and $\mathcal{P}(S)$ to be the space of probability measures on S. $\mathcal{M}(S)$ and $\mathcal{P}(S)$ are topologized by weak convergence.

Let $\mathcal{L}_t(S) = \mathcal{M}(S \times [0,t])$. We define $\mathcal{L}(S)$ to be the space of measures ξ on $S \times [0,\infty)$ such that $\xi(S \times [0,t]) < \infty$, for each t, and topologized so that $\xi_n \to \xi$ if and only if $\int f d\xi_n \to \int f d\xi$, for every $f \in \overline{C}(S \times [0,\infty))$ with $\mathrm{supp}(f) \subset S \times [0,t_f]$ for some $t_f < \infty$. Let $\xi_t \in \mathcal{L}_t(S)$ denote the restriction of ξ to $S \times [0,t]$. Note that a sequence $\{\xi^n\} \subset \mathcal{L}(S)$ converges to a $\xi \in \mathcal{L}(S)$ if and only if there exists a sequence $\{t_k\}$, with $t_k \to \infty$, such that, for each t_k, $\xi^n_{t_k}$ converges weakly to ξ_{t_k}, which in turn implies ξ^n_t converges weakly to ξ_t for each t satisfying $\xi(S \times \{t\}) = 0$. Finally, we define $\mathcal{L}^{(m)}(S) \subset \mathcal{L}(S)$ to be the set of ξ such that $\xi(S \times [0,t]) = t$ for each $t > 0$.

Throughout, we will assume that the state space E is a complete, separable, metric space.

Let $A, B : \mathcal{D} \subset \overline{C}(E) \to C(E)$ and $\nu_0 \in \mathcal{P}(E)$. Let X be an E-valued process and Γ be an $\mathcal{L}(E)$-valued random variable. Let Γ_t denote the $\mathcal{L}(E)$-random variable which, for each realization, is defined by restricting Γ to $E \times [0,t]$. Then (X,Γ) is a solution of the *singular martingale problem* for (A, B, ν_0) if there exists a filtration $\{\mathcal{F}_t\}$ such that (X, Γ_t) is $\{\mathcal{F}_t\}$-progressive, $X(0)$ has distribution ν_0, and for every $f \in \mathcal{D}$,

$$f(X(t)) - f(X(0)) - \int_0^t Af(X(s))\, ds - \int_{E\times[0,t]} Bf(x)\Gamma(dx \times ds) \tag{3}$$

is an $\{\mathcal{F}_t\}$-martingale.

Note we refer to A as being the absolutely continuous generator of X and B as the singular generator. This is an intuitive labelling. Strictly speaking, the random measure Γ may have an absolutely continuous part.

2.2 Conditions on A and B

We assume that the absolutely continuous generator A and the singular generator B have the following properties.

Condition 2.1

i) $A, B : \mathcal{D} \subset \overline{C}(E) \to C(E)$, $1 \in \mathcal{D}$, *and* $A1 = 0, B1 = 0$.

ii) Defining $(A_0, B_0) = \{(f, Af, Bf) : f \in \mathcal{D}\}$, (A_0, B_0) *is* separable *in the sense that there exists a countable collection* $\{g_k\} \subset \mathcal{D}$ *such that* (A_0, B_0) *is contained in the bounded, pointwise closure of the linear span of* $\{(g_k, Ag_k, Bg_k)\}$.

iii) $\mathcal{D}$ *is closed under multiplication and separates points.*

Example 2.2 *Reflecting diffusion processes.*

The most familiar class of processes of the kind we consider are reflecting diffusion processes satisfying equations of the form

$$X(t) = X(0) + \int_0^t \sigma(X(s))dW(s) + \int_0^t b(X(s))ds + \int_0^t m(X(s))d\xi(s),$$

where X is required to remain in the closure of a domain D (assumed smooth for the moment) and ξ increases only when X is on the boundary of D. Then

$$Af(x) = \frac{1}{2}\sum_{i,j} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} f(x) + b(x)\cdot\nabla f(x),$$

where $a(x) = ((a_{ij}(x))) = \sigma(x)\sigma(x)^T$. In addition, under reasonable conditions ξ will be continuous, so

$$Bf(x) = m(x)\cdot\nabla f(x).$$

Example 2.3 *Diffusion with jumps away from the boundary.*

Assume that D is an open domain and that for $x \in \partial D$, $m(x)$ satisfies $x + m(x) \in D$. Also assume that

$$X(t) = X(0) + \int_0^t \sigma(X(s))dW(s) + \int_0^t b(X(s))ds + \int_0^t m(X(s-))d\xi(s),$$

where ξ is required to be the counting process that "counts" the number of times that X has hit the boundary of D, that is, assuming $X(0) \in D$, X diffuses until the first time τ_1 that X hits the boundary ($\tau_1 = \inf\{s > 0 : X(s-) \in \partial D\}$) and then jumps to $X(\tau_1) = X(\tau_1-) + m(X(\tau_1-))$. The diffusion then continues until the next time τ_2 that the process hits the boundary, and so on. (In general, this model may not be well-defined since the $\{\tau_k\}$ may have a finite limit point, but we will not consider that issue.) Then A is the ordinary diffusion operator, $Bf(x) = f(x + m(x)) - f(x)$, and $\Gamma(H \times [0,t]) = \int_0^t I_H(X(s-))d\xi(s)$.

Suppose that (X,Γ) is a solution of the singular martingale problem for (A, B, ν_0) and let τ be a stopping time satisfying the conditions of the optional sampling theorem (see [3, Theorem 2.2.13]). It follows from (X,Γ) satisfying (3) that when we augment a time component to the state space by considering the process (t,X),

$$\begin{aligned} \gamma(t)f(X(t)) - \gamma(0)f(X(0)) &- \int_0^t \left[\gamma(s)Af(X(s)) + \gamma'(s)f(X(s))\right] ds \\ &- \int_{\mathbb{R}^+\times E\times[0,t]} \gamma(s)Bf(x)\Gamma(ds\times dx\times dr) \end{aligned} \tag{4}$$

is also an $\{\mathcal{F}_t\}$-martingale for $\gamma \in \widehat{C}(\mathbb{R}^+)$ and $f \in \mathcal{D}$. It then follows from the optional sampling theorem that for each $\gamma \in \widehat{C}(\mathbb{R}^+)$ and $f \in \mathcal{D}$,

$$\begin{aligned} E\Bigg[\gamma(\tau)f(X(\tau)) - \gamma(0)f(X(0)) &- \int_0^\tau \tilde{A}[\gamma f](s, X(s))\, ds \\ &- \int_{E\times[0,\tau]} \tilde{B}[\gamma f](s,x)\,\Gamma(ds\times dx\times dr)\Bigg] = 0, \end{aligned} \tag{5}$$

where $\tilde{A}[\gamma f] = \gamma Af + \gamma' f$ and $\tilde{B}[\gamma f] = \gamma Bf$. Let ν_τ denote the distribution of $(\tau, X(\tau))$ and define the expected "occupation measures" on $\mathbb{R}^+ \times E$ by

$$\begin{aligned} \mu_0(G) &= E\left[\int_0^\tau I_G(s, X(s))\, ds\right], \qquad \forall G \in \mathcal{B}(\mathbb{R}^+\times E), \\ \mu_1(G) &= E\left[\int_{\mathbb{R}^+\times E\times[0,\tau]} I_G(s,x)\,\Gamma(ds\times dx\times dr)\right], \qquad \forall G \in \mathcal{B}(\mathbb{R}^+\times E). \end{aligned}$$

Notice that μ_0 is the occupation measure of the process according to "regular" time (Lebesgue measure on $\mathbb{R}^+$) and μ_1 is the occupation measure of X according to the singular set of times at which Γ increases. We refer to μ_1 as the "singular" occupation measure. It then follows immediately from (5) that the measures ν_τ, μ_0 and μ_1 satisfy

$$\begin{aligned} 0 = \int \gamma(t) f(x)\, \nu_\tau(dx) - \gamma(0) \int f(x)\, \nu_0(dx) \quad &- \quad \int \tilde{A}[\gamma f](r,x)\, \mu_0(dr \times dx) \\ &- \quad \int \tilde{B}[\gamma f](x)\, \mu_1(dr \times dx), \qquad \forall f \in \mathcal{D}. \end{aligned} \tag{6}$$

Thus given any solution (X, Γ) to the singular martingale problem for (A, B) and any sufficiently nice τ, the measures ν_τ, μ_0 and μ_1 will satisfy (6). The main theoretical result is the converse, namely that any measures ν_τ, μ_0 and μ_1 satisfying (6) are related to some solution (X, Γ) of the singular martingale problem and some stopping time τ. We use the notation μ_2 for ν_τ in the theorem since the existence of a stopping time τ corresponding to the measure μ_2 is part of the result and this choice of notation will not give the impression that τ exists *a priori.*

Theorem 2.4 *Suppose E is a complete, separable metric space and A and B satisfy Conditions 2.1. Let $\nu_0 \in \mathcal{P}(E)$. Define $\mathcal{D}_1 = \{\gamma f : \gamma \in \widehat{C}(\mathbb{R}^+), f \in \mathcal{D}\}$. Suppose $\mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^+ \times E)$ and $\mu_2 \in \mathcal{P}(\mathbb{R}^+ \times E)$ satisfy*

$$\begin{aligned} \int_{\mathbb{R}^+ \times E} \tilde{A}[\gamma f](s,x)\, \mu_0(ds \times dx) + \int_{\mathbb{R}^+ \times E} \tilde{B}[\gamma f](s,x) \mu_1(ds \times dx) \\ + \gamma(0) \int f\, d\nu_0 - \int_{\mathbb{R}^+ \times E} \gamma(s) f(x) \mu_2(ds \times dx) = 0, \qquad \forall \gamma f \in \mathcal{D}_1. \end{aligned} \tag{7}$$

Then there exist a process X adapted to a filtration $\{\mathcal{F}_t\}$, a random measure Γ on $\mathbb{R}^+ \times E \times [0, \infty)$ and an $\{\mathcal{F}_t\}$-stopping time τ such that

$$\gamma(t \wedge \tau) f(X(t \wedge \tau)) - \int_0^{t \wedge \tau} \tilde{A}[\gamma f](s, X(s))\, ds - \int_{\mathbb{R}^+ \times E \times [0, t \wedge \tau)} \tilde{B}[\gamma f](s,x) \Gamma(ds \times dx \times dv)$$

is an $\{\mathcal{F}_t\}$-martingale for every $\gamma f \in \mathcal{D}_1$, and

$$\begin{aligned} E\left[\int_0^\tau c_0(s, X(s))\, ds + \int_{E \times [0,\tau)} c_1(s,x)\, \Gamma(ds \times dx \times dv) + c_2(\tau, X(\tau)) \right] \\ = \int_{\mathbb{R}^+ \times E} c_0(s,x)\, \mu_0(ds \times dx) + \int_{\mathbb{R}^+ \times E} c_1(s,x)\, \mu_1(ds \times dx) + \int_{\mathbb{R}^+ \times E} c_2(s,x) \mu_2(ds \times dx) \end{aligned} \tag{8}$$

for every $c_0, c_1, c_2 \in M(\mathbb{R}^+ \times E)$ that are bounded below.

Proof. Let $r \in \mathbb{R}^+$ be an additional state component and let $u \in U = \{0, 1\}$ be a control variable. We wish to demonstrate the existence of a state process $(\overline{R}, \overline{S}, \overline{X})$ such that $\overline{S}$ measures regular time up to some stopping time, $\overline{X}$ gives the original dynamics up to this stopping time and $\overline{R}$ is a new time process which starts at the stopping time, lasts for an exponentially distributed length of time, and during this time $(\overline{S}, \overline{X})$ are fixed. At the occurance of the exponential time, $\overline{S}$ jumps to 0, $\overline{X}$ jumps to a new point according to the initial distribution ν_0 and $\overline{R}$ jumps to 0. The desired solution is then obtained using both a time change and a change of measures.

Define the new generators $\overline{A}$ and $\overline{B}$ on the domain $\mathcal{D}_2 = \{\psi\gamma f : \psi \in \widehat{C}^1(\mathbb{R}^+), \gamma f \in \mathcal{D}_1\}$ such that

$$\begin{aligned}\overline{A}[\psi\gamma f](r,s,x,u) &= u\psi(r)[\gamma(s)Af(x) + \gamma'(s)f(x)] \\ &\quad +(1-u)\left[\psi(0)\gamma(0)\int f\, d\nu_0 - \psi(r)\gamma(s)f(x) + \psi'(r)\gamma(s)f(x)\right] \\ \overline{B}[\psi\gamma f](r,s,x) &= \psi(r)\gamma(s)Bf(x).\end{aligned}$$

Observe that $\overline{A}$ is the generator of a *controlled* process with control variable u. Also define the measures $\overline{\mu}_0 \in \mathcal{P}(\mathbb{R}^+ \times \mathbb{R}^+ \times E \times U)$ and $\overline{\mu}_1$ on $\mathcal{M}(\mathbb{R}^+ \times \mathbb{R}^+ \times E)$ satisfying

$$\begin{aligned}&\int h(r,s,x,u)\overline{\mu}_0(dr \times ds \times dx \times du) \qquad (9)\\ &= K^{-1}\left(\int_{\mathbb{R}^+\times E} h(0,s,x,1)\mu_0(ds \times dx) + \int_{\mathbb{R}^+\times E}\int_0^\infty e^{-r}h(r,s,x,0)\, dr\, \mu_2(ds \times dx)\right)\end{aligned}$$

$$\int h(r,s,x)\,\overline{\mu}_1(dr \times ds \times dx) = K^{-1}\int_{\mathbb{R}^+\times E} h(0,s,x)\,\mu_1(ds \times dx) \qquad (10)$$

for each bounded, continuous h, where $K = \mu_0(\mathbb{R}^+ \times E) + 1$ is a normalizing constant. Note that the conditional distribution of u given (r,s,x) under $\overline{\mu}_0$ is

$$\overline{\eta}(r,s,x,du) = \delta_{\{1\}}(du)I_{\{0\}}(r)\frac{d\mu_0}{d\overline{\mu}_0}(s,x) + \delta_{\{0\}}(du)e^{-r}I_{(0,\infty)}(r)\frac{d\mu_2}{d\overline{\mu}_0}(s,x).$$

Since μ_0, μ_1 and μ_2 satisfy (7), it immediately follows from integration by parts that for $\psi\gamma f \in \mathcal{D}_2$

$$\int \overline{A}[\psi\gamma f]\, d\overline{\mu}_0 + \int \overline{B}[\psi\gamma f]\, d\overline{\mu}_1 = 0.$$

Since the control space $U = \{0,1\}$ is compact, the conditions of Theorem 1.7 of [14] are satisfied by $\overline{A}$, $\overline{B}$, $\overline{\mu}_0$ and $\overline{\mu}_1$, which implies the existence of a stationary $\mathbb{R}^+ \times \mathbb{R}^+ \times E$-valued process $(\overline{R}, \overline{S}, \overline{X})$ and a random measure $\overline{\Gamma}$ having stationary increments such that

$$\begin{aligned}\psi(\overline{R}(t))\gamma(\overline{S}(t))f(\overline{X}(t)) &- \int_0^t\int_U \overline{A}[\psi\gamma f]((\overline{R}(s), \overline{S}(s), \overline{X}(s), u)\,\overline{\eta}(\overline{R}(s), \overline{S}(s), \overline{X}(s), du)\, ds \\ &- \int_{\mathbb{R}^+\times\mathbb{R}^+\times E\times[0,t)} \overline{B}[\psi\gamma f](r,s,x,v)\,\overline{\Gamma}(dr \times ds \times dx \times dv) \qquad (11)\end{aligned}$$

is an $\{\mathcal{F}_t^{\overline{R},\overline{S},\overline{X}}\}$-martingale for all $\psi\gamma f \in \mathcal{D}_2$, $(\overline{R}, \overline{S}, \overline{X})$ has stationary distribution $\overline{\mu}_0(\cdot \times U)$ and for each t, $E[\overline{\Gamma}(\cdot\times[0,t])] = t\overline{\mu}_1$. Without loss of generality, we may assume that $(\overline{R}, \overline{S}, \overline{X})$ is defined for all $t \in \mathbb{R}$ and that $\overline{\Gamma}$ is a measure on $\mathbb{R}^+ \times \mathbb{R}^+ \times E \times (-\infty, \infty)$. Observe that $\overline{\eta}(r,s,x,\cdot)$ places full mass on $u = 1$ when $r = 0$ and full mass on $u = 0$ when $r > 0$.

For each $t \geq 0$, define the random variable $\sigma_0^t = \sup\{r < t : \overline{S}(r) = 0, \overline{R}(r) = 0\}$ as the last time previous to t that the time components are both 0, and define the sequences of stopping times $\sigma_1^t = \inf\{r \geq t : \overline{R}(r) = 0\}$, $\tau_1^t = \int\{r > \sigma_1^t : \overline{R}(r) > 0\}$, and $\sigma_2^t = \inf\{r > \tau_{k-1}^t : \overline{R}(r) = 0\}$. Observe that σ_1^t is the first time after t that the time component $\overline{R}$ is 0, τ_1^t is the first time after σ_1^t at which $\overline{R}$ begins to increase and σ_2^t is the first time after τ_1^t that $\overline{R}$ is again 0. Thus for $s \in [\sigma_1^t, \tau_1^t)$, Lemma 4.4 of [13] implies that $\overline{S}(s) = \int_{\sigma_1^t}^s I_{\{0\}}(\overline{R}(r))\, dr = s - \sigma_1^t$. On the other hand, for $s \in [\tau_1^t, \sigma_2^t)$, [13, Lemma 4.4] implies $\overline{R}(s) = \int_{\tau_1^t}^s I_{(0,\infty)}(\overline{R}(r))\, dr = s - \tau_1^t$ a.s. and conditional on $\mathcal{F}_{\tau_1^t}$, $\sigma_2^t - \tau_1^t$ is exponentially distributed with mean 1, and again by [13, Lemma 4.4] $\overline{S}(s) = \overline{S}(\tau_1^t) + \int_{\tau_1^t}^s I_{\{0\}}(\overline{R}(r))\, dr = \overline{S}(\tau_1^t) = \tau_1^t - \sigma_1^t$. Now observe that taking $\psi(r) = e^{-\alpha r}$, $\gamma(s) = e^{-\alpha s}$ and $f(x) \equiv 1$ and letting $g(z) = e^{-\alpha z}$, (11) implies

$$g(\overline{S}(t) + \overline{R}(t)) - \int_0^t \left[g'(\overline{S}(r) + \overline{R}(r)) + (1 - \overline{u}(\overline{R}(r), \overline{S}(r), \overline{X}(r)))\left\{g(0) - g(\overline{S}(r) + \overline{R}(r))\right\}\right] dr \quad (12)$$

is a martingale, where $\overline{u}(\overline{R}(r), \overline{S}(r), \overline{X}(r)) = \int u\, \overline{\eta}(\overline{R}(r), \overline{S}(r), \overline{X}(r), du)$. In particular, notice that since $f \equiv 1$, $Bf = 0$ eliminates the singular term in (11). Now approximating more general g by linear combinations of the exponentials $e^{-\alpha z}$, we see that (12) holds for C^1 functions with g and g' bounded. Now letting $\tilde{\sigma}_2^t = \inf\{r > \tau_1^t : \overline{S}(r) + \overline{R}(r) = 0\}$, Lemma 4.4 of [13] implies

$$\begin{aligned} &P\left(\int_{\tau_1^t}^{\sigma_2^t} (1 - \overline{u}(\overline{R}(r), \overline{S}(r), \overline{X}(r)))\, dr > x | \mathcal{F}_{\tau_1^t}^{\overline{R}, \overline{S}, \overline{X}}\right) \\ &= \quad e^{-x} = P\left(\int_{\tau_1^t}^{\tilde{\sigma}_2^t} (1 - \overline{u}(\overline{R}(r), \overline{S}(r), \overline{X}(r)))\, dr > x | \mathcal{F}_{\tau_1^t}^{\overline{R}, \overline{S}, \overline{X}}\right), \end{aligned}$$

and since $\sigma_2^t \leq \tilde{\sigma}_2^t$, we must have $\sigma_2^t = \tilde{\sigma}_2^t$ a.s.; in particular, $\overline{S}(\sigma_s^t) = 0$ a.s.. Finally, defining $Z(r) = (\overline{R}(\tau_1^t + r), \overline{S}(\tau_1^t + r), \overline{X}(\tau_1^t + r))$ for $r < \sigma_2^t - \tau_1^t$, we can extend Z to be a solution of the martingale problem for the generator $\tilde{C}$ defined by

$$\tilde{C}g(r, s, x) = \int g(0, 0, x)\, \nu_0(dy) - g(r, s, x) + \frac{\partial}{\partial r} g(r, s, x).$$

Since any solution of this martingale problem has the property that the final component is constant except for jumps that occur when the first two components jump to 0, it follows that $\overline{X}(r) = \overline{X}(\tau_1^t)$ for $\tau_1^t \leq r < \sigma_2^t$.

Now let h be a fixed, bounded, continuous function and, for $\epsilon > 0$, define

$$H_\epsilon(r) = \int_U e^{-\epsilon(\overline{R}(r) + \overline{S}(r))} h(\overline{R}(r), \overline{S}(r), \overline{X}(r))\, \overline{\eta}(\overline{R}(r), \overline{S}(r), \overline{X}(r), du).$$

As a process in t,

$$(\sigma_1^t - \sigma_0^t)^{-1} \int_{\sigma_1^t}^{\sigma_2^t} H_\epsilon(r)\, dr$$

is stationary, and for each $s \in [\sigma_0^t, \sigma_1^t)$

$$(\sigma_1^s - \sigma_0^s)^{-1} \int_{\sigma_1^s}^{\sigma_2^s} H_\epsilon(r)\, dr = (\sigma_1^t - \sigma_0^t)^{-1} \int_{\sigma_1^t}^{\sigma_2^t} H_\epsilon(r)\, dr.$$

Using stationarity, we have

$$E\left[(\sigma_1^t-\sigma_0^t)^{-1}\int_{\sigma_1^t}^{\sigma_2^t} H_\epsilon(r)\,dr\right]=T^{-1}\int_0^T E\left[(\sigma_1^t-\sigma_0^t)^{-1}\int_{\sigma_1^t}^{\sigma_2^t} H_\epsilon(r)\,dr\right]\,dt. \tag{13}$$

These expressions may be infinite but, in fact, the following argument shows that both terms are finite and identifies their common value.

Let $N(T)$ denote the number of jumps of the process $(\overline{R},\overline{S},\overline{X})$ in the interval $[0,T]$, and let $\{\sigma_k : k=1,\ldots,N(T)\}$ denote these jump times. Let $\sigma_{N(T)+1}$ denote the first jump time after time T and σ_0 denote the last jump before time $t=0$. Then the right-hand side of (13) equals

$$\begin{aligned}
&T^{-1}E\left[\sum_{k=1}^{N(T)+1}\frac{T\wedge\sigma_k-\sigma_{k-1}\vee 0}{\sigma_k-\sigma_{k-1}}\int_{\sigma_k}^{\sigma_{k+1}}H_\epsilon(r)\,dr\right]\\
&= T^{-1}E\left[\int_0^T H_\epsilon(r)\,dr\right]\\
&\qquad -T^{-1}E\left[\left(1-\frac{T\wedge\sigma_1}{\sigma_1-\sigma_0}\right)\int_0^{\sigma_1\wedge T}H_\epsilon(r)\,dr\right]\\
&\qquad +T^{-1}E\left[I_{\{N(T)=1\}}\left(\frac{\sigma_1}{\sigma_1-\sigma_0}\right)\int_T^{\sigma_2}H_\epsilon(r)\,dr\right]\\
&\qquad +T^{-1}E\left[I_{\{N(T)>1\}}\int_T^{\sigma_{N(T)+1}}H_\epsilon(r)\,dr\right]\\
&\qquad +T^{-1}E\left[I_{\{N(T)>0\}}\left(\frac{T-\sigma_{N(T)}}{\sigma_{N(T)+1}-\sigma_{N(T)}}\right)\int_{\sigma_{N(T)+1}}^{\sigma_{N(T)+2}}H_\epsilon(r)\,dr\right].
\end{aligned}$$

Since $\overline{\mu}_0$ is the stationary distribution of $(\overline{R}(t),\overline{S}(t),\overline{X}(t))$, the first term is

$$\int e^{-\epsilon(r+s)}h(r,s,x,u)\overline{\mu}_0(dr\times ds\times dx\times du)$$

and the other terms are bounded above by $4||h||/(\epsilon T)$. Letting $T\to\infty$ implies that

$$\begin{aligned}
&E\left[(\sigma_1^t-\sigma_0^t)^{-1}\int_{\sigma_1^t}^{\sigma_2^t}\int_U e^{-\epsilon(\overline{R}(r)+\overline{S}(r))}h(\overline{R}(r),\overline{S}(r),\overline{X}(r),u)\overline{\eta}(\overline{R}(r),\overline{S}(r),\overline{X}(r),du)\,dr\right]\\
&\qquad\qquad = \int e^{-\epsilon(r+s)}h(r,s,x,u)\overline{\mu}_0(dr\times ds\times dx\times du)
\end{aligned}$$

and hence letting $\epsilon\to 0$,

$$\begin{aligned}
&E\left[(\sigma_1^t-\sigma_0^t)^{-1}\int_{\sigma_1^t}^{\sigma_2^t}\int_U h(\overline{R}(r),\overline{S}(r),\overline{X}(r),u)\overline{\eta}(\overline{R}(r),\overline{S}(r),\overline{X}(r),du)\,dr\right]\\
&\qquad\qquad = \int h(r,s,x,u)\overline{\mu}_0(dr\times ds\times dx\times du)
\end{aligned} \tag{14}$$

for each bounded, continuous h and therefore for each bounded, measurable h.

Now setting $h(r,s,x,u) = I_{\{0\}}(u)$ in (14) yields

$$K^{-1} = E\left[(\sigma_1^t - \sigma_0^t)^{-1}(\sigma_2^t - \tau_1^t)\right] = E\left[(\sigma_1^t - \sigma_0^t)^{-1}\right] \tag{15}$$

in which the last equality follows from the facts that σ_1^t and σ_0^t are $\mathcal{F}_{\tau_1^t}^{\overline{R},\overline{S},\overline{X}}$-measurable and, conditional on $\mathcal{F}_{\tau_1^t}^{\overline{R},\overline{S},\overline{X}}$, $\sigma_2^t - \tau_1^t$ is a mean 1 exponential random variable.

We are now in a position to define the desired process that is a solution of the martingale problem and the corresponding stopping time. Take $t = 0$ in the above and define the process (R,S,X) by $X(r) = \overline{X}(\sigma_1^0 + r)$, $R(r) = \overline{R}(\sigma_1^0 + r)$, $S(r) = \overline{S}(\sigma_1^0 + r)$, for $r \geq 0$, and set the filtration $\{\mathcal{F}_t\}$ to be $\mathcal{F}_t = \mathcal{F}_{\sigma_0+t}^{\overline{R},\overline{S},\overline{X}}$ for each t. Let $\tau = \inf\{r \geq 0 : R(r) > 0\}$ and $\sigma = \inf\{r > \tau : R(r) = 0\}$. Note that $X(r) = X(\tau)$ for $\tau \leq r < \sigma$. Since σ_1^0 and σ_0^0 are both $\mathcal{F}_0$-measurable, (15) allows us to define a new probability measure $\tilde{P}$ having Radon-Nikodym derivative $\frac{d\tilde{P}}{dP} = K(\sigma_1^0 - \sigma_0^0)^{-1}$. It then follows from (9), (14) and (15) that for each bounded, measurable h,

$$\begin{aligned}
E^{\tilde{P}}&\left[\int_0^\sigma \int_U h(R(r),S(r),X(r),u)\,\overline{\eta}(R(r),S(r),X(r),du)\,dr\right] \\
&= \int h(r,s,x,u)\,\overline{\mu}_0(dr\times ds\times dx\times du)\,/\,E[(\sigma_1^0 - \sigma_0^0)^{-1}] \\
&= \int h(0,s,x,1)\,\mu_0(ds\times dx) \\
&\qquad + \int_0^\infty \int e^{-r} h(r,s,x,0)\,\mu_2(ds\times dx)\,dr.
\end{aligned} \tag{16}$$

Now taking $h(r,s,x,u) = I_{\{0\}}(r)I_{\{0\}}(u)$ in (16), shows that $\overline{\eta}(R(r),S(r),X(r),\cdot)$ places unit mass at $\{1\}$ a.s. for $0 \leq r < \tau$, and similarly, $h(r,s,x,u) = I_{(0,\infty)}(r)I_{\{1\}}(u)$ establishes that $\overline{\eta}(R(r),S(r),X(r),\cdot)$ places unit mass at $\{0\}$ a.s. for $\tau \leq r < \sigma$. Furthermore, for $G \in \mathcal{B}(\mathbb{R}^+ \times E)$ and $h(r,s,x,u) = I_{(0,\infty)\times G\times\{0\}}(r,s,x,u)$, (16) implies

$$\mu_2(G) = E^{\tilde{P}}\left[\int_\tau^\sigma I_G(S(r),X(r))\,dr\right] = E^{\tilde{P}}[I_G(\tau,X(\tau))(\sigma-\tau)] = E^{\tilde{P}}[I_G(\tau,X(\tau))].$$

Similarly, for $G \in \mathcal{B}(\mathbb{R}^+ \times E)$ and $h(r,s,x,u) = I_{\{0\}\times G\times\{1\}}(r,s,x,u)$, (16) implies

$$\mu_0(G) = E^{\tilde{P}}\left[\int_0^\tau I_G(r,X(r))\,dr\right].$$

It is necessary to determine the corresponding results related to the random measure $\overline{\Gamma}$. We adapt the argument that established (14). For bounded, continuous h on $\mathbb{R}^+ \times \mathbb{R}^+ \times E$,

$$(\sigma_1^t - \sigma_0^t)^{-1} \int_{\mathbb{R}^+\times\mathbb{R}^+\times E\times[\sigma_1^t,\sigma_2^t)} h(r,s,x)\,\overline{\Gamma}(dr\times ds\times dx\times dv)$$

is stationary in t. As a result,

$$
\begin{aligned}
E&\left[(\sigma_1^t-\sigma_0^t)^{-1}\int_{\mathbb{R}^+\times\mathbb{R}^+\times E\times[\sigma_1^t,\sigma_2^t)} h(r,s,x)\,\overline{\Gamma}(dr\times ds\times dx\times dv)\right]\\
&= T^{-1}\int_0^T E\left[(\sigma_1^t-\sigma_0^t)^{-1}\int_{\mathbb{R}^+\times\mathbb{R}^+\times E\times[\sigma_1^t,\sigma_2^t)} h(r,s,x)\,\overline{\Gamma}(dr\times ds\times dx\times dv)\right]\,dt\\
&= T^{-1}E\left[\sum_{k=1}^{N(T)+1}\frac{T\wedge\sigma_k-\sigma_{k-1}\vee 0}{\sigma_k-\sigma_{k-1}}\int_{\mathbb{R}^+\times\mathbb{R}^+\times E\times[\sigma_k,\sigma_{k+1})} h(r,s,x)\,\overline{\Gamma}(dr\times ds\times dx\times dv)\right]\\
&= T^{-1}E\left[\int_{\mathbb{R}^+\times\mathbb{R}^+\times E\times[0,T)} h(r,s,x)\,\overline{\Gamma}(dr\times ds\times dx\times dv)\right]\\
&\qquad -T^{-1}E\left[\left(1-\frac{T\wedge\sigma_1}{\sigma_1-\sigma_0}\right)\int_{\mathbb{R}^+\times\mathbb{R}^+\times E\times[\sigma_1\wedge T)} h(r,s,x)\,\overline{\Gamma}(dr\times ds\times dx\times dv)\right]\\
&\qquad +T^{-1}E\left[I_{\{N(T)=1\}}\left(\frac{\sigma_1}{\sigma_1-\sigma_0}\right)\int_{\mathbb{R}^+\times\mathbb{R}^+\times E\times[T,\sigma_2)} h(r,s,x)\,\overline{\Gamma}(dr\times ds\times dx\times dv)\right]\\
&\qquad +T^{-1}E\left[I_{\{N(T)>1\}}\int_{\mathbb{R}^+\times\mathbb{R}^+\times E\times[T,\sigma_{N(T)+1})} h(r,s,x)\,\overline{\Gamma}(dr\times ds\times dx\times dv)\right]\\
&\qquad +T^{-1}E\left[I_{\{N(T)>0\}}\left(\frac{T-\sigma_{N(T)}}{\sigma_{N(T)+1}-\sigma_{N(T)}}\right)\right.\\
&\qquad\qquad\left.\cdot\int_{\mathbb{R}^+\times\mathbb{R}^+\times E\times[\sigma_{N(T)+1},\sigma_{N(T)+2})} h(r,s,x)\,\overline{\Gamma}(dr\times ds\times dx\times dv)\right].
\end{aligned}
$$

Again, the first term equals $\int h\,d\overline{\mu}_1$ and the other terms converge to 0 as $T\to\infty$ and therefore

$$
\begin{aligned}
E&\left[(\sigma_1^t-\sigma_0^t)^{-1}\int_{\mathbb{R}^+\times\mathbb{R}^+\times E\times[\sigma_1^t,\sigma_2^t)} h(r,s,x)\,\overline{\Gamma}(dr\times ds\times dx\times dv)\right]\\
&= \int h(r,s,x)\,\overline{\mu}_1(dr\times ds\times dx)\\
&= K^{-1}\int h(0,s,x)\,\mu_1(ds\times dx) \qquad (17)
\end{aligned}
$$

where the last equality follows from (10).

Define the $\mathcal{L}(\mathbb{R}^+\times\mathbb{R}^+\times E)$-valued random measure Γ such that for each $s,t\geq 0$, $\Gamma(\cdot\times\cdot\times\cdot\times[s,t))=\overline{\Gamma}(\cdot\times\cdot\times\cdot\times[\sigma_1^0+s,\sigma_1^0+t))$. Noting the fact that $R(r)>0$ for $r\in[\tau,\sigma)$, (17) implies first that $\Gamma(\cdot\times[\tau,\sigma))\equiv 0$ a.s. and secondly that

$$
\mu_1(G)=E^{\tilde{P}}\left[\int_{\mathbb{R}^+\times E\times[0,\tau)} I_G(s,x)\,\Gamma(ds\times dx\times dv)\right].
$$

Finally, taking $\psi \equiv 1$ in (11) and recalling that $\overline{\eta}(R(r), S(r), X(r), \cdot) = \delta_{\{0\}}(\cdot)$ for $r \in [0, \tau)$, the optional sampling theorem implies that under $\tilde{P}$

$$\gamma(t \wedge \tau) f(X(t \wedge \tau)) - \int_0^{t\wedge\tau} \tilde{A}[\gamma f](s, X(s))\, ds - \int_{E\times[0,t\wedge\tau)} \tilde{B}[\gamma f](s,x)\Gamma(ds \times dx \times dv)$$

is a martingale with respect to the filtration $\{\mathcal{F}_t\}$. □

The previous analysis focuses on the dynamics of the process and demonstrates the fact that the identity (7) characterizes the solutions (X, Γ) of the martingale problem up to some stopping time τ. We now use this characterization to rewrite the optimal stopping problem.

Let $c_0, c_1, c_2 \in M(\mathbb{R}^+ \times E)$ be measurable and bounded below, and represent the time-dependent running cost of the process according to regular time, the running cost according to the singular time and the stopping cost, respectively. The goal of the decision maker is to select a stopping rule τ so as to minimize the expect cost of the process up to time τ:

$$E\left[\int_0^\tau c_0(s, X(s))\, ds + \int_{\mathbb{R}^+\times E\times[0,\tau)} c_1(s,x)\,\Gamma(ds \times dx \times dv) + c_2(\tau, X(\tau))\right]. \tag{18}$$

The discussion prior to Theorem 2.4 shows that for each solution (X, Γ) and stopping time τ, the occupation measures μ_0, μ_1 and the distribution ν_τ of $(\tau, X(\tau))$ satisfy (7) and moreover, the expected cost (18) is given by (8). Conversely, for each set of measures μ_0, μ_1 and μ_2 satisfying (7), Theorem 2.4 shows the existence of a solution (X, Γ) and a stopping time τ whose associated expected cost up to time τ is given by (8). The expected costs can therefore be determined by the following infinite-dimensional linear program.

Theorem 2.5 *The optimal stopping problem of selecting a stopping time τ^* so as to minimize*

$$E\left[\int_0^\tau c_0(X(s))\, ds + \int_{E\times[0,\tau]} c_1(x)\,\Gamma(dx \times ds) + c_2(X(\tau))\right]$$

subject to (X, Γ) being a solution of the singular martingale problem for (A, B, ν_0) is equivalent to the infinite dimensional linear program

$$\begin{array}{lll} Min. & \int c_0\, d\mu_0 + \int c_1\, d\mu_1 + \int c_2\, d\nu_\tau & \\ S.t. & 0 = \int \gamma f\, d\nu_\tau - \gamma(0)\int f\, d\nu_0 - \int \tilde{A}[\gamma f]\, d\mu_0 - \int \tilde{B}[\gamma f]\, d\mu_1, & \forall \gamma f \in \mathcal{D}_1 \\ & \nu_\tau \in \mathcal{P}(\mathbb{R}^+ \times E), & \\ & \mu_0, \mu_1 \in \mathcal{M}(\mathbb{R}^+ \times E). & \end{array} \tag{19}$$

3 Drifted Brownian Motion

Let X be a drifted Brownian motion with drift rate $-a$ starting from the point x_0 and let Y be its running maximum starting with an initial value of y_0 for some value of $y_0 \geq x_0$. Then the pair (X,Y) satisfies

$$\begin{aligned} X(t) &= x_0 - at + \sigma W(t), \\ Y(t) &= y_0 \vee \max_{\{0\le s\le t\}} X(s). \end{aligned} \tag{20}$$

Observe the paired process (X,Y) evolves over the region $\mathcal{H} = \{(x,y) : x \le y, y \ge y_0\}$. Let $\{\mathcal{F}_t\}$ denote the filtration generated by X so $\mathcal{F}_t = \sigma(X(s) : 0 \le s \le t)$.

The goal is to select an $\{\mathcal{F}_t\}$-stopping time τ so as to maximize (2).

3.1 The Exact Solution

In order to compare the numerical results from our linear programming formulation, we derive the value function and identify both the continuation and stopping regions. It is instructive to solve the two-dimensional problem in order to observe how the singular operator is handled in the variational argument. We note that the boundary between the continuation and stopping regions is the line $y = x - b^*$ with b^* being the solution to a transcendental equation. Thus, the exact solution must also be numerically computed.

We solve this problem by solving a variational inequality which includes the singular operator. To do so, we first observe that the absolutely continuous generator of the two dimensional process (x,y) is $A : \mathcal{D} \to C(\mathcal{H})$ defined by

$$Af(x,y) = \frac{1}{2}\,\sigma^2 f_{xx}(x,y) - a f_x(x,y) \tag{21}$$

and the singular generator $B : \mathcal{D} \to C(\mathcal{H})$ is

$$Bf(x,y) = f_y(x,y) \tag{22}$$

with $\mathcal{D}$ restricted to the test functions

$$\mathcal{D} = \{f \in C^{2,0}(\mathcal{H}) : f_y(x,y) \text{ exists on } y = x\}. \tag{23}$$

The notation $C^{(i,j)}(\mathcal{H})$ denotes the space of continuous functions on $\mathcal{H}$ having i continuous derivatives in the first component and j continuous derivatives in the second component. Note, in particular, that this singular formulation of the problem only requires $f \in \mathcal{D}$ to have f_y exist on the diagonal $\{(x,y) : x = y\}$. This contrasts with the typical absolutely continuous formulation in which reflection is obtained by requiring $f_y(y,y) = 0$.

We seek regions $\mathcal{C}$ and $\mathcal{S}$ and a function V satisfying the following conditions:

(i) $\mathcal{C} \cup \mathcal{S} = \mathcal{H}$, $\mathcal{B} = \overline{\mathcal{C}} \cap \mathcal{S}$,

(ii) $V \in \mathcal{D}_{\mathcal{C}} \cap C^{(1,0)}(\mathcal{H})$ where $\mathcal{D}_{\mathcal{C}}$ is defined as in (23) but with $\mathcal{H}$ replaced by $\mathcal{C}$,

(iii) for $(x,y) \in \mathcal{S}$, $V(x,y) = y - x$, and

(iv) for $(x,y) \in \mathcal{C}$, $V(x,y) \ge y - x$,

(v) for $(x,y) \in \mathcal{C}$, $AV(x,y) - \lambda V(x,y) = 0$,

(vi) for $(x,y) \in \mathcal{S}\backslash\mathcal{B}$, $AV(x,y) - \lambda V(x,y) \leq 0$,

(vii) for (x,y) on the line $y = x$, $V_y(y,y)$ exists and $V_y(y,y) = 0$,

(viii) $\tau_{\mathcal{S}} := \inf\{t \geq 0 : (x_t, y_t) \in \mathcal{S}\} < \infty$ a.s. and the stochastic integral $\int_0^{t\wedge\tau_{\mathcal{S}}} V_x(x_s, y_s)\, dw_s$ is a martingale.

The region $\mathcal{C}$ denotes the continuation region and $\mathcal{S}$ denotes the stopping region. The requirements (ii), (iii) and (iv) on the function V are known as "smooth pasting" since V is defined differently on the regions $\mathcal{C}$ and $\mathcal{S}$ and is required to have continuous first partial derivative in x across the boundary $\mathcal{B}$.

We remark that there are two ways one may view condition (vii). It is the natural requirement on the singular generator applied to V in the continuation region $\mathcal{C}$, similar to condition (v). Since the reflection occurs on the diagonal where $y = x$ and the evaluation of the objective function at these points is 0, these points cannot be in the stopping region. It may also be viewed as requiring V to be in the domain of the generator for the drifted Brownian motion *reflected* at the diagonal, which reduces the associated martingale to an absolutely continuous martingale.

Let $\alpha_1 = (-a - \sqrt{a^2 + 2\lambda\sigma^2})/\sigma^2$ and $\alpha_2 = (-a + \sqrt{a^2 + 2\lambda\sigma^2})/\sigma^2$ and observe that $\alpha_1 < 0 < \alpha_2$ with $\alpha_2 - \alpha_1 = 2\sqrt{a^2 + 2\lambda\sigma^2}/\sigma^2$ and $\alpha_1 \cdot \alpha_1 = -2\lambda/\sigma^2$. To find V, we begin by observing that the general solution to the pde in (v) is given by

$$f(x,y) = c_1(y)e^{\alpha_1 x} + c_2(y)e^{\alpha_2 x}. \tag{24}$$

We conjecture that $\mathcal{C} = \{(x,y) : 0 \leq y - x < b, y \geq y_0\}$ and $\mathcal{S} = \{(x,y) : y - x \geq b, y \geq y_0\}$ so that the boundary is given by a line of the form $y = x + b$ for some value of b. The principle of smooth fit requires

$$\begin{aligned} b &= (e^{\alpha_1 x})\, c_1(x+b) + e^{\alpha_2 x} c_2(x+b) \\ -1 &= (\alpha_1 e^{\alpha_1 x})\, c_1(x+b) + (\alpha_2 e^{\alpha_2 x})\, c_2(x+b) \end{aligned}$$

which upon solving for $c_1(x+b)$ and $c_2(x+b)$ and making the substitution $y = x + b$, we obtain

$$\begin{aligned} c_1(y) &= \frac{\alpha_2 b + 1}{\alpha_2 - \alpha_1}\, e^{-\alpha_1(y-b)} \\ c_2(y) &= -\frac{\alpha_1 b + 1}{\alpha_2 - \alpha_1}\, e^{-\alpha_2(y-b)}\ . \end{aligned}$$

The function V thus has the form

$$V(x,y) = \frac{\alpha_2 b + 1}{\alpha_2 - \alpha_1}\, e^{\alpha_1(x+b-y)} - \frac{\alpha_1 b + 1}{\alpha_2 - \alpha_1}\, e^{\alpha_2(x+b-y)}$$

with the value of b still to be determined using condition (vii), namely $V_y(y,y) = 0$. Since

$$V_y(x,y) = -\frac{\alpha_1(\alpha_2 b + 1)}{\alpha_2 - \alpha_1}\, e^{\alpha_1(x+b-y)} + \frac{\alpha_2(\alpha_1 b + 1)}{\alpha_2 - \alpha_1}\, e^{\alpha_2(x+b-y)},$$

setting $x = y$ yields the requirement that b satisfy

$$e^{(\alpha_2-\alpha_1)b} = \frac{\alpha_1(\alpha_2 b+1)}{\alpha_2(\alpha_1 b+1)} = \frac{b+\frac{1}{\alpha_2}}{b+\frac{1}{\alpha_1}}. \tag{25}$$

Analyzing the graphs of the left hand side and the right hand side of (25) indicates that there are two solutions with one being positive and the other negative. The positive solution b^* is required in order for the process to be able to reach $\mathcal{B}$ and hence for condition (viii) to be satisfied. In addition, the drift rate $-a$ must be nonpositive for the stopping time $\tau_{\mathcal{S}}$ to be finite a.s., again as required by condition (viii), implying $a \geq 0$.

Finally, elementary calculations again show that V satisfies condition (iv) and condition (vi) is satisfied since on $\mathcal{S}\backslash\mathcal{B}$, $y - x > b > 0$ and

$$0 \geq AV - \lambda V = -a - \lambda(y - x).$$

Summarizing, the function V which satisfies the smooth pasting conditions is

$$V(x,y) = \begin{cases} \dfrac{\alpha_2 b^*+1}{\alpha_2-\alpha_1}\, e^{-\alpha_1(y-x-b^*)} - \dfrac{\alpha_1 b^*+1}{\alpha_2-\alpha_1}\, e^{-\alpha_2(y-x-b^*)} & \text{for } (x,y) \in \mathcal{C}, \\ \\ y-x & \text{for } (x,y) \in \mathcal{S}, \end{cases} \tag{26}$$

where b^* is the positive solution of (25), $\mathcal{C} = \{(x,y) : 0 \leq y - x < b^* \leq y, y \geq y_0\}$ is the continuation region and $\mathcal{S} = \{(x,y) : b^* \leq y - x, y \geq y_0\}$ is the stopping region.

Theorem 3.1 *The function V given in (26) with b^* specified as the positive solution of (25) is the value function for the optimal stopping problem specified by the dynamics (20) and the objective (2). Moreover, $\tau_{\mathcal{S}} = \inf\{t \geq 0 : (x_t, y_t) \in \mathcal{S}\}$ is an optimal stopping rule.*

In the case when there is no drift ($a = 0$), the formula for the value function and the transcendental equation simplify. We display the simpler expressions in the following corollary.

Corollary 3.2 *The value function for the optimal stopping problem specified by (2) and (20) when $a = 0$ is*

$$V(x,y) = \begin{cases} b^* \cosh\left(\frac{\sqrt{2\lambda}}{\sigma}(y-x-b^*)\right) + \frac{\sigma}{\sqrt{2\lambda}} \sinh\left(\frac{\sqrt{2\lambda}}{\sigma}(y-x-b^*)\right), & (x,y) \in \mathcal{C} \\ y-x, & (x,y) \in \mathcal{S}, \end{cases}$$

where b^* is the unique positive solution to the transcendental equation

$$\frac{\sqrt{2\lambda}}{\sigma} b = \coth\left(\frac{\sqrt{2\lambda}}{\sigma} b\right),$$

the continuation region is $\mathcal{C} = \{(x,y) : 0 \leq y - x < b^* \leq y, y \geq y_0\}$ and the stopping region is $\mathcal{S} = \{(x,y) : b^* \leq y - x, y \geq y_0\}$.

An optimal stopping rule is to stop at $\tau_{\mathcal{S}} = \inf\{t \geq 0 : (x_t, y_t) \in \mathcal{S}\}$.

Proof. Observe that on the set $\{\tau_{\mathcal{S}} > t\}$,

$$\begin{aligned} |V(X(t),Y(t))| &= \left| \frac{\alpha_2 b^* + 1}{\alpha_2 - \alpha_1}\, e^{-\alpha_1(Y(t)-X(t)-b^*)} - \frac{\alpha_1 b^* + 1}{\alpha_2 - \alpha_1}\, e^{-\alpha_2(Y(t)-X(t)-b^*)} \right| \\ &\le \frac{\alpha_2 b^* + 1}{\alpha_2 - \alpha_1} + \frac{|\alpha_1 b^* + 1|}{\alpha_2 - \alpha_1}\, e^{\alpha_2 b^*} \end{aligned}$$

since $\tau_{\mathcal{S}} > t$ implies $0 \le Y(t) - X(t) < b^*$. A similar analysis shows that on the set $\{\tau_{\mathcal{S}} > t\}$, $|V_x(X(t),Y(t))| \le \frac{|\alpha_1|(\alpha_2 b^*+1)}{\alpha_2-\alpha_1} + \frac{\alpha_2|\alpha_1 b^*+1|}{\alpha_2-\alpha_1}\, e^{\alpha_2 b^*}$ and hence $\int_0^{t\wedge\tau_{\mathcal{S}}} V_x(X(s),Y(s))\, dW(s)$ is a martingale. Using condition (ii), Itô's formula implies that

$$e^{-\lambda(t\wedge\tau_{\mathcal{S}})} V(X(t\wedge\tau_{\mathcal{S}}), Y(t\wedge\tau_{\mathcal{S}})) - \int_0^{t\wedge\tau_{\mathcal{S}}} e^{-\lambda s}(AV(X(s),Y(s)) - \lambda V(X(s),Y(s)))\, ds$$

is an $\{\mathcal{F}_t\}$-martingale. Thus

$$\begin{aligned} V(x_0,y_0) &= E\Big[e^{-\lambda(t\wedge\tau_{\mathcal{S}})} V(X(t\wedge\tau_{\mathcal{S}}), Y(t\wedge\tau_{\mathcal{S}})) \\ &\qquad - \int_0^{t\wedge\tau_{\mathcal{S}}} e^{-\lambda s}(AV(X(s),Y(s)) - \lambda V(X(s),Y(s)))\, ds\Big] \\ &= E\left[e^{-\lambda(t\wedge\tau_{\mathcal{S}})} V(X(t\wedge\tau_{\mathcal{S}}), Y(t\wedge\tau_{\mathcal{S}}))\right] \end{aligned}$$

since $AV - \lambda V = 0$ on $\mathcal{C}$. Now observe that

$$\begin{aligned} E\left[e^{-\lambda(t\wedge\tau_{\mathcal{S}})} V(X(t\wedge\tau_{\mathcal{S}}), Y(t\wedge\tau_{\mathcal{S}}))\right] &= E\left[e^{-\lambda\tau_{\mathcal{S}}} V(X(\tau_{\mathcal{S}}), Y(\tau_{\mathcal{S}})) I_{\{\tau_{\mathcal{S}} \le t\}}\right] \\ &\quad + E\left[e^{-\lambda t} V(X(t),Y(t)) I_{\{\tau_{\mathcal{S}} > t\}}\right]. \end{aligned} \tag{27}$$

Hence recalling that $V(X(t),Y(t))$ is bounded on the set $\{\tau_{\mathcal{S}} > t\}$, the second term in (27) converges to 0 as $t \to \infty$. Using the bounded convergence theorem on the first term, we have

$$V(x_0,y_0) = E\left[e^{-\lambda\tau_{\mathcal{S}}}(Y(\tau_{\mathcal{S}}) - X(\tau_{\mathcal{S}}))\right].$$

It still remains to show that $\tau_{\mathcal{S}}$ is optimal. To do so, we employ a mollification argument.

Let τ be any stopping time. We first observe that the law of the iterated logarithm (see [12, Theorem 2.9.23]) implies that on the set $\{\tau = \infty\}$, $e^{-\lambda t}$ dominates $(Y(t) - X(t))$ and hence $\lim_{t\to\infty} E[e^{-\lambda t}(Y(t) - X(t)) I_{\{\tau=\infty\}}] = 0$. Thus we may, without loss of generalization, assume that $\tau < \infty$ a.s.. In fact, the objective improves by setting τ finite.

Next observe that $E\left[\int_0^\infty I_{\mathcal{B}}(X(t),Y(t))\, dt\right] = \int_0^\infty P((X(t),Y(t)) \in \mathcal{B})\, dt = 0$ so the set of times at which the process (X,Y) is on $\mathcal{B}$ has Lebesgue measure 0.

By Theorem D.1 of Øksendal [16], there exists a sequence $\{V_j\}_{j=1}^\infty$ of functions $V_j \in C^2(\mathcal{H})$ such that

- $V_j \to V$ uniformly on compact subsets of $\mathcal{H}$ as $j \to \infty$,
- $AV_j \to AV$ uniformly on compact subsets of $\mathcal{H}\backslash\mathcal{B}$, as $j \to \infty$, and
- $\{AV_j\}_{j=1}^\infty$ is locally bounded on $\mathcal{H}$.

(For example, let $\phi(x) = e^{-1/(1-x^2)}$ and define $V_j(x,y) = \int \epsilon_j^{-1} V(x,y)\phi(x/\epsilon_j)dx$, where $\epsilon_j \searrow 0$ as $j \to \infty$.) Itô's formula then implies that

$$\begin{aligned} V_j(x_0,y_0) &= E[e^{-\lambda(t\wedge\tau)} V_j(X(t\wedge\tau), Y(t\wedge\tau))] \\ &\qquad - E\left[\int_0^{t\wedge\tau} e^{-\lambda s}[AV_j - \lambda V_j](X(s),Y(s))\,ds\right] \\ &\qquad - E\left[\int_0^{t\wedge\tau} e^{-\lambda s} BV_j(X(s),Y(s))\,dY(s)\right]. \end{aligned}$$

Letting $j \to \infty$ and using the uniform convergence along with the fact that the set of times (X,Y) is on $\mathcal{B}$ has measure 0, Fatou's Lemma implies

$$\begin{aligned} V(x_0,y_0) &\geq E[e^{-\lambda(t\wedge\tau)} V(X(t\wedge\tau), Y(t\wedge\tau))] \\ &\qquad - E\left[\int_0^{t\wedge\tau} e^{-\lambda s}[AV - \lambda V](X(s),Y(s))\,ds\right] \\ &\qquad - E\left[\int_0^{t\wedge\tau} e^{-\lambda s} BV(X(s),Y(s))\,dY(s)\right] \\ &\geq E\left[e^{-\lambda(t\wedge\tau)}[Y(t\wedge\tau) - X(t\wedge\tau)]\right]. \end{aligned}$$

Finally, letting $t \to \infty$, employing Fatou's Lemma again yields

$$V(x_0,y_0) \geq E\left[e^{-\lambda\tau}(Y(\tau) - X(\tau))\right]$$

establishing the result.

□

3.2 Dimension Reduction and LP Formulation

Observe that the objective of the decision maker is to maximize the expected discounted value of $Y(\tau) - X(\tau)$. Thus it is only necessary to track the process $Z(t) = Y(t) - X(t)$ to be able to know the value. The process Z which starts at $z_0 = y_0 - x_0 \geq 0$ has the same distribution as a (drifted) Brownian motion process that is reflected at $\{0\}$ (see [12]). We can therefore change the stopping problem to one of maximizing

$$E\left[e^{-\lambda\tau} Z(\tau)\right] \tag{28}$$

where Z has absolutely continuous and singular generators

$$\begin{aligned} Af(z) &= -af'(z) + (\sigma^2/2)f''(z), \text{ and} \\ Bf(z) &= f'(z). \end{aligned}$$

Theorem 2.4 is stated with μ_0, μ_1 and μ_2 being measures on $\mathbb{R}^+ \times E$ since the proof required a time component in order to identify the stopping time σ_1^0 when the solution $(\overline{X}, \overline{\Gamma})$ needs to be started to obtain the desired solution (X, Γ). The only time dependence in the specific problem under consideration, however, is through discounting. We can simplify the formulation of the LP in the following manner.

For a solution (Z, Γ) of the singular martingale problem for (A, B, ν_0), it follows from taking $\gamma(t) = e^{-\lambda t}$ that

$$e^{-\lambda t} f(Z(t)) - f(z_0) - \int_0^t e^{-\lambda s}(Af - \lambda f)(Z(s))\, ds - \int_{\mathbb{R}^+ \times \mathbb{R}^+ \times [0,t]} e^{-\lambda s} Bf(z)\, \Gamma(ds \times dz \times dr)$$

is a martingale. Now define the "discounted" measures on Borel sets G by

$$\begin{aligned}
\mu_0(G) &= E\left[\int_0^\tau e^{-\lambda s} I_G(Z(s))\, ds\right], \\
\mu_1(G) &= E\left[\int_{\mathbb{R}+ \times \mathbb{R}^+ \times [0,\tau]} e^{-\lambda s} I_G(z)\, \Gamma(ds \times dz \times dr)\right], \\
\nu_\tau(G) &= E\left[e^{-\lambda \tau} I_G(Z(\tau))\right].
\end{aligned}$$

It then follows from the martingale property and the optional sampling theorem that μ_0, μ_1 and ν_τ satisfy the conditions of Theorem 2.4. Note in particular that the discounting is incorporated into the measures, which also implies that ν_τ is a sub-probability measure.

Applying Theorem 2.5, this optimal stopping problem is equivalent to the infinite-dimensional linear program

$$Max. \quad \langle z, \nu_\tau \rangle \tag{29}$$

$$\begin{aligned}
S.t. \quad & f(z_0) = \langle f, \nu_\tau \rangle - \langle Af - \lambda f, \mu_0 \rangle - \langle Bf, \mu_1 \rangle, \qquad \forall f \in \mathcal{D} \\
& 0 \le \langle 1, \nu_\tau \rangle \le 1, \\
& \langle 1, \mu_0 \rangle \ge 0, \\
& \langle 1, \mu_1 \rangle \ge 0,
\end{aligned} \tag{30}$$

where for a measurable function g and measure μ, the notation $\langle g, \mu \rangle = \int g\, d\mu$. Notice that the linear program determines the optimal value corresponding to the initial position $Z(0) = z_0$. Denote this optimal value by $\tilde{V}(z_0)$ and observe that, in principle, the value function is obtained from the linear program by varying the initial value z_0.

First observe that the measure μ_1 captures the expected local time of z at $\{0\}$ and thus is a point mass p_0 at $\{0\}$.

Now consider the distribution ν_τ of $Z(\tau)$. In general, ν_τ can be any distribution on $[0, \infty)$. Often, as in this example, ν_τ has bounded support and the LP formulation can be used to obtain more information about this support. Helmes [7] employs the following procedure. Assume ν_τ has its support on an interval $[0, M]$ for some $M > 0$. Next, partition the interval into three segments $[0, M_1]$, $[M_1, M_2]$ and $[M_2, M]$, break the measure ν_τ into separate measures ν_τ^1, ν_τ^2 and ν_τ^3 on each subinterval and run the LP (29)-(30) using the three measures. Typically, only one of these measures has mass indicating that ν_τ has its support on the smaller subinterval. Now break this subinterval into three further subintervals and repeat the LP. (Note that the choices of M and M_2 should be large enough that ν_τ^3 has zero mass. When ν_τ^3 has positive mass, one should begin anew with different values of M and

M_2.) In this manner, the support of ν_τ can be determined to be in a very small interval and thus is likely to be at a single value b.

For clarity of exposition, we assume the optimal stopping rule will be one of stopping at the first time the process Z attains some level b. Considering for a moment the process under this stopping rule, it is clear that Z will always be in the interval $[0, b]$ and the stopping rule is enforced when Z exits the interval $[0, b)$. Thus the optimal stopping problem can be thought of as a family of exit problems, with the optimal stopping rule determined from the choice of b^* which maximizes the value (29). A search technique can be used to determine the value of b^*.

Notice that the measure μ_1 is also a point mass p_b on $\{b\}$. The above linear program, therefore, has a measure μ_0 on $[0, b]$ and point masses p_0 and p_b as its variables.

We seek to yet again rephrase the linear program in terms of a new set of variables which more readily lead to computable (finite dimensional) linear programs. With this in mind, use the function $f(z) = z^n$, with $n \geq 0$, in (30) to obtain

$$\begin{aligned} z_0^n &= \int z^n\, \nu_\tau(dz) - \int \left(\frac{n(n-1)}{2}\, z^{n-2} - a z^{n-1} - \lambda z^n \right) \mu_0(dz) - \int n z^{n-1}\, \mu_1(dz) \\ &= b^n p_b - \frac{n(n-1)}{2}\, m_0(n-2) + a m_0(n-1) - \lambda m_0(n) - n 0^{n-1} p_0, \end{aligned} \tag{31}$$

where $m_0(n)$ denotes the nth moment of the measure μ_0 and 0^0 is defined to equal 1. The objective function and other constraints can also be expressed in terms of the moments of μ_0 and the point masses.

At this point, care must be taken to ensure that the variables $\{m_0(n) : n = 0, 1, 2, \ldots\}$ are not just any sequences but are the moments of μ_0. Necessary and sufficient conditions for a sequence to be the moments of a measure on $[0, 1]$ are derived by Hausdorff [6] (see also Feller [4] for a probabilistic proof). The minor extension to the interval $[0, b]$ (using a change of variable) arises from using linearity in the identities

$$0 \leq \int z^i (b-z)^k \mu(dz) = \sum_{j=0}^{k} \binom{k}{j} (-1)^j b^{k-j} m_0(i+j), \qquad \forall i, k \geq 0. \tag{32}$$

To be computable, the infinite collection of conditions in (31) must be reduced to a finite number of conditions. We accomplish this by truncating the number of moments to the set $M_0 = \{m_0(n) : n \leq M\}$ and dropping those constraints using moments that are not in M_0. This truncation introduces an approximation to the original stopping time problem since the set of (finite-dimensional) feasible points now includes values that do *not* arise as moments of some measures. Note carefully however that all initial terms of moments sequences are elements of the feasible set. The feasible set being larger therefore implies that the optimal value of the finite LP will provide an upper bound on the value for the original optimal stopping problem.

Now consider the geometry of the feasible set. The set is determined in part by the Hausdorff moment conditions, each of which specifies a half-space in the appropriate dimension. The intersection of these half-spaces is therefore a convex set having a finite number of extreme points. Helmes and Röhl [8] determine the precise coordinates of these points. As a

	LP-results		NEV	
λ	b^*	$\tilde{V}(0)$	b^*	$V(0,0)$
0.100	2.13753483	1.27860034	2.13753	1.2786
0.200	1.60558283	0.94111933	1.60558	0.941119
0.300	1.34854165	0.78280418	1.34854	0.782804
0.400	1.18827618	0.68562758	1.18828	0.685628
0.500	1.07568729	0.61805321	1.07569	0.618053
0.600	0.99083011	0.56749719	0.99083	0.567497
0.700	0.92382189	0.52780051	0.923822	0.527801
0.800	0.86911883	0.49553996	0.869119	0.49554
0.900	0.82333255	0.46863857	0.823333	0.468639
1.000	0.78425732	0.44575236	0.784257	0.445752

Table 1: Optimal stopping points and optimal values as functions of the discount factor λ for the case BM with drift; other parameters are $\sigma = 1$, $a = 0.2$, $x_0 = y_0 = z_0 = 0$ and $M = 40$.

result, the Hausdorff moment conditions can be enforced by replacing the inequalities (32) by convex combinations of the corner points of the Hausdorff polytope. This substitution once again changes the variables in the LP, now to the coefficients of the convex combinations of the extreme points.

Table 1 compares the results of using the linear programming formulation (29)-(30) to compute the value function $\tilde{V}(0)$ at $z_0 = 0$ with "exact values" numerically evaluated (NEV) using Mathematica. Similar levels of accuracy are obtained when the values of σ and z_0 are varied while holding the other parameters fixed.

4 Geometric Brownian Motion

The drifted Brownian motion does not provide a reasonable model for the price of an asset. Rather, a model that is often used in the mathematical finance literature is that of geometric Brownian motion, also called the Black-Scholes model since it is the model for which the celebrated Black-Scholes option pricing formula was determined.

Consider the pair of processes (X,Y) satisfying

$$\begin{aligned} X(t) &= x_0 + \int_0^t \mu X(s)\,ds + \int_0^t \sigma X(s)\,dW(s) \\ Y(t) &= y_0 \vee \max_{0\leq s\leq t} X(s) \end{aligned} \tag{33}$$

with the goal of maximizing (2). It is well-known that the geometric Brownian motion process X can be expressed as

$$X(t) = x_0 \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W(t)\right\}$$

and thus, assuming that the initial value $x_0 > 0$, the process X only takes values in $(0, \infty)$. Therefore, for this section we redefine the set $\mathcal{H}$ as

$$\mathcal{H} = \{(x, y) : 0 < x \le y, y \ge y_0\}. \tag{34}$$

The generators for (X, Y) are

$$\begin{aligned} Af(x,y) &= \frac{\sigma^2}{2} x^2 f_{xx}(x,y) + \mu x f_x(x,y) \text{ and} \\ Bf(x,y) &= f_y(x,y), \end{aligned} \tag{35}$$

with domain $\mathcal{D}$ specified in (23).

4.1 The Exact Solution

As before we seek regions $\mathcal{S}$, $\mathcal{C}$, boundary $\mathcal{B}$ and a function V satisfying conditions (i) - (viii).

Theorem 4.1 *Let $0 \le \mu < \lambda$. Define $\beta_1 = \frac{1}{2} - \frac{\mu}{\sigma^2} - \sqrt{\left(\frac{\mu}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2\lambda}{\sigma^2}}$ and $\beta_2 = \frac{1}{2} - \frac{\mu}{\sigma^2} + \sqrt{\left(\frac{\mu}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2\lambda}{\sigma^2}}$. Then the value function to the problem of finding a stopping time so as to maximize (2) with dynamics given by (33) is*

$$V(x,y) = \begin{cases} \frac{\beta_2 k^* + (1-\beta_2)}{\beta_2 - \beta_1}\, x^{\beta_1} \left(\frac{y}{k^*}\right)^{1-\beta_1} - \frac{\beta_1 k^* + (1-\beta_1)}{\beta_2 - \beta_1}\, x^{\beta_2} \left(\frac{y}{k^*}\right)^{1-\beta_2}, & (x,y) \in \mathcal{C}, \\ y - x, & (x,y) \in \mathcal{S}, \end{cases} \tag{36}$$

where k^* is the positive solution (greater than 1) of

$$\frac{\beta_2(k-1)+1}{\beta_1(k-1)+1} = \frac{(1-\beta_2)}{(1-\beta_1)}\, k^{\beta_2 - \beta_1}\,. \tag{37}$$

The continuation region is $\mathcal{C} = \{(x,y) : \frac{y}{k^*} < x \le y, y \ge y_0\}$; the stopping region is $\mathcal{S} = \{(x,y) : x \le \frac{y}{k^*}, y \ge y_0\}$, with boundary $\mathcal{B}$ being the line $y = k^* x$. An optimal stopping rule is $\tau_{\mathcal{S}} = inf\{t \ge 0 : (x_t, y_t) \in \mathcal{S}\}$.

Proof. The construction of V is obtained similarly to the case of a drifted Brownian motion in the previous section. Conditions (i), (ii), (iii), (v) and (vii) are used to construct V so are straightforward to verify. It is still necessary to verify the other conditions.

First, we remark that the model requires $k \ge 1$ to make sense. Next observe that the left hand side of (37) can be written as

$$\frac{\beta_2}{\beta_1} \frac{k + \frac{1-\beta_2}{\beta_2}}{k + \frac{1-\beta_1}{\beta_1}}$$

which indicates that its graph has a horizontal asymptote at $\frac{\beta_2}{\beta_1}$ and a vertical asymptote at $-\frac{1-\beta_1}{\beta_1} = \frac{1-\beta_1}{-\beta_1} > 1$. The left hand expression in (37) is an increasing function so increases from $-\infty$ to $\frac{\beta_2}{\beta_1}$ as k increases from $-\frac{1-\beta_1}{\beta_1}$. Since $\lambda > \mu$, the coefficient in the right hand side of (37) is negative, indicating that the right hand expression will decrease from $\frac{1-\beta_2}{1-\beta_1}$ when $k = 1$ to $-\infty$. Hence there is a unique solution $k^* > 1$ of (37).

Turning to condition (vi), on the set $\mathcal{S}\backslash\mathcal{B}$, we have

$$AV(x,y) - \lambda V(x,y) = -\mu x - \lambda(y-x) \leq -[\lambda(k^*-1)+\mu]x \leq 0,$$

which requires $\mu \geq -\lambda(k^*-1)$. This inequality is satisfied since we assume μ to be non-negative, requiring that the asset price will trend higher. (In fact, for μ slightly negative $(-\lambda(k^*-1) \leq \mu < 0)$ the result still holds.) In the case $\mu < -\lambda(k^*-1)$, we suspect the optimal stopping rule will be to stop immediately!

For the verification of (iv), first observe that $V_y(x, kx) = 1$; that is, $V \in C^{1,1}(\mathcal{H})$. For $x < y < k^*x$, two applications of the intermediate value theorem (expanding about the point (x, k^*x)) yield for some $y \leq \eta, \xi < k^*x$,

$$\begin{aligned}
V_y(x,y) &= \frac{(1-\beta_1)[\beta_2(k^*-1)+1]}{k^*(\beta_2-\beta_1)}\left(1+\left(\beta_1\eta^{\beta_1-1}\right)\left(\frac{k^*x}{y}-1\right)\right) \\
&\quad -\frac{(1-\beta_2)[\beta_1(k^*-1)+1]}{k(\beta_2-\beta_1)}\left(1+\left(\beta_2\xi^{\beta_2-1}\right)\left(\frac{k^*x}{y}-1\right)\right) \\
&= 1+\beta_1\cdot\left\{\frac{(1-\beta_1)[\beta_2(k^*-1)+1]}{k^*(\beta_2-\beta_1)}\cdot\eta^{\beta_1-1}\left(\frac{k^*x}{y}-1\right)\right\} \\
&\quad -(1-\beta_2)[\beta_1(k^*-1)+1]\left\{\frac{\beta_2}{k(\beta_2-\beta_1)}\cdot\xi^{\beta_2-1}\left(\frac{k^*x}{y}-1\right)\right\}.
\end{aligned}$$

Observe the factors inside each brace are positive. Note also that β_1, $1-\beta_2$ and $\beta_1(k^*-1)+1$ are negative (rewrite (37) to obtain the last observation). Hence $V_y(x,y) < 1$ for all $x \leq y < k^*x$, implying condition (iv).

The final condition to be verified is (viii). Let $\alpha = -\frac{\beta_1\beta_2(k^*)^{\beta_1+\beta_2}}{\lambda(\beta_2(k^*)^{\beta_2}-\beta_1(k^*)^{\beta_1})}$ and define the function

$$f(x,y) = \frac{1}{\lambda} + \frac{\alpha}{\beta_1}\left(\frac{x}{y}\right)^{\beta_1} - \frac{\alpha}{\beta_2}\left(\frac{x}{y}\right)^{\beta_2}.$$

The function f is constructed to solve

$$\begin{cases} Af - \lambda f = -1 & \text{on } \mathcal{C}, \\ f = 0 & \text{on } \mathcal{B}, \\ Bf = 0 & \text{on the diagonal } y = x. \end{cases}$$

Itô's formula in conjunction with the optional sampling theorem implies

$$\begin{aligned}
f(x_0,y_0) &= E\left[f(X(t\wedge\tau_{\mathcal{S}}), Y(t\wedge\tau_{\mathcal{S}})) - \int_0^{t\wedge\tau_{\mathcal{S}}}(Af-\lambda f)(X(s),Y(s))\,ds \right.\\
&\qquad\left. - \int_0^{t\wedge\tau_{\mathcal{S}}} Bf(X(s),Y(s))\,dy_s\right] \\
&= E[\tau_{\mathcal{S}} I_{\{\tau_{\mathcal{S}}\leq t\}}] + E\left[(f(X(t),Y(t))+t)I_{\{\tau_s>t\}}\right]. \qquad (38)
\end{aligned}$$

Since f is bounded on $\mathcal{C}$, the equation (38) implies that $P(\tau_{\mathcal{S}} > t) \to 0$ as $t \to \infty$. The verification that $\int_0^{t\wedge\tau_{\mathcal{S}}} V_x(X(s),Y(s))\,dW(s)$ is a martingale follows as in Theorem 3.1. The proof of optimality of $\tau_{\mathcal{S}}$ now follows as in the proof of Theorem 3.1. $\square$

4.2 Dimension Reduction and LP Formulation

We begin this section on the numerical solution of the optimal stopping problem by reducing the problem from two dimensions to one dimension.

Proposition 4.2 *The optimal stopping problem for the perpetual lookback put option specified by maximizing (2) over stopping times τ when the process dynamics are given by (33) is equivalent to the stopping problem of maximizing*

$$x_0 \tilde{E}\left[e^{-\tilde{\lambda}\tau}\left(\frac{Y_\mu(\tau)}{X_\mu(\tau)} - 1\right)\right]$$

where $Z := \frac{Y}{X} := \left\{\frac{Y_\mu(t)}{X_\mu(t)} : t \geq 0\right\}$ is a geometric Brownian motion process having drift $-\mu$ that is reflected at 1.

Proof. Since X is a geometric Brownian motion process it can be expressed in the form $X(t) = x_0 e^{(\mu-\sigma^2/2)t+\sigma W(t)}$. As a result,

$$\begin{aligned}
E\left[e^{-\lambda\tau}(Y(\tau) - X(\tau))\right] &= E\left[e^{-\lambda\tau}X(\tau)\left(\frac{Y(\tau)}{X(\tau)} - 1\right)\right] \\
&= E\left[e^{-\lambda\tau} \cdot x_0 e^{(\mu-\sigma^2/2)\tau+\sigma W(\tau)}\left(\frac{Y(\tau)}{X(\tau)} - 1\right)\right] \\
&= x_0 \cdot \tilde{E}\left[e^{-(\lambda-\mu)\tau}\left(\frac{Y(\tau)}{X(\tau)} - 1\right)\right]
\end{aligned}$$

where $\tilde{E}$ represents the expectation under the measure $\tilde{P}$ having Radon-Nikodym derivative $\frac{d\tilde{P}}{dP} = e^{\sigma W_\tau - (\sigma^2/2)\tau}$ on $\mathcal{F}_\tau$.

The characterization of the process $\frac{Y}{X}$ as a reflected geometric Brownian motion is due to Shepp and Shiryaev [18]. □

Note that the discount factor $\tilde{\lambda}$ for the one-dimensional problem is $\tilde{\lambda} = \lambda - \mu$ so there will only be a finite value when $\mu < \lambda$; that is, the discount rate must dominate the mean growth rate of the stock.

The stopping problem can therefore be expressed as a problem of finding a stopping time τ so as to maximize

$$\tilde{E}\left[e^{-\tilde{\lambda}\tau}(Z(\tau) - 1)\right] \tag{39}$$

where Z has generators

$$\begin{aligned}
Af(z) &= -\mu z f'(z) + (\sigma^2/2)z^2 f''(z), \\
Bf(z) &= f'(z).
\end{aligned}$$

The LP formulation of the problem is

$$Max. \qquad \langle z - 1, \nu_\tau \rangle$$

$$
\begin{aligned}
S.t. \quad & f(z_0) = \langle f, \nu_\tau \rangle - \langle Af - \tilde{\lambda} f, \mu_0 \rangle - \langle Bf, \mu_1 \rangle, \qquad \forall f \in \mathcal{D}_1 \\
& \langle 1, \nu_\tau \rangle \leq 1, \\
& \langle 1, \mu_0 \rangle < \infty, \\
& \langle 1, \mu_1 \rangle < \infty.
\end{aligned} \tag{40}
$$

As in the drifted Brownian motion example, the process Z is reflected, this time at $\{1\}$ so the measure μ_1 is a point mass p_1 at $\{1\}$. In addition, the stopping rule will take the form of a first hitting time of some level k so μ_τ will also be a point mass p_k at $\{k\}$. Thus the optimal stopping problem becomes one of finding an optimal stopping level $\{k^*\}$ and determining the value from a corresponding exit problem of the process Z from the bounded interval $[1, k^*]$, with reflection occurring at $\{1\}$. A search procedure can again be employed to determine the level k^* by iteration.

Using $\lambda - \mu = \tilde{\lambda}$ and $f(z) = z^n$, $n = 0, 1, 2, \ldots$ as the test functions, the LP takes the form

$$
\begin{aligned}
Max. \quad & (k-1)p_k \\
\\
S.t. \quad & z_0^n = k^n p_k + [n\mu - n(n-1)\sigma^2/2 + (\lambda - \mu)]m_0(n) - np_1, \qquad n = 0, 1, 2, \ldots, \\
& 0 \leq p_k \leq 1, \\
& p_1 \geq 0, \\
& m_0(n) \geq 0, \qquad n = 0, 1, 2, \ldots, \\
& \{m_0(n) : n = 0, 1, 2, \ldots\} \text{ satisfy the Hausdorff moment conditions.}
\end{aligned}
$$

The variables $\{m_0(n)\}$ are the moments of the measure μ_0 on the bounded interval $[1, k]$.

As with the drifted Brownian motion model, we truncate the number of moments to obtain a finite dimensional LP and use the representation of the pseudo-moments as convex combinations of the extreme points in the resulting Hausdorff polytope. Table 2 compares the numerical results obtained using the linear programming formulation with values computed from the exact formulas. Since k^* satisfies the transcendental equation (37), the exact formula must be numerically estimated. Figure 1 displays the two-dimensional value function V obtained from the LP formulation, whereas Figure 2 graphs the one-dimensional section of the value function with $x = 1$ fixed.

5 Geometric Brownian Motion with Regime Switches

The previous two sections have demonstrated the accuracy of the LP approach on examples for which the theoretical solution is known. This section extends the example of the previous sections to include regime switches in order to demonstrate that the LP method can successfully solve stopping problems having complex dynamics.

Consider the pair of processes (X, Y) satisfying

$$
\begin{aligned}
X(t) &= x_0 + \int_0^t \mu(\epsilon(s))X(s)\,ds + \int_0^t \sigma(\epsilon(s))X(s)\,dW(s) \\
Y(t) &= y_0 \vee \max_{0 \leq s \leq t} X(s)
\end{aligned} \tag{41}
$$

	LP-results		NEV	
λ	k^*	$V(1,1)$	k^*	$V(1,1)$
0.100	3.398180	0.998403	3.397804	0.998390
0.200	2.247672	0.571501	2.248057	0.571497
0.300	1.911591	0.433497	1.911461	0.433497
0.400	1.742510	0.360870	1.742503	0.360870
0.500	1.638242	0.314702	1.638239	0.314702
0.600	1.566217	0.282189	1.566334	0.282188
0.700	1.513171	0.257763	1.513165	0.257762
0.800	1.471929	0.238578	1.471926	0.238578
0.900	1.438808	0.223011	1.438805	0.223011
1.000	1.411494	0.210062	1.411491	0.210062

Table 2: Optimal stopping points and optimal values as functions of the discount factor λ; the other parameters are $\sigma = 0.4$, $\mu = 0.03$, $x_0 = y_0 = 1$. Number of moments used is $M = 70$. The N(umerically) EV(aluated) numbers are based on formulae (36) and (37).

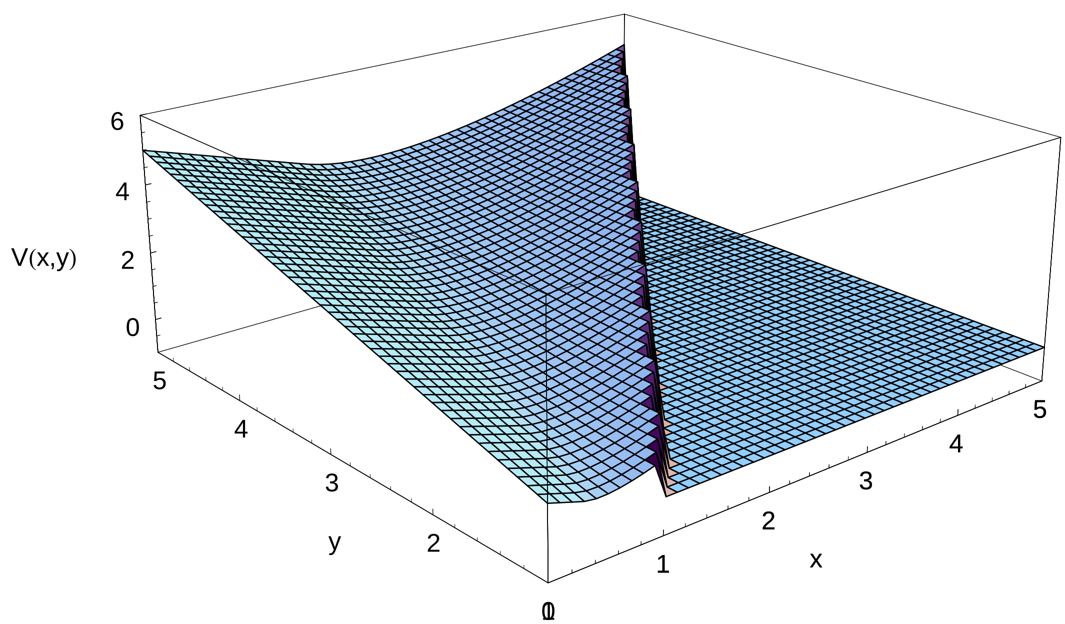

Figure 1: The value function $V(x, y)$, cf. (36), for the case $\lambda = 0.1$, $\sigma = 0.4$ and $\mu = 0.03$.

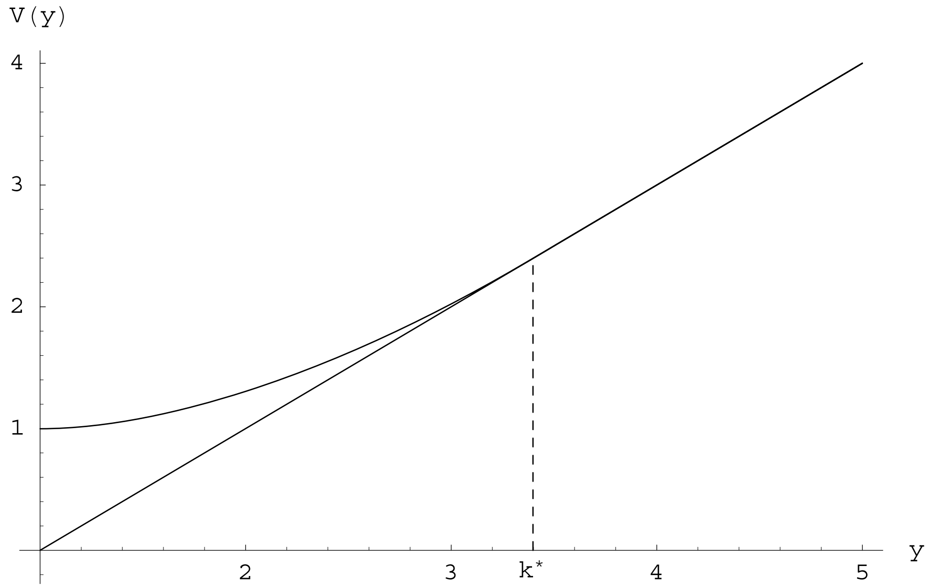

Figure 2: The value function $y \mapsto V(1,y)$. The parameters are the same as in Figure 1. Note, $V(1,1) \doteq 0.998$, cf. Table 2.

in which ϵ is a finite state, continuous-time Markov chain having states $\mathbf{M} = \{\epsilon_1, \ldots, \epsilon_m\}$ and generator $Qg(\epsilon_i) = \sum_{j\in\mathbf{M}}(g(\epsilon_j) - g(\epsilon_i))q_{ij}$ defined on functions g on $\mathbf{M}$, where $q_{ij} > 0$ gives the rate of change of the Markov chain from state ϵ_i to state ϵ_j, $i \neq j$. The goal is to find an optimal stopping time so as to maximize (2).

As in the case of simple geometric Brownian motion, the dimension can be reduced.

Proposition 5.1 *The optimal stopping problem for the perpetual lookback put option specified by maximizing (2) over stopping times τ when the process dynamics are given by (41) is equivalent to the stopping problem of maximizing*

$$x_0\tilde{E}\left[e^{-\int_0^\tau(\lambda-\mu(\epsilon(t)))\,dt}\,(Z(\tau)-1)\right] \tag{42}$$

where Z is a regime switching geometric Brownian motion process with reflection satisfying

$$dZ(t) = Z(t)\left[-\mu(\epsilon(t))\,dt + \sigma(\epsilon(t))\,d\tilde{W}(t)\right] + d\xi(t),$$

in which ξ is a monotone increasing process which increases only when $Z(t) = 1$.

Proof. The proof follows exactly as in Proposition 4.2. Note in particular that the characterization of the process $\frac{Y}{X}$ extends to the regime switching case with the proof being a minor modification of the argument given in [18]. □

Notice that the discount rate in (42) is now random and like the drift and diffusion coefficients depend on the state of the regime. It is therefore necessary to account for changes to the infinitesimal generator due to a switch in the regime. The generators of the pair process (Z,ϵ) are defined for $f \in \mathcal{D}$ and bounded, measurable functions g on $\mathbb{M}$ by

$$\begin{aligned} A[fg](z,\epsilon_i) &= \left[-\mu(\epsilon_i) z f'(z) + (1/2)\sigma^2(\epsilon_i) z^2 f''(z)\right] g(\epsilon_i) + f(z) Q g(\epsilon_i), \\ B[fg](z,\epsilon_i) &= f'(z) g(\epsilon_i). \end{aligned}$$

Let $\mathcal{D}_3 = \{fg : f \in \mathcal{D}, g \text{ bounded, measurable}\}$ denote the domain of the generators. Also, assume $Z(0) = z_0$ and $\epsilon(0) = \epsilon_0$, where $\epsilon_0 \in \mathbb{M}$ is given.

We now identify the LP formulation. It follows that for $fg \in \mathcal{D}_3$,

$$\begin{aligned} & e^{-\int_0^t (\lambda - \mu(\epsilon(s)))\, ds} f(Z(t)) g(\epsilon(t)) - f(z_0) g(\epsilon_0) \\ & - \int_0^t e^{-\int_0^s (\lambda - \mu(\epsilon(r)))\, dr} \left(A[fg] - (\lambda - \mu(\epsilon(s))) fg\right)(Z(s), \epsilon(s))\, ds \\ & - \int_0^t e^{-\int_0^s (\lambda - \mu(\epsilon(r)))\, dr} B[fg](Z(s-), \epsilon(s-))\, d\xi(s) \end{aligned}$$

is a martingale. Letting τ be a stopping time such that the optional sampling theorem applies, we have

$$\begin{aligned} 0 \;=\; & \tilde{E}\left[e^{-\int_0^\tau (\lambda - \mu(\epsilon(s)))\, ds} f(Z(\tau)) g(\epsilon(\tau)) - f(z_0) g(\epsilon_0)\right] \\ & - \tilde{E}\left[\int_0^\tau e^{-\int_0^s (\lambda - \mu(\epsilon(r)))\, dr} \left(A[fg] - (\lambda - \mu(\epsilon(s))) fg\right)(Z(s), \epsilon(s))\, ds\right] \\ & - \tilde{E}\left[\int_0^\tau e^{-\int_0^s (\lambda - \mu(\epsilon(r)))\, dr} B[fg](Z(s-), \epsilon(s-))\, d\xi(s)\right]. \end{aligned} \tag{43}$$

Defining the discounted distribution at time τ and the discounted occupation measures by

$$\begin{aligned} \nu_\tau(G) &= \tilde{E}\left[e^{-\int_0^\tau (\lambda - \mu(\epsilon(s)))\, ds} I_G(Z(\tau), \epsilon(\tau))\right], \\ \mu_0(G) &= \tilde{E}\left[\int_0^\tau e^{-\int_0^s (\lambda - \mu(\epsilon(r)))\, dr} I_G(Z(s), \epsilon(s))\, ds\right], \\ \mu_1(G) &= \tilde{E}\left[\int_0^\tau e^{-\int_0^s (\lambda - \mu(\epsilon(r)))\, dr} I_G(Z(s-), \epsilon(s-))\, d\xi(s)\right], \end{aligned}$$

the identity (43) becomes

$$f(z_0) g(\epsilon_0) = \int fg\, d\nu_\tau - \int (A[fg] - (\lambda - \mu) fg)\, d\mu_0 - \int B[fg]\, d\mu_1, \qquad \forall fg \in \mathcal{D}_3.$$

Note that μ in the above identity is a function of ϵ. The LP formulation of the problem is

$$\begin{aligned} & Max. \quad \langle z - 1, \nu_\tau \rangle \\ & S.t. \quad f(z_0) g(\epsilon_0) = \langle fg, \nu_\tau \rangle - \langle A[fg] - (\lambda - \mu) fg, \mu_0 \rangle - \langle B[fg], \mu_1 \rangle, \quad \forall fg \in \mathcal{D}_2 \\ & \qquad\;\; \langle 1, \nu_\tau \rangle \le 1, \\ & \qquad\;\; \langle 1, \mu_0 \rangle < \infty, \\ & \qquad\;\; \langle 1, \mu_1 \rangle < \infty. \end{aligned} \tag{44}$$

switching rates $(q_{12},\, q_{21})$	k_1^*	k_2^*	LP-value for (31)
(1, 1)	1.91	2.38	0.508862
(1.5, 1)	1.99	2.42	0.547346
(2, 1)	2.06	2.49	0.580254
(2.5, 1)	2.11	2.55	0.608618
(3, 1)	2.17	2.58	0.633340
(1, 1.5)	1.86	2.26	0.471539
(1, 2)	1.84	2.14	0.449419
(1, 2.5)	1.82	2.10	0.434750
(1, 3)	1.81	2.04	0.424254

Table 3: Optimal stopping threshholds (to 2 decimals) and the value function for geometric BM with regime switching as functions of pairs of switching rates; other parameters are $\lambda_1 = 0.4$, $\lambda_2 = 0.1$, $\sigma_1 = \sigma_2 = 0.4$, $\mu_1 = \mu_2 = 0.03$, $x_0 = y_0 = 1$ and $M = 70$.

Another rephrasing of the LP problem is beneficial. In place of the measures ν_τ, μ_0 and μ_1 on $[1, \infty) \times \mathbb{M}$, we break each of these measures into m (the number of regimes) measures $\nu_{\tau i}$, μ_{0i} and μ_{1i} on $[1, \infty) \times \{\epsilon_i\}$, respectively, by conditioning on the state of $\epsilon(t)$. This change combines with selecting $f(z) = z^n$ and $g(\epsilon) = I_{\{\epsilon_j\}}(\epsilon)$ to express (44) in terms of the new measures.

Note first that each μ_{1i} is still a point mass at $\{1\}$ since μ_1 measures the expected local time of Z at $\{1\}$. But in contrast to the case of geometric Brownian motion, the stopping measures $\nu_{\tau i}$ no longer have support at a single point k_i because the process (Z, ϵ) could enter the interior of the stopping region due to a switch in regime. We must therefore continue to use the moments of $\nu_{\tau i}$ in the LP formulation.

This last observation means that care must be taken in correctly formulating the LP. It is still reasonable to expect that for each regime, there will be a level k_i at or beyond which one would stop the process. The implication is that the measure $\nu_{\tau i}$ may have its support beyond $\{k_i\}$. However, letting $k_{max} = \max_{0 \leq i \leq m} k_i$, it follows that each of the measures will have support on the set $[k_i, k_{max}]$. Thus Hausdorff moment conditions (32) need to be suitably adjusted to the new intervals. This formulation now enables search procedures to be used to find the optimal levels for k_i^*, $i = 1, \ldots, m$.

For illustrative purposes, we use $m = 2$ regimes. Table 3 displays the optimal stopping threshholds k_1^* and k_2^* for a variety of switching rates between regimes. The optimal stopping levels for each of the regimes *when there is a single regime* are $k^* = 1.7425$ when $\lambda = 0.4$, and $k^* = 3.3978$ when $\lambda = 0.1$. Thus the optimal stopping levels are affected by the regime switching dynamics. Figure 3 displays the LP objective function as a function of the two switch points (k_1, k_2); the grid points are $\{(k_1, k_2) = (1.74 + ih, 2.20 + jh) : i, j = -5, -4, ..., 4, 5\}$ with $h = 0.01$.

A careful explanation is helpful to understand the numerical results. As observed before, the infinite-dimensional LP problem that uses all the moments will have exactly one feasible

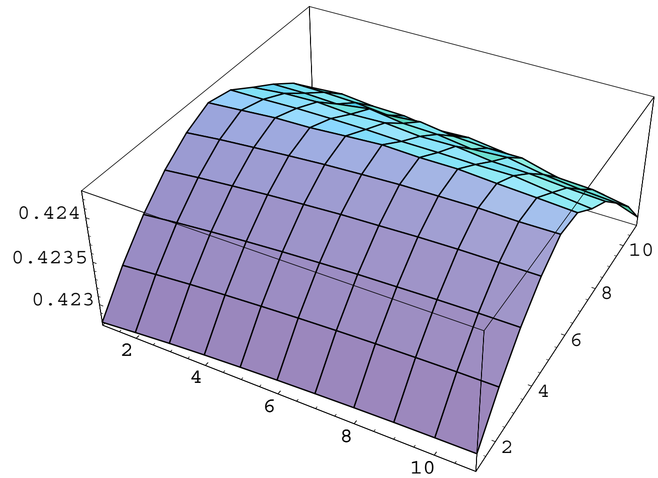

Figure 3: LP values as a function of (k_1, k_2) for geometric Brownian motion with regime switching for the case $q_{12} = 1$ and $q_{21} = 3$; all other parameters are as given in Table 3. Each square indicates a point on the grid $\{(k_1, k_2) = (1.74 + ih, 2.20 + jh) : i, j = -5, -4, ..., 4, 5\}$ with $h = 0.01$.

solution. When the number of variables is truncated at some level M, the feasible set becomes the Hausdorff polytope which includes the truncated moment sequence. Bounds on the value of the exit problem for specific k_1 and k_2 are obtained by solving a minimization and maximization problem for the objective function. The numerical results displayed in Table 3 and Figure 3 employ the search technique to find values k_1^*, k_2^* so as to maximize the lower bound on the value. This level of precision is only mentioned for the regime switching example since the search method for the stopping levels in the single regime case resulted in essentially the same values when maximizing the lower bound and the upper bound. In this example having regime switching, the results are not as numerically stable. We therefore choose to report the more conservative results associated with the lower bounds on the true optimal value.

References

[1] A. G. Bhatt and V. S. Borkar, Occupation measures for controlled Markov processes: Characterization and optimality, *Ann. Probab.*, **24** (1996), 1531–1562.

[2] M.J. Cho and R.H. Stockbridge, Linear programming formulation for optimal stopping problems, *SIAM J. Control Optim.* **40** (2002), 1965–1982.

[3] S.N. Ethier and T.G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, New York (1986).

[4] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 2, Wiley, New York (1965).

[5] S.E. Graverson and G. Peskir, On the Russian option: The expected waiting time, *Theor. Probab. Appl.*, **42** (1997), 564–575.

[6] F. Hausdorff, Momentprobleme für ein endliches intervall, *Mathematische Zeitschrift,* **16** (1923), 220–248.

[7] K. Helmes, Numerical Methods for Optimal Stopping using Linear and Nonlinear Programming, in *Stochastic Theory and Control*, B. Pasik-Duncan, ed., Lect. Notes Control Inform. Sci. **280**, Springer-Verlag, Berlin, (2002), 185–203.

[8] K. Helmes and S. Röhl, A Geometrical characterization of the multi-dimensional Hausdorff and Dale polytopes with applications to exit time problems, preprint.

[9] O. Hernandez-Lerma and J.B. Lasserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria,* Springer, New York (1996).

[10] O. Hernandez-Lerma and J.B. Lasserre, Approximation schemes for infinite linear programs, *SIAM J. Optim.,* **8** (1998), 973–988.

[11] O. Hernandez-Lerma and J.B. Lasserre, *Further Topics on Discrete-Time Markov Control Processes,* Springer, New York (1999).

[12] I. Karatzas and S.E. Shreve, Brownian Motion and Stochastic Calculus, 2nd ed., Springer-Verlag, New York, 1996.

[13] T.G. Kurtz and R.H. Stockbridge, Existence of Markov controls and characterization of optimal Markov controls, *SIAM J. Control. Optim.,* **36** (1998), 609–653.

[14] T.G. Kurtz and R.H. Stockbridge, Stationary solutions and forward equations for controlled and singular martingale problems, *Elect. J. Probab.,* **6** (2001), Paper No. 14, 1-52.

[15] A.S. Manne, Linear programming and sequential decisions, *Management Sci.*, **6** (1960), 259–267.

[16] B. Øksendal, Stochastic Differential Equations, 5th ed., Springer-Verlag, Berlin, 1998.

[17] G. Peskir, The Russian option: Finite horizon, *Finance Stoch.,* **9** (2005), 251-267.

[18] L.A. Shepp and A.N. Shiryaev, A new look at pricing of the "Russian Option", *Theory Probab. Appl.* **39** (1993), 103-119.

[19] S.E. Shreve, An introduction to singular stochastic control. *Stochastic differential systems, stochastic control theory and applications (Minneapolis, Minn., 1986), 513–528, IMA Vol. Math. Appl.*, **10**, Springer, New York-Berlin, 1988.

[20] R.H. Stockbridge, Time-average control of martingale problems: Existence of a stationary solution, *Ann. Probab.*, **18** (1990), 190–205.

[21] R.H. Stockbridge, Time-average control of martingale problems: A linear programming formulation, *Ann. Probab.*, **18** (1990), 206–217.

[22] L.C. Young, On approximation by polygons in the calculus of variations, *Proc. Royal Society, Series A*, **141** (1933), pp. 325–341.