Jukka Lempa
**On Infinite Horizon
Optimal Stopping of General
Random Walk**

Aboa Centre for Economics

Discussion Paper No. 3
Turku 2006

Jukka Lempa
On Infinite Horizon Optimal Stopping of General Random Walk

Aboa Centre for Economics
Discussion Paper No. 3
April 2006

ABSTRACT

The objective of this study is to provide an alternative characterization of the optimal value function of a certain Black-Scholes-type optimal stopping problem where the underlying stochastic process is a general random walk, i.e. the process constituted by partial sums of an IID sequence of random variables. Furthermore, the pasting principle of this optimal stopping problem is studied.

Contact information

Jukka Lempa
Turku School of Economics and Business Administration
Department of Economics
Quantitative Methods in Management
Rehtorinpellonkatu 3
20500 Turku
Finland
tel. +358 2 4814 314
fax +358 2 4814 302
jukka.lempa@tukkk.fi

1   Introduction

The purpose of this paper is to study a certain Black-Scholes-type infinite horizon optimal stopping problem where the underlying process is a general random walk. In more precise terms, let X be a random variable on $\mathbf{R}$ with a continuous law λ, mean $\mu > 0$ and variance $\sigma^2 < \infty$ and define the general random walk W on $\mathbf{R}$ as partial sums of IID random variables $X, X_1, X_2, \ldots$; i.e. let $W_n = X_1 + \cdots + X_n$, where $W_0 = 0$. For technical reasons it is assumed that

$$\exists\, \varepsilon > 0 \;:\; \mathbf{P}\{X > \varepsilon\} > 0 \text{ and } \mathbf{P}\{X < -\varepsilon\} > 0. \tag{1.1}$$

In other words, we assume that the distribution λ is not concentrated on either side of the origin. Note that in the case where $X \sim N(\mu, \sigma^2)$, the process W can expressed as $W_n = \mu n + \sigma \hat{W}\sqrt{n}$, where $\hat{W} \sim N(0,1)$. Given the process W, define the expected present value of exercise payoff gained after potentially infinite waiting period as

$$J(n, x) = \mathbf{E}\left[\beta^{-n}\left(e^{x+W_n} - c\right)^+\right] \tag{1.2}$$

and pose the optimal stopping problem

$$V(x) = \sup_{\eta \in \mathcal{N}} J(\eta, x), \tag{1.3}$$

where $\mathcal{N}$ is the set of all W-stopping times, $\beta > 1$ is the discount factor satisfying the condition $\mathbf{E}\left[\beta^{-1}e^X\right] < 1$ and $c > 0$ is the exercise cost. Note that in the case where $X \sim N(\mu, \sigma^2)$ the increments in the geometric random walk $Y_n =: e^{x+W_n}$ are log-normally distributed, which is a typical assumption in investment theoretical applications. The formulation (1.3) is well-established in mathematical finance. In particular, it is closely related to finding the value and exercise policy of a perpetual American call option in a Black-Scholes-type market driven in this case by a general random walk. Using this analogy, the increment $e^{W_{n+1}-W_n} = e^{X_{n+1}}$ can be interpreted as the relative price change in the period $n + 1$ and η is the date when the option is immediately and irreversibly exercised.

When studying maximization problems of the form (1.3) (possibly for a more general payoff structure) there is a number of different approaches to adopt. Perhaps the most general and fundamental approach is a direct application of principle of dynamic programming; for a recent treatment of dynamic programming in discrete time stochastic control, see Bertsekas and Shreve [4]. Another straightforward approach is to derive a set of conditions under which value function satisfies suitable monotonicity properties and growth rate restrictions and then utilize general martingale or other probabilistic techniques to establish the existence of a unique optimal stopping rule; see e.g. McKean [14] and Dubins and Teicher [13]. This set of conditions typically include convexity assumptions on the payoff. Yet another possible approach is the utilization of a powerful technique known as the

Snell envelope; see e.g. Snell [19] and Dalang and Hongler [9]. However, these approaches suffer from a downside, namely that they yield very little tangible information on the optimal characteristics, i.e. the optimal stopping rule or the value and typically they are accompanied by complementary techniques in order to gain more detailed information on the problem. In [10], Darling et al. solve the problem (1.3) and present a probabilistic characterization of the optimal characteristics in terms of the historical maximum of the driving random up to a certain independent, geometrically distributed random time. The characterizations by Darling et al. will be used as the starting point of our study.

The content of the paper is organized as follows. In section 2 the mathematical apparatus required by our analysis is presented. In section 3 closed-form representations of the optimal characteristics of the problem (1.3) are presented. In section 4 the pasting principle of the optimal value function is investigated. In section 5 the results are illustrated numerically and the study is concluded in section 6.

2   On the Minimal Functions of W

Denote as $L_1(W)$ the set of all functions $f : \mathbf{R} \to \mathbf{R}$ such that the expectation $\mathbf{E}\left[\beta^{-1}f(x+X)\right]$ exists and define the averaging Markov operator $\mathcal{P}_W$ on $L_1(W)$ as

$$(\mathcal{P}_W f)(x) := \mathbf{E}\left[\beta^{-1}f(x+X)\right] = \int_{-\infty}^{\infty}\beta^{-1}f(z)p(x,z)dz,$$

where $p(x,z) := \lambda(z-x)$ is the single-step transition density. A measurable function $u : \mathbf{R} \to \mathbf{R}_+ \cup \{\infty\}$ satisfying the condition $\mathcal{P}_W u(x) \leq u(x)$ for all $x \in \mathbf{R}$ is called β-excessive; in the case of an equality, the function u is called β-harmonic. A 1-excessive function is simply called excessive and, similarly, 1-harmonic function is called harmonic. Moreover, if a β-harmonic function h has the property that any β-harmonic function u with $u(x) \leq h(x)$ for all $x \in \mathbf{R}$ is proportional to h, then h is called β-minimal. Assume that h is β-excessive and define the function $p^h : \mathbf{R}^2 \to \mathbf{R}$ as

$$p^h(x,y) = \frac{h(y)}{\beta h(x)}p(x,y).$$

Since h is β-excessive, the function p^h is a transition density. Thus it constitutes a stochastic process. This process will be denoted as W^h and called the h-process of W.

The purpose of this section is to present a characterization of the minimal functions of the general random walk W and then utilize this characterization to actually determine the minimal functions. The characterization formulated in Theorem 2.1 is essentially due to Doob et al. [12]. In [12], the case where the driving general random walk is spatially discrete is considered in the absence of discounting. However, the proof they present can be straightforwardly generalized to

cover the present case by simply replacing their corresponding definitions with the ones presented above and carrying out the exactly same computations.

Theorem 2.1. *Assume, that the function $h : \mathbf{R} \to \mathbf{R}_+$ satisfies the condition $h(0) = 1$. Then h is β-minimal for the general random walk W if and only if it satisfies condition*

(A) $\mathbf{E}[\beta^{-1} h(X)] = 1$,

(B) $h(x + y) = h(x)h(y)$, *for all* $x, y \in \mathbf{R}$.

Theorem 2.1 is a forceful result on a general mathematical level. It is known from the theory of Martin boundaries that there is a fundamental connection between the minimal functions of a stochastic process and the minimal Martin compactification of the state space of the process (see e.g. Revuz [18], Chapter 7). Roughly speaking, this compactification is attained by embedding the state space in a suitable way in to a certain infinite-dimensional function space. In this light, there is no guarantee *a priori* that the minimal Martin compactification concurs with the elementary two-point compactification of the state-space. However, theorem 2.1 implies that in the case of a general random walk W, these two compactifications concur. This is equivalent to saying that there exist exactly two β-minimal functions of the general random walk W. This statement is now proved by utilizing Theorem 2.1.

Lemma 2.2. *There exists exactly two real numbers $-a$ and b, $a, b > 0$, determined by the condition $\mathbf{E}[e^{tX}] = \beta$ such that the functions $\psi : \mathbf{R} \to \mathbf{R}_+$ and $\varphi : \mathbf{R} \to \mathbf{R}_+$ defined as $\psi(x) = e^{bx}$ and $\varphi(x) = e^{-ax}$ are the only β-minimal functions of the general random walk W.*

Proof. It is well known that all positive solutions of the functional equation $h(x+y) = h(x)h(y)$ can be expressed in the form $h(x) = e^{tx}$ for $t \in \mathbf{R}$. For β-harmonicity the condition $\mathbf{E}[e^{tX}] = \beta$ must also be satisfied. Let M be the moment-generating function of X. Then $h(x) = e^{tx}$ is β-minimal if and only if $M(t) = \beta$. Define the function $\theta : \mathbf{R} \to \mathbf{R}$ as $\theta(t) = M(t) - \beta$. First note that θ is convex and $\theta(0) = 1 - \beta < 0$. Furthermore, by virtue of the standing assumption (1.1)

$$\lim_{t \to -\infty} \theta(t) \geq \lim_{t \to -\infty} \mathbf{E}[e^{tX}; X < -\varepsilon] - \beta > \lim_{t \to -\infty} e^{-t\varepsilon} \mathbf{P}\{X < -\varepsilon\} = \infty.$$

Using an identical argument, the condition $\lim_{t \to \infty} \theta(t) = \infty$ can also be established. This observation completes the proof. $\qquad \square$

Lemma 2.2 has a nice analogue in the theory of continuous time Markov processes. To point this out, consider the continuous time counterpart of the process W, namely the drifting Brownian motion $B_t^{(\mu)} = \exp\{(\mu + \sigma^2/2) t + \sigma B_t\}$ satisfying the stochastic differential equation $dB_t^{(\mu)} = \mu t + \sigma dB_t$, where B is a standard Brownian motion. It is well known that the β-minimal functions

3

of the process $B_t^{(\mu)}$ are the so-called fundamental solutions of the ordinary differential equation $\frac{1}{2}\sigma^2 h''(x) + \mu h'(x) - (\ln\beta)h(x) = 0$, in other words the increasing fundamental solution $\psi_B(x) = e^{\gamma x}$ and the decreasing fundamental solution $\varphi_B(x) = e^{-\delta x}$, where γ and $-\delta$ are the positive and the negative root of the characteristic equation $\frac{1}{2}\sigma^2 t^2 + \mu t - \ln\beta = 0$, respectively (see e.g. [5], pp. 17-18). Two interesting observations can now be made. First, for any particular choice of random variable X, the β-minimal functions of the processes $B^{(\mu)}$ and W are of the same functional form $x \mapsto e^{tx}$ for some parameter $t \in \mathbf{R}$. Moreover, if $X \sim N(\mu, \sigma^2)$, then the condition $\mathbf{E}\left[e^{bX}\right] = \beta$ from Theorem 2.2 can be written as $e^{b\mu + \frac{1}{2}\sigma^2 b^2} = \beta$, which implies that $\gamma = b$ and $\delta = a$. In other words, if $X \sim N(\mu, \sigma^2)$, then the β-minimal functions of the processes $B^{(\mu)}$ and W are the same.

To close the section, a scaled-down version of the integral representation theorem for harmonic functions of a general Markov chain is presented. For the complete formulation of the result and the proof, see Revuz [18], Corollary 3.11, pp. 257.

Theorem 2.3. *Assume that the process W has exactly two β-minimal functions, say ψ and φ, and that h is β-harmonic. Then there exists a unique pair (c_1, c_2) of non-negative constants such that $c_1 + c_2 = 1$ and $h(x) = c_1\psi(x) + c_2\varphi(x)$ for all $x \in \mathbf{R}$.*

3   On the Optimal Stopping Rule and Value Function

Typically, optimal stopping rules for maximization problems of the form (1.3) are characterized as passage times of the underlying stochastic process into the stopping region. This set is potentially very complex and its boundary, the optimal stopping threshold, can be virtually impossible to determine. However, in a number of practically meaningful cases it can be established that the stopping region is of the form (s^*, ∞), where $s^* \in \mathbf{R}$. In many cases, the threshold s depends on the properties of either the process itself or some random variable closely related to it, for example the historical maximum or minimum of the underlying process. First of these cases cases appears to be connected with processes having almost surely continuous sample paths (see e.g. Alvarez [2], Dayanik and Karatzas [11], Øksendal [16]) and second with processes exhibiting jump behavior (see e.g. Alili and Kyprianou [1], Boyarchenko and Levendorskiĭ [7], Darling et al. [10], Mordecki [15]). Given this observation, the study of the problem (1.3) is now continued by defining the random variable M as

$$M = \max_{0 \le k < \tau} W_k, \tag{3.1}$$

where τ is a random time which is independent of $\{X_i\}$ and geometrically distributed with $\mathbf{P}\{\tau > k\} = \beta^{-k}$ for $k \ge 0$. In other words, the random variable M is the historical maximum of the general random walk W up to a certain independent random time. Note that $M \ge 0$, since $W_0 = 0$. The information of the random variable M required by our analysis is now formulated in the following

two results.

Theorem 3.1. *Let $H_0^+ = \inf\{n \geq 0 : W_n > 0\}$. Then the following are equivalent:*

(A) $\mathbf{E}\left[\beta^{-1} e^X\right] < 1$

(B) $\mathbf{E}\left[\beta^{-H_0^+} \exp\left(W_{H_0^+}\right)\right] < 1$

(C) $\mathbf{E}\left[e^M\right] < \infty.$

Proof. See [10], pp. 1367. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Corollary 3.2. *The random variable M has an atom at origin; i.e. $\mathbf{P}\{M = 0\} > 0$.*

Proof. Recall the definition of random time H_0^+ from Theorem 3.1. Since $\exp\left(W_{H_0^+}\right) > 1$ almost surely, the part *(B)* of Theorem 3.1 implies that $\mathbf{P}\{M = 0\} = 1 - \mathbf{P}\{\tau > H_0^+\} = 1 - \mathbf{E}[\beta^{-H_0^+}] > 0$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

The next theorem gives a probabilistic characterization of the optimal characteristics of the problem (1.3). This theorem is essentially due to Darling et al. [10], where they consider the case $c = 1$ on pages 1367-8. However, their treatment generalizes easily to the case of general c, see also [15], Theorem 1.

Theorem 3.3. *The optimal stopping rule is to stop at time $H_{s^*} = \min\{n \geq 0 : x + W_n \geq s^*\}$, where $s^* = \ln\left(c\mathbf{E}\left[e^M\right]\right) < \infty$ and M is the random variable defined in (3.1). Moreover, the optimal value reads as*

$$V(x) = \mathbf{E}\left[\beta^{-H_{s^*}}\left(e^{x+W_{H_{s^*}}} - c\right)^+\right] = \frac{\mathbf{E}\left[\left(e^{x+M} - c\mathbf{E}\left[e^M\right]\right)^+\right]}{\mathbf{E}\left[e^M\right]}. \tag{3.2}$$

Useful information can be extracted from the representation (3.2). First of all, notice that V is nondecreasing and convex and that the first equation in the expression (3.2) implies that $V(x) = e^x - c$ for all $x \geq s^*$. On the other hand, the latter equation in the expression (3.2) implies that $V(x) = 0$ if and only if $e^M \leq e^{-x}c\mathbf{E}[e^M]$ almost surely which does not hold for any finite $x \leq s^*$. By combining this observation with the fact that the value V satisfies the principle of dynamic programming, i.e. that $V(x) = \max\left\{e^x - c, (\mathcal{P}_W V)(x)\right\}$, yields the condition $V(x) = (\mathcal{P}_W V)(x)$ for all $x \leq s^*$. Theorem 2.3 implies now that there exists a unique constant $K > 0$ such that $V(x) = Ke^{bx}$ for all $x \leq s^*$. Finally, since $V(x)$ is continuous in s^*, the constant $K = \frac{e^{s^*}-c}{e^{bs^*}}$. These results are now summarized in the following theorem, which is the first of our main results.

Theorem 3.4. *The optimal value function $V(x)$ reads as*

$$V(x) = \begin{cases} e^x - c, & x \geq s^* \\ \frac{e^{s^*}-c}{e^{bs^*}}e^{bx}, & x \leq s^*, \end{cases} \tag{3.3}$$

5

where b is the unique positive solution of the equation $\mathbf{E}\left[e^{bX}\right] = \beta$.

The function $V(x)$ is constructed from the functions $x \mapsto \frac{e^{s^*}-c}{e^{bs^*}}e^{bx}$ and $x \mapsto e^x - c$ by pasting them (possibly smoothly) together in the threshold s^*. Generalize now this function with respect to both the threshold s^* and the exponent b. More precisely, generate a whole family $\{G_{y,\alpha}(x)\}_{y\in\mathbf{R}_+,\ \alpha\in\mathbf{R}}$ of functions of the form (3.3) by first replacing the optimal stopping threshold s^* with a free boundary y and the critical exponent b with an arbitrary exponent α and defining the function $G_{y,\alpha} : \mathbf{R} \to \mathbf{R}_+$ as

$$G_{y,\alpha}(x) = \begin{cases} e^x - c, & x \geq y \\ \frac{e^y - c}{e^{\alpha y}}e^{\alpha x}, & x \leq y. \end{cases} \tag{3.4}$$

Using this notation, $V(x) = G_{s^*,b}(x)$. It is now natural to ask the question when the function $G_{y,\alpha}$ is continuously differentiable in y for a given $\alpha > 1$. Elementary differentiation yields that the function $G_{x^*,\alpha} =: G_{x^*_\alpha}$ is continuously differentiable in $x^*_\alpha \in \mathbf{R}_+$ if and only if $e^{x^*_\alpha} = \frac{\alpha c}{\alpha - 1}$. If now $\alpha = b$, this condition is the smooth pasting principle of the problem (1.3) and the threshold $x^*_b =: x^*$ is called the smooth pasting threshold. With this information, it is natural to pose the following question about the pasting principle: Is $s^* = x^*$?

For the sake of comparison to the continuous time setting, consider again the particular problem (1.3) where $X \sim N(\mu, \sigma^2)$ and define the continuous time version of the problem (1.3) where the underlying process is the drifting Brownian motion $B^{(\mu)}$, in other words the problem

$$V_B(x) = \sup_{\rho \in \mathcal{R}} \mathbf{E}\left[\beta^{-\rho}\left(e^{x+B_\rho^{(\mu)}} - c\right)^+\right], \tag{3.5}$$

where $B_t^{(\mu)}$ is the continuous time process introduced in Section 2 and $\mathcal{R}$ is the set of all $B^{(\mu)}$-stopping times. It has been established already in [14], Section 3 by McKean that in this case the optimal value V_B reads as $V_B(x) = G_{x^*_\gamma}(x)$, where the constant $\gamma > 1$ is positive solution of the equation $\frac{1}{2}\sigma^2 t^2 + \mu t - \ln\beta = 0$. Recall from Section 2 that $b = \gamma$. This implies that optimal stopping threshold x^*_γ of the continuous time problem (3.5) coincides with the smooth pasting threshold x^* of the discrete time problem (1.3) when $X \sim N(\mu, \sigma^2)$. Moreover, given that the problem (1.3) satisfies the smooth pasting principle, this would imply that values V and V_B satisfy the (rather counter-intuitive) condition $V_B(x) = V(x)$ for al $x \in \mathbf{R}$.

On a conceptual level, the representation (3.3) is analogous to the one presented in Alvarez [2], where the representation is given in terms of minimal functions of the underlying linear diffusion. From the point of view of the pasting principle, the representation in [2] is particularly convenient, since it gives the smooth pasting as a simple consequence. Moreover, the representation (3.3) is not completely unfamiliar to the literature of temporally discrete optimal stopping. In [20], Taylor considers essentially the same stopping problem as (1.3) and proves that the optimal value is bounded

6

from above by a function of the form (3.3). However, he makes no comment on whether or not the actual value of the problem is of the same form or on its connection to the Martin boundary theory.

4 Continuous Pasting vs. Smooth Pasting

Consider again the optimal stopping problem (3.5). It is a classical result (see McKean [14], Section 3) that for this problem the optimal value function $V_B(x)$ is continuously differentiable on the optimal stopping threshold x_γ^*. During the recent years, many authors have discussed the pasting principles of various optimal stopping problems and there is an increasing number of research articles reporting a failure of smooth pasting in the optimal value function, see e.g. Alili and Kyprianou [1], Asmussen et al. [3], Boyarchenko and Levendorskiĭ [6], Dalang and Hongler [9], Peskir and Shiryaev [17]. While going through these articles, one observes that the breakdown of smooth pasting appears to be connected to the cases when there is a chance that the underlying process can jump discontinuously into the stopping region. In this light, it is reasonable to guess that the smooth pasting fails also in the problem (1.3). In [1], Theorem 6, Alili and Kyprianou present an elegant characterization of the pasting principle for a problem of the form (1.3) in the case where the underlying process is a general Lévy process. They characterize the pasting principle in terms of the random variable M_L defined analogously to (3.1) for the driving Lévy process. Conveniently, this characterization holds also for the problem (1.3).

Theorem 4.1. *The optimal value function $V(x)$ exhibits smooth-pasting if and only if $M \neq 0$ almost surely.*

Proof. First, recall the expression (3.2) for the optimal value $V(x)$ from Theorem 3.3. Elementary manipulations yield

$$V(x) = c\mathbf{E}\left[\left(e^{-(s^*-x-M)} - 1\right); M > s^* - x\right] = c\left(e^{-(s^*-x)} - 1\right)\mathbf{E}\left[e^M; M > s^* - x\right] + c\mathbf{E}\left[\left(e^M - 1\right); M > s^* - x\right].$$

The last equality implies that

$$\frac{\left(e^{s^*} - c\right) - V(x)}{s^* - x} = \frac{c\left(\mathbf{E}\left[e^M\right] - 1\right) - V(x)}{s^* - x} = -c\frac{e^{-(s^*-x)} - 1}{s^* - x}\mathbf{E}\left[e^M; M > s^* - x\right] + c\frac{\mathbf{E}\left[\left(e^M - 1\right); M \leq s^* - x\right]}{s^* - x}.$$

In order to simplify the notation, denote last two terms on the right hand side as $A(x)$ and $B(x)$ respectively. Application of L'Hospitals rule yields

$$\lim_{x \to s^*-} A(x) = c\mathbf{E}\left[e^M; M > 0\right] = e^{s^*} - c\mathbf{E}\left[e^M; M \leq 0\right].$$

On the other hand, partial integration yields

$$
\begin{aligned}
B(x) =& c\frac{\mathbf{E}\left[\left(e^M - 1\right); 0 < M \leq s^* - x\right]}{s^* - x} = c\int_{0+}^{s^* - x} \frac{e^z - 1}{s^* - x} m(dz) \\
=& c\frac{e^{s^* - x} - 1}{s^* - x} m\{(0, s^* - x)\} + \frac{c}{s^* - x}\int_{0+}^{s^* - x} e^z m\{(0, z)\} dz \to 0, \text{ as } x \to s^*-,
\end{aligned}
$$

where in the first expectation the atom at origin is removed, this can be done because the integrand $e^M - 1 = 0$ when $M = 0$. Combination of these results yields

$$
\lim_{x \to s^*-} V'(x) = c\mathbf{E}\left[e^M; M > 0\right] = e^{s^*} - c\mathbf{E}\left[e^M; M = 0\right] = e^{s^*} - c\mathbf{P}\{M = 0\},
$$

which is clearly equivalent to the claim. $\qquad\square$

Coupled with Corollary 3.2, Theorem 4.1 yields immediately that smooth pasting fails in the problem (1.3). This result is our second main theorem.

Theorem 4.2. *The optimal value function V exhibits only continuous pasting on the optimal stopping threshold s^*. In other words, $s^* < x^*$.*

As was indicated earlier, the characterization of smooth pasting in Theorem 4.1 holds for a general Lévy processes, in particular for the drifting Brownian motion $B^{(\mu)}$. It is a well-known fact from the theory of Brownian motion that the sample paths of $B^{(\mu)}$ are regular in the sense that $\mathbf{P}\{H_{0,B}^+ = 0\} = 1$, where $H_{0,B}^+ = \inf\{t \geq 0 \mid B_t^{(\mu)} > 0\}$. This condition is clearly equivalent to the statement $\mathbf{P}\{M_B = 0\} = 0$, where the random variable M_B is defined analogously to (3.1) for $B^{(\mu)}$. It is also quite clear that the sample paths of the general random walk W are not regular in the previous sense. This is simply because of the fact that the random variable X admits negative values with positive probability.

5   An Illustration

In Section 4 it was established that the optimal value V does not exhibit smooth pasting in the optimal stopping boundary s^*. The aim of this section is to illustrate to size of the error being made in the case where $X \sim N(\mu, \sigma^2)$ if the smooth pasting principle is used as a basis of decision making. This error is illustrated using a simple quantity, namely the relative distance $D := \frac{x^*}{s^*}$ of the smooth pasting threshold x^* and the threshold s^*. Recall from Section 3 that in the current special case the threshold x^* coincides with the optimal stopping threshold x_γ^* of the problem (3.5) and therefore the distance D can also be seen as a difference between the discrete time model (1.3) and the continuous time model (3.5). This difference is of interest especially in investment-theoretical applications, where in many of cases the modelled phenomena (for example the evolution

of a stock price and the option pricing decisions based on this evolution) evolves in discrete time but the utilized mathematical model (for example the Black-Scholes model) takes place in continuous time. The difference between the models (1.3) and (3.5) can also be illustrated using another simple quantity, namely the relative point-wise distance $D_V(x) := \frac{V_B(x)}{V(x)}$ of the optimal value functions V_B and V. In the sequel, the quantities D and $D_V(x)$ will be illustrated graphically.

For simplicity, assume that $c = 1$. Then it is known from Darling et al. [10], pp. 1368 that the threshold s^* can be expressed as

$$s^* = \sum_{n=1}^{\infty} \frac{\beta^{-n}}{n} \mathbf{E}\left[e^{W_n^+} - 1\right] =: \sum_{n=1}^{\infty} \frac{\beta^{-n}}{n} s_n. \tag{5.1}$$

Since $W_n = \mu n + \sigma Y \sqrt{n}$, where $Y \sim N(0,1)$, the coefficient s_n can be expressed as

$$s_n = \frac{1}{\sqrt{2n\pi\sigma^2}} \int_0^{\infty} (e^z - 1) e^{\frac{-(z-n\mu)^2}{2n\sigma^2}} dz$$

$$= \frac{1}{2}\left(\exp\left[\frac{n}{2\sigma^2}\left(\left(\mu + \sigma^2\right)^2 - \mu^2\right)\right] \operatorname{erfc}\left[-\frac{\mu + \sigma^2}{2\sigma}\sqrt{2n}\right] - \operatorname{erfc}\left[-\frac{\mu}{2\sigma}\sqrt{2n}\right]\right),$$

where $\operatorname{erfc} : \mathbf{R} \to \mathbf{R}_+$ is the complementary error function defined as $\operatorname{erfc}(x) := \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-y^2} dy$. The thresholds x^* and s^* and the relative distance D are now illustrated in Figure 1 as functions of standard deviance σ under the assumption that $\mu = 0.03$ and $\beta = 1.07$. The approximations of s^* are computed from series (5.1) such that the reminder term $R < 10^{-7}$.
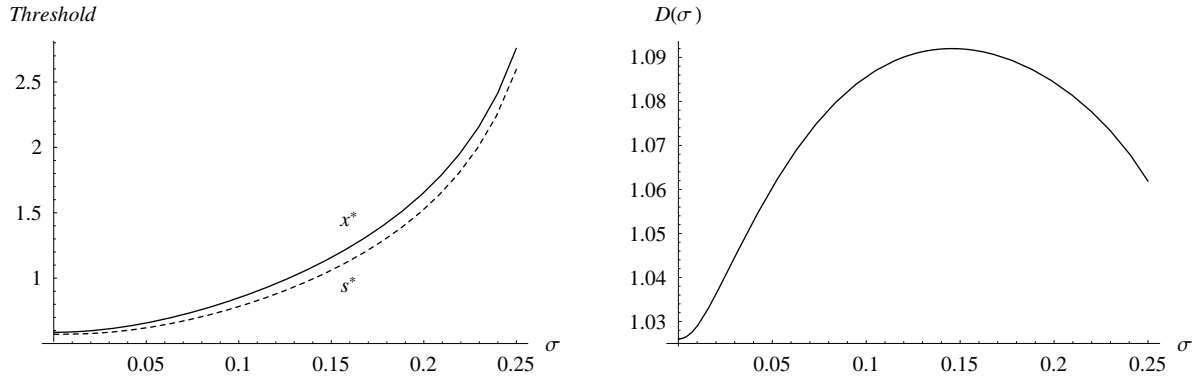

Figure 1. The smooth-fit threshold x^* (continuous curve), optimal stopping threshold s^* (dashed curve) and the relative error D as functions of volatility σ under the assumption $\mu = 0.03$ and $\beta = 1.07$.

The left hand side of Figure 1 indicates that for this specific example the thresholds x^* and s^* are both convex as functions of standard deviance σ but interestingly they are not equally "convex", as the right hand side clearly shows. In other words, the right hand side indicates that for small values of σ, the threshold x^* grows with a faster rate increasing the relative distance D until σ reaches a critical value σ^* at which D is maximal; for this specific parameter configuration the

9

critical $\sigma^* = 0.147$ at which $D(\sigma^*) = 1.092$. Above this critical value, the threshold s^* starts to gain on x^* and the relative distance D starts to decrease.

The behavior of the optimal value functions $V_B(x)$ and $V(x)$ and the relative distance $D_V(x)$ are now illustrated in Figure 2 under the assumption $\mu = 0.03$ and $\beta = 1.07$ in the case of maximal relative distance $D(\sigma^*)$, in other words when $\sigma = \sigma^* = 0.147$.
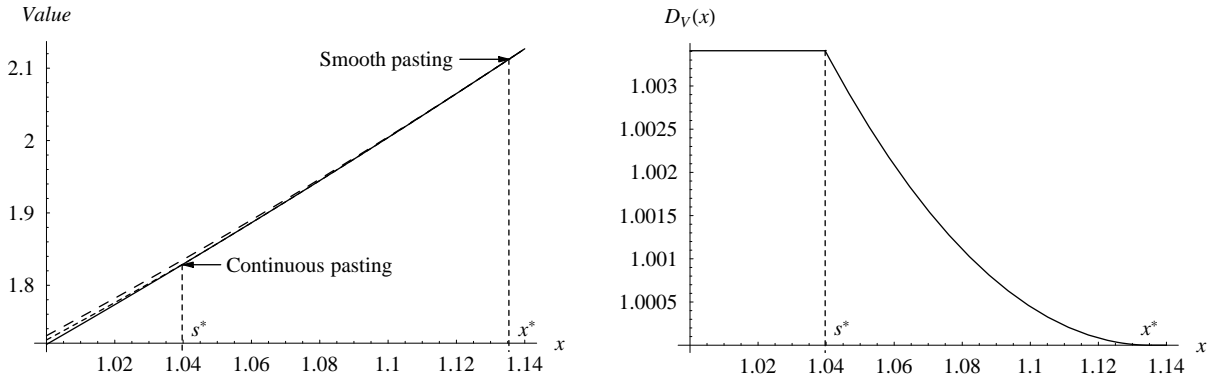
Figure 2. The optimal value function V_B of the problem (3.5) (upper dashed curve) and V of the problem (1.3) (lower dashed curve) and the reward $x \mapsto e^x - 1$ (continuous curve) under the assumption $\mu = 0.03$, $\sigma = 0.147$ and $\beta = 1.07$.

Figure 2 indicates that even though the thresholds x^* and s^* are relatively far away from each other $(D(\sigma^*) = 1.092)$, the curves $V_B(x)$ and $V(x)$ are quite close by in relative scale $(D_V(x) < 1.004$ for all $x \in \mathbf{R})$. Note that when $x < s^*$, the distance $D_V(x)$ is actually independent of x, in other words $D_V(x) = \frac{(e^{x^*}-1)e^{bs^*}}{(e^{s^*}-1)e^{bx^*}}$. On the interval (s^*, x^*) the distance $D_V(x)$ starts to decrease as the curve $x \mapsto e^x - 1$ gains on the curve $x \mapsto \frac{e^{s^*}-c}{e^{bs^*}}e^{bx}$ until D_V hits 1 at x^*.

The $D_V(x)$ can be seen as a simple relative measure of incompleteness of the model (1.3). Put somewhat differently, the difference $D_V(x)$ measures the point-wise relative loss of value caused by the restriction that the information on the underlying stochastic process and the exercise opportunities do not realize in continuous but in discrete time. Moreover, it is important to stress from the applications point of view that the difference $D_V(x)$ is a relative quantity and that small relative differences can sum up into substantial losses in the absolute scale.

6   Concluding Remarks

The presented paper considers the infinite horizon optimal stopping problem (1.3) of general random walk and it contains two main results. First, it presents an explicit formula (3.3) for the optimal value function in terms of β-minimal functions of the driving random walk. This representation of the value appears to be new in the discrete-time setting. In [20], Taylor analyzes a problem equivalent

to (1.3) and proves that for his problem the optimal value is dominated by a function of the form (3.3). However, he makes no comment about the actual value of the problem or its connections to the Martin boundary theory. The representation (3.3) is analogous to the one presented in Alvarez [2] in the case where the underlying process is a linear diffusion. In the case of a linear diffusion, where sample paths exhibit almost sure continuity, Alvarez's representation of the value is particularly convenient, since it gives smooth pasting as a simple consequence and therefore simplifies the characterization of the optimal stopping rule significantly. However, in the context of the current study, this representation does not yield any implication on the pasting principle of the problem. Therefore other techniques must be used in order to investigate the pasting principle. To this end, a characterization of smooth pasting is adopted from Alili and Kyprianou [1] and utilized to prove that the optimal value (3.3) is not differentiable in the optimal stopping threshold s^*.

The analysis of this study has a number of possible interesting extensions. First, a natural extension would be to consider a wider class of admissible control policies than just single stopping policies. More precisely, it would be of interest to extend the results of this paper to the sequential stopping problems appearing in the stochastic impulse control. These type of control problems appear for example in economics of renewable resources and cash flow management. Given the infinite horizon setting, a second natural extension would be the introduction of a stochastic interest rate structure. There is also room for generalization with respect to underlying stochastic dynamic structure. One possible way of extending the results in to this direction could be a development of some transformation technique of the underlying stochastic process (with respect to either time or scale) in order to be able to tackle more complicated dynamical systems, for example mean-reverting dynamics. However, these investigations are out of the scope of this study and are therefore left for future research.

References

[1] Alili, L. and Kyprianou, A.E. *Some Remarks on First Passage of Lévy Processes, the American Put and Pasting Principles*, 2005, *The Annals of Applied Probability*, **15(3)**, 2062–2080

[2] Alvarez, L. H. R. *Reward Functionals, Salvage Values, and Optimal Stopping*, 2001, *Mathematical Methods of Operations Research*, **54**, 315 – 337

[3] Asmussen S., Avram, F. and Pistorius M. *Russian and American Put Options Under Exponential Phase-Type Lévy Models*, 2004, *Stochastic Processes and Their Applications*, **109**, 79–111

[4] Bertsekas, D.P. and Shreve, S.E. *Stochastic Optimal Control: The Discrete-Time Case*, 1996, Athena Scientific, Belmont, Massachusetts

[5] Borodin, A. and Salminen, P. *Handbook of Brownian Motion - Facts and Formulæ*, 2002, 2nd Edition, Birkhäuser, Basel

[6] Boyarchenko, S. and Levendorskiĭ, S. *Pricing American Options Under Lévy Processes*, 2002, *SIAM Journal of Control and Optimization*, **40**, 1663–1696

[7] Boyarchenko, S. and Levendorskiĭ, S. *Practical Guide to Real Options in Discrete Time*, 2004, Social Science Research Network, http://ssrn.com/abstract=510324

[8] Black, R. and Scholes, M. *The Pricing of Options and Corporate Liabilities*, 1973, *Journal of Political Economy*, **81**, 637–659

[9] Dalang, R.C. and Hongler M.-O. *The Right Time to Sell a Stock Whose Price is Driven by Markovain Noise*, 2004, *The Annals of Applied Probablitity*, **14(4)**, 2176–2201

[10] Darling D. A., Liggett T. and Taylor H. M. *Optimal Stopping for Partial Sums*, 1972, *The Annals of Mathematical Statistics* **43**, 1363 – 1368

[11] Dayanik S. and Karatzas I. *On the Optimal Stopping Problem for One-Dimensional Diffusions*, 2003, *Stochastic Processes and their Applications*, **107(2)**, 173–212

[12] Doob, J. L., Snell J. L. and Williamson R. E. *Application of boundary theory to sums of random variables*, 1960, *Contributions to probability and statistics*, Stanford University Press, Stanford, California, 182 – 197

[13] Dubins, L. E. and Teicher, H. *Optimal Stopping When the Future is Discounted*, 1967, *The Annals of Mathematical Statistics*, Vol. **38(2)**, 601 –6 05

[14] McKean, H.P, Jr. *Appendix: A Free Boundary Problem for the Heat Equation Arising from a Problem of Mathematical Economics*, 1965, *Industrial Management Review*, **6**, 32–39

[15] Mordecki, E. *Optimal Stopping and Perpetual Options for Lévy Processes*, 2002, *Finance and Stochastics*, **6(4)**, 473–493

[16] Øksendal, B. *Stochastic Differential Equations*, 2000, 5th edition, 2nd Printing, Springer-Verlag, Berlin

[17] Peskir, G. and Shiryaev, A.N. *Sequential Testing Problems for Poisson Problems*, 2000, *The Annals of Statistics*, **28**, 837–859

[18] Revuz, D. *Markov Chains*, 1984, 2nd edition, North-Holland Publishing, New York

[19] Snell, J.L. *Applications of Martingale System Theorems*, 1952, *Transactions of American Mathematical Society*, **73(2)**, 293 – 312

[20] Taylor, H.M. *Bounds for Stopped Partial Sums*, 1972, *The Annals of Mathematical Statistics*, **43(3)**, 733-747