A computational analysis of the detection problem for Brownian motion with exponential penalty based on linear programming

T. Decker and K. Helmes

Humboldt University of Berlin, Institute of Operations Research, Germany

Abstract

The quickest detection problem of a Wiener process for the case of an exponential delay-penalty was recently solved by Beibel. He derived an explicit solution to the problem exploiting the equivalence of this detection problem to an optimal stopping problem of a 2-dimensional degenerate diffusion process. In this publication we shall compute the minimal risk and the optimal stopping rule – with and without additional constraints – using linear programming models. These models are derived from a general LP approach to optimal stopping. This approach is based on a characterization of a stopped Markov process through a family of equations which relate the generator of the process to a pair of measures representing the expected occupation of the process and the distribution of the state when the process is stopped. The computational analysis of the detection problem with exponential delay-penalty leads to bounds on the minimal risk and to a range for the optimal stopping threshold. In the case of no constraints the accuracy of the numerical results will be illustrated by comparing the numerical values with the known analytical ones. While we shall prove that each *constrained* problem is equivalent to a particular unconstrained detection problem this correspondence does not lead to an analytical characterization of the optimal stopping rule of such problems. We shall thus computationally analyze the constrained detection problems using the aforementioned LP technique and compute optimal values and optimal stopping thresholds.

1 Introduction.

In Beibel (2000) the quickest detection problem of Shiryayev (1975, chapter 4) is analyzed for the case of an exponential delay-penalty and for data which arrive continuously in time. The case of discrete time observations had been analyzed by Poor (1998). But the characterization of the optimal stopping rule and the optimal value are less explicit for the discrete case than for the case of continuous time data. In this paper we shall numerically solve the continuous time optimal stopping problem using linear programming (LP) techniques. Specifically, we will apply the golden section rule to a family of exit time problems whose values are sandwiched between the maximum- and the minimum-solution values of appropriately defined LP-problems. These LP-problems capture the dynamics of a two-dimensional degenerate diffusion which is associated with the detection problem with exponential penalty, see below. We shall describe the diffusion and the associated optimal stopping time problem adopting the notation used by Beibel.

Let $B = (B_t)_{t\geq 0}$ denote a standard Wiener process, let θ be a fixed, known real number and let τ be a nonnegative random variable. We assume that τ is independent of B and is exponentially distributed, i. e. $P[\tau > t] = \exp\{-\lambda t\}$ for all $t > 0$, where $\lambda > 0$ is known. The model assumes the data $W = (W_t)_{t\geq 0}$ to be sequentially observed and to be defined as, $0 \leq t < \infty$,

$$W_t = B_t + \theta(t-\tau)^+; \tag{1}$$

let $\mathcal{F}_t = \sigma(W_s : 0 \leq s \leq t)$. The detection problem consists of detecting τ as soon as possible, striking a balance between "false alarms", i. e. announcing the onset of the drift term before it happens, and long delays after the event has occured. Specifically, the objective is to find a stopping time T which has finite expectation, is adapted to $(\mathcal{F}_t)_{t\geq 0}$ and minimizes the risk

$$R(T) = P[T < \tau] + cE\left[e^{\alpha(T-\tau)^+} - 1\right], \tag{2}$$

where $c > 0$ and $\alpha > 0$ are given parameters.

The following two theorems which will be exploited in Section 2 summarize the main results of Beibel (2000).

Theorem 1.1 *The stochastic processes* $\pi = (\pi_t)_{t\geq 0}$*,* $\pi_t = P[\tau \leq t \mid \mathcal{F}_t]$*, and* $\tilde{\Psi} = (\tilde{\Psi}_t)_{t\geq 0}$*,* $\tilde{\Psi}_t = (1-\pi_t)^{-1}E\left[e^{\alpha(t-\tau)^+} \mid \mathcal{F}_t\right] - 1$*, satisfy the stochastic differential equations*

$$\begin{aligned} d\pi_t &= \lambda(1-\pi_t)dt + \theta\pi_t(1-\pi_t)d\bar{W}_t, \quad \pi_0 = 0, & (3)\\ d\tilde{\Psi}_t &= \left[(\lambda+\alpha+\theta^2\pi_t)\tilde{\Psi}_t + \lambda\right]dt + \theta\tilde{\Psi}_t d\bar{W}_t, \quad \tilde{\Psi}_0 = 0, & (4) \end{aligned}$$

where $\bar{W}_t := W_t - \int_0^t \pi_s ds$. *Moreover, for every stopping time* T *such that* $R(T) < \infty$ *the following representation of* $R(T)$ *holds:*

$$R(T) = E\left[(1-\pi_T)\Big(1 + c\big(1+\tilde{\Psi}_T\big)\Big)\right] - c. \tag{5}$$

To formulate the second result the following abbreviations will be used, see Beibel (2000) for additional details.

Let $\bar{\alpha} = 2\alpha/\theta^2$, $\bar{\lambda} = 2\lambda/\theta^2$ and, for $x \geq 0$, let

$$g(x) = \frac{\int_0^\infty e^{-u} u^{\gamma_1 - 1}(\bar{\lambda} + xu)^{\gamma_2 - 1} du}{(\bar{\lambda})^{\gamma_2 - 1}\Gamma(\gamma_1)}\,, \tag{6}$$

where $\Gamma(z) = \int_0^\infty e^{-u} u^{z-1} du$,

$$\gamma_1 = \frac{1}{2}(\bar{\lambda} + \bar{\alpha} - 1) + \sqrt{\frac{1}{4}(\bar{\lambda} + \bar{\alpha} - 1)^2 + \bar{\lambda}} \tag{7}$$

and

$$\gamma_2 = 1 - \frac{1}{2}(\bar{\lambda} + \bar{\alpha} - 1) + \sqrt{\frac{1}{4}(\bar{\lambda} + \bar{\alpha} - 1)^2 + \bar{\lambda}}. \tag{8}$$

Put $f(x) = [1 + c(1+x)]/g(x)$, $x \geq 0$, and let $v^* \in (0, \infty)$ be the unique minimizer of f, i. e.

$$f(v^*) = \min_{0 \leq v < \infty} f(v) > 0.$$

Define $S_{v^*} = \inf\{t \geq 0 \mid \tilde{\Psi}_t \geq v^*\}$ the first time the process $\tilde{\Psi}$ hits the threshold v^*.

Theorem 1.2 *For all stopping times* T *which are adapted to* $(\mathcal{F}_t)_{t\geq 0}$,

$$R(T) \geq R(S_{v^*}) = f(v^*) - c =: R^*\,, \tag{9}$$

i.e. S_{v^*} *is the optimal stopping time determined by the threshold value* v^*, *and the righthand side of (9) equals the minimal objective value, cf. (2).*

Remark 1.3 *For further use we shall collect some helpful facts about the function g which were also used by Beibel (2000, pp. 1698). The proof of these facts is straightforward.*
Fact 1: For all $x > 0$, $g'(x) > 0$ and $g''(x) < 0$.
Fact 2: $g(x) = (x)^{\gamma_2 - 1} \frac{\Gamma(\gamma_1 + \gamma_2 - 1)}{(\lambda)^{\gamma_2 - 1}\Gamma(\gamma_1)}(1 + o(1))$ *as* $x \to \infty$
Fact 3: $\lim_{x\to\infty} g(x) = \infty$ *.*
Fact 4: $\lim_{x\to\infty} g(x)/x = 0$ *.*

This paper is organized as follows. Section 2 presents two numerical methods for analyzing Beibel's detection problem and illustrates the accuracy of the numerical results. Method I uses the information, cf. Theorem 1.2, that the optimal stopping time of Beibel's problem is characterized by a stopping line in the $(\pi_t, \tilde{\Psi}_t)$-space which runs parallel to the π-axis. For each such stopping line we formulate finite dimensional linear programs which yield upper and lower bounds for the corresponding value of the objective value. We then apply line search techniques, e.g. the golden section rule, to find the optimal line. Together with Method I we give a brief survey of the relevant literature. A novel feature compared to previous numerical work on Shiryayev's detection problem by Helmes (2002a) is the use of corner point formulas of special polytopes to speed up computing time and to enhance numerical accuracy. Method II does not require prior information about the structure of the optimal stopping time but allows to deduce structural properties from the numerical results. In Section 3 we report on our computational analysis of the *constrained* detection problems that can be associated with Beibel's problem. Using Method II one finds that the optimal stopping time for any constrained problem is again specified by a threshold value. A careful analysis of the numerical results suggests a one-to-one correspondence between constrained and unconstrained problems. The formulation and proofs of the characterization theorems are given in Section 3 as well.

2 The LP method and first numerical results

Different numerical methods based on infinite dimensional linear programs for solving optimal stopping problems of Markov processes have been proposed by Cho (2001) and Röhl (2001). These methods build on the LP-techniques proposed by Helmes et al. (2000) for analyzing exit time problems of Markov processes; for related work see Cho and Stockbridge (2001), Helmes and Stockbridge (1999, 2000) and, more recently, Lasserre and Rumeau (2003). Modifications of Röhl's method using linear and nonlinear optimization techniques have been given by Helmes (2002a). The power of these methods for solving 1-dimensional stopping time problems were demonstrated by numerically analyzing Shiryayev's quickest detection problem and the pricing of perpetual Russian options, see Helmes (2002a,2002b). Here we apply these techniques to a 2-dimensional problem, cf. Section 1. The following theorem, Theorem 2.1 below, which is a special case of more general results which were proved by Cho and Röhl provides the analytical underpinning of our computational analysis. To formulate the theorem we introduce the following shorthand notation related to Beibel's problem, cf. Section 1:

- The generator A associated with the 2-dimensional diffusion $(\pi_t, \tilde{\Psi}_t)$ is given by, f a twice differentiable function on $(0,1) \times (0, \infty)$,

  $$\begin{aligned} Af(\pi, \tilde{\Psi}) &= \lambda(1-\pi)f_\pi + [(\lambda + \alpha + \theta^2\pi)\tilde{\Psi} + \lambda]f_{\tilde{\Psi}} \\ &\quad + 1/2 \cdot (\theta\pi(1-\pi))^2 f_{\pi\pi} + \theta^2\tilde{\Psi}\pi(1-\pi)f_{\pi\tilde{\Psi}} + 1/2 \cdot (\theta\tilde{\Psi})^2 f_{\tilde{\Psi}\tilde{\Psi}}\,; \end{aligned}$$

  the domain of the operator will be denoted by $\mathcal{D}(A)$ or $\mathcal{D}$. Note, A maps polynomial functions onto polynomial functions.

- $G := (0,1) \times (0, K)$, where $K > 0$ is some fixed number; the closure of G will be denoted by $\bar{G}$.

- For measures ν with support in $\bar{G}$ and integrable functions f defined on $\bar{G}$

  $$\langle f, \nu \rangle := \iint f(\pi, \tilde{\Psi})\, d\nu(\pi, \tilde{\Psi})\,.$$

- $\mathbb{1}$ denotes the function on $\bar{G}$ which is identical to 1.

- $L = (1-\pi)(1 + c(1+\tilde{\Psi})) - c$, and $x_0 = (0,0)$; note, L is a polynomial function.

- $\mathcal{T} := \{T \,|\, T \text{ is a feasible stopping time, cf. Theorem 1.1}\}$

- $\mathcal{M} := \{\nu \,|\, \nu \text{ is a finite measure on } \bar{G}\}$

- For $v > 0$, $S_v := \inf\{t \geq 0 \mid \tilde{\Psi}_t \geq v\}$, cf. Section 1

Theorem 2.1 *The optimal stopping problem defined in Section 1 is equivalent to the following infinite dimensional linear program:*

$$\inf_{T\in\mathcal{T}} R(T) = \inf_{\mu,\mu^T\in\mathcal{M}} \Big\{ \langle \mu^T, L\rangle \quad \Big| \quad \langle \mu^T, 1\rangle = 1, \\ \forall f\in\mathcal{D}, \langle f,\mu^T\rangle - f(x_0) - \langle Af,\mu\rangle = 0 \Big\} .$$

It follows from the extension of Hausdorff's characterization theorem of measures, e.g. Shohat and Tamarkin (1943), that any finite measure ν on $\bar{G}$ is uniqely determined by its moment sequence, $0\le i,j<\infty$,

$$m_{ij} = \iint_{\bar{G}} x^i y^j \, d\nu(x,y) .$$

Moreover, see Röhl (2001), the finite subsequence $(m_{ij})_{0\le i,j\le M}$, $M\in\mathbb{N}$ fixed, of any moment sequence $(m_{ij})_{0\le i,j<\infty}$ of a *probability measure* on the *unit square* in $\mathbb{R}^2$ is an element of the convex hull of the set $\{\vec{e}_{kl}\}_{0\le k,l\le M} \subset \mathbb{R}^{M+1}\times\mathbb{R}^{M+1}$, where the (i,j)-th coordinate of $\vec{e}_{kl}$ is given by

$$\vec{e}_{kl}(i,j) = \begin{cases} \binom{M-i}{k-i}\binom{M-j}{l-j} \big/ \binom{M}{k}\binom{M}{l} & , \quad k\ge i,\ l\ge j , \\ 0 & , \quad \text{else} . \end{cases} \tag{10}$$

We shall call the convex hull $\mathcal{H}^{M,2}$ the 2-dimensional Hausdorff polytope of order M. This characterization of $\mathcal{H}^{M,2}$ is equivalent to the one based on difference sequences which has previously been exploited by Helmes(2002a,2002b). The characterization of $\mathcal{H}^{M,2}$ by its corner points is a generalization of the anlogous result for *probability measures* on the *unit interval* on $\mathbb{R}$ which had been proved by Karlin and Shapley (1953). The corner points $\vec{f}_k$, $0\le k\le M$, of the 1-dimensional Hausdorff polytope $\mathcal{H}^{M,1}$ are given by

$$\vec{f}_k(i) = \begin{cases} \binom{M-i}{k-i} \big/ \binom{M}{k} & , \quad 0\le i\le k , \\ 0 & , \quad \text{else} . \end{cases} \tag{11}$$

We have experienced that using the explicit formulas for the corner points of $\mathcal{H}^{M,2}$ speeds up computing time and allows us to use larger values of M for the approximating finite dimensional LPs than would otherwise be possible.

Moment sequences of measures on bounded intervals in $\mathbb{R}^2$ other than the unit square, e.g. on $\bar{G}$ or subintervals of $\bar{G}$, are but scaled and garbled sequences corresponding to probability measures on $[0,1]\times[0,1]$; the detailed transformations will be exhibited below.

To numerically analyze Beibel's detection problem and the associated restricted versions we have used the following two different methods to approximately solve the corresponding optimal stopping problems, cf. Theorem 2.1:

2.1 Method I: The exit time approach

In case one has a-priori information about the structure of the optimal stopping policy, for instance, the process $(\pi_t, \tilde{\Psi}_t)_t$ should be stopped the first time the second component hits a particular value v^*, this information can be exploited as follows:

For every $v \in (0, K)$, K sufficiently large, consider the exit time problem of $(\pi_t, \tilde{\Psi}_t)_t$ from the domain $(0,1) \times (0, v)$, and compute

$$R(S_v) = E[L(\pi_{S_v}, \tilde{\Psi}_{S_v})]\,. \tag{12}$$

The main benefit of the linear programming approach is that the associated finite dimensional LP problems, see below, provide upper and lower bounds for the value (12). To formulate the finite dimensional LP programs we define transformations Λ_v, Ω_v resp., on $\tilde{\mathcal{H}}^{M,2}$, the cone generated by $\mathcal{H}^{M,2}$, and on $\mathcal{H}^{M,1}$ resp., as follows:
For any given $v > 0$, $h \in \tilde{\mathcal{H}}^{M,2}, g \in \mathcal{H}^{M,1}$, and $0 \le i, j \le M$, let

$$\Lambda_v h(i,j) := v^j h(i,j)\,, \tag{13}$$

and

$$\Omega_v g(j) := v^j g(j)\,. \tag{14}$$

The image of $\tilde{\mathcal{H}}^{M,2}$ under Λ_v, $\mathcal{H}^{M,1}$ under Ω_v resp., includes all truncated moment sequences of measures on $(0,1) \times (0,v)$, $(0,v)$ resp. If we only choose monomials (up to order M) as test functions f, Theorem 2.1 suggests to consider the following set of feasible points (note the definition of the generator A):

$$\mathcal{S}_{M,v} := \left\{ \begin{array}{l} (\vec{\mu}, \vec{\mu}^T) \in (\mathbb{R}^{M+1} \times \mathbb{R}^{M+1}) \times \mathbb{R}^{M+1} \;\Big| \\ \qquad\qquad \vec{\mu} \in \Lambda_v(\tilde{\mathcal{H}}^{M,2})\,,\ \vec{\mu}^T \in \mathcal{H}^{M,1} \text{ satisfy } (*) \text{ below} \end{array} \right\};$$

for all $0 \le i,j \le M-2$,

$$0^i 0^j = v^j \vec{\mu}^T(i) - \Big[i\lambda\vec{\mu}(i-1,j) + j\lambda\vec{\mu}(i,j-1) + \Big(\big[i(i-1)/2 + ij + j(j-1)/2\big]\theta^2 + j(\lambda+\alpha) - i\lambda\Big)\vec{\mu}(i,j) + \big(j - i(i-1) - ij\big)\theta^2\vec{\mu}(i+1,j) + i(i-1)\theta^2/2\vec{\mu}(i+2,j)\Big]. \qquad (*)$$

Next, we consider the two LP problems (note the definition of the pay-off function and recall that the distribution of the stopping location is concentrated on $[0,1] \times \{v\}$):

$$\underline{L}(v) := \min_{(\vec{\mu},\vec{\mu}^T) \in \mathcal{S}_{M,v}} \big\{(1+cv)\vec{\mu}^T(0) - (1+c(1+v))\vec{\mu}^T(1)\big\} \qquad (15)$$

and

$$\overline{L}(v) := \max_{(\vec{\mu},\vec{\mu}^T) \in \mathcal{S}_{M,v}} \big\{(1+cv)\vec{\mu}^T(0) - (1+c(1+v))\vec{\mu}^T(1)\big\}. \qquad (16)$$

By construction, the following two inequalities hold:

$$\inf_{v>0} \underline{L}(v) \le \inf_{\mu,\mu^T} \langle \mu^T, L \rangle \le \inf_{v>0} \overline{L}(v).$$

Assuming $\underline{L}$ and $\overline{L}$ to be unimodal functions we can use a line search technique, e.g. the golden section rule, to find bounds for $\inf_{\mu,\mu^T} \langle \mu^T, L \rangle$ as well as ϵ-optimal stopping times. If, up to numerical accuracy,

$$\inf_{v>0} \underline{L}(v) \doteq \underline{L}(v^*) \doteq \overline{L}(v^*) \doteq \inf_{v>0} \overline{L}(v),$$

then, for all practical purposes, we have actually identified the optimal threshold.

In Table 1, see subsection 2.3 , column "Method I" displays some values of $\inf_{v>0} \overline{L}(v)$ together with estimates of v^* for some parameter constellations. Note the excellent agreement of the LP bounds with the values of $R(S_{v^*})$ derived from Beibel's analytical results.

2.2 Method II: The refinement technique

While Method 1 exploits prior knowledge about the structure of the optimal stopping time, Method 2 provides this structural information. To be able to extraxt information from the numerical results we shall not only approximate the 2-dimensional occupation measures μ by transformed elements of $\tilde{\mathcal{H}}^{M,2}$, but we shall also approximate the distributions μ^T this way. Furthermore, we shall exploit a one-to-one correspondence between the corner points $\vec{e}_{kl}, 0 \leq k, l \leq M$, and the Dirac-measures $\delta_{(k/M,l/M)}(\cdot,\cdot)$ to obtain information about the support of the optimal exit distribution μ^T from the solution of the finite-dimensional LPs defined below. Moreover, we will represent any measure μ^T on $\bar{G}$ as

$$\mu^T = \mu^T_{|I_1} + \mu^T_{|I_2} + \mu^T_{|I_3}\,,$$

where, $0 = v_1 < v_2 < v_3 < v_4 = K, I_i, i = 1, 2, 3$, are the indicator functions

$$I_i(\pi,\tilde{\Psi}) := I_{[0,1]\times[v_i,v_{i+1}]}(\pi,\tilde{\Psi})\,.$$

The truncated moment sequences of measures defined on intervals like $[0,1]\times[v_2,v_3]$, are elements of the image of $\tilde{\mathcal{H}}^{M,2}$ under a generalization of the transformation Λ_v. For general $0 \leq a < b < \infty, h \in \tilde{\mathcal{H}}^{M,2}$ we define, $0 \leq i,j \leq M$,

$$\Lambda_{a,b}h(i,j) := \sum_{r=0}^{j}\binom{j}{r}(b-a)^r a^{j-r} h(i,r)\,;$$

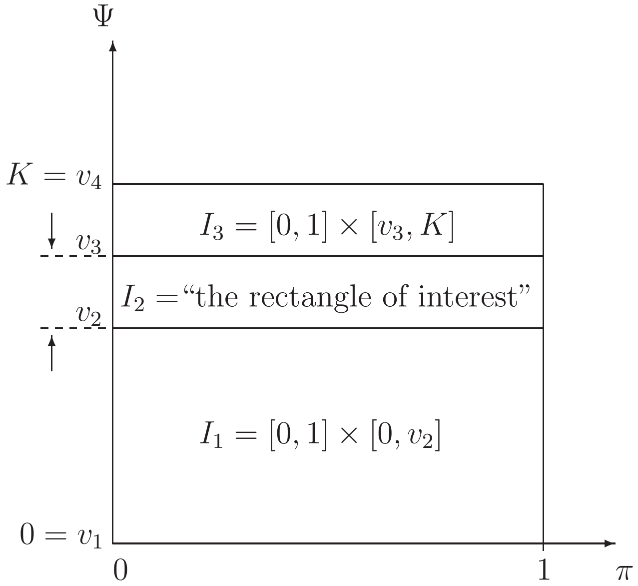
note, $\Lambda_{0,v} = \Lambda_v$. In case of Method II the following set of feasible points will be considered,

$$\mathcal{S}^{II}_{M,\{v_s\}} := \left\{ \begin{array}{r} (\vec{\mu},\vec{\mu}^T)\,\big|\,\vec{\mu}\in\Lambda_K(\tilde{\mathcal{H}}^{M,2}),\ \vec{\mu}^T_s\in\Lambda_{v_s,v_{s+1}}(\tilde{\mathcal{H}}^{M,2}), 1\leq s\leq 3,\ \vec{\mu}^T=\sum_{s=1}^{3}\vec{\mu}^T_s\,, \\ \vec{\mu}^T(0,0)=1,\ \text{and}\ (\vec{\mu},\vec{\mu}^T)\ \text{satisfy (**) below} \end{array} \right\};$$

for all $0 \leq i, j \leq M-2$,

$$\begin{aligned} 0^i 0^j = \vec{\mu}^T(i,j) - \Big[& i\lambda\vec{\mu}(i-1,j) + j\lambda\vec{\mu}(i,j-1) \\ & + \Big(\big[i(i-1)/2 + ij + j(j-1)/2\big]\theta^2 + j(\lambda+\alpha) - i\lambda\Big)\vec{\mu}(i,j) \\ & + \big(j - i(i-1) - ij\big)\theta^2\vec{\mu}(i+1,j) + [i(i-1)\theta^2/2]\vec{\mu}(i+2,j)\Big]\,. \end{aligned} \tag{**}$$

Figure 1: Partitioning of $\bar{G} = I_1 \cup I_2 \cup I_3$



With the notation introduced above the LP-problems that will be solved for every refinement $\{v_s\}_{1 \le s \le 4}, 0 = v_1 < v_2 < v_3 < v_4 = K$, K sufficiently large, will look very much like problems (15) and (16) although these are actually very different problems. To be specific, we shall consider for a given $M \ge 1$ and refinement $\{v_s\}$,

$$L_M^{II}(\{v_s\}) := \min_{(\vec{\mu}, \vec{\mu}^T) \in \mathcal{S}^{II}_{M,\{v_s\}}} \left\{ 1 - (1+c)\vec{\mu}^T(1,0) + c\vec{\mu}^T(0,1) - c\vec{\mu}^T(1,1) \right\}. \quad (17)$$

By construction, the following inequality holds:

$$L_M^{II}(\{v_s\}) \le \inf_{\mu, \mu^T} \langle \mu^T, L \rangle \,. \quad (18)$$

Note, Method II will only provide (good) lower bounds for the minimization problem. A trivial upper bound, which will typically not be very tight, is given by any feasible pair of measures (μ, μ^T), cf. Theorem 2.1.

The final ingredience of Method II is the way refinements are chosen. To this end, we define one further transformation on $\tilde{\mathcal{H}}^{M,2}$, viz. for $h \in \tilde{\mathcal{H}}^{M,2}, 0 \le$

$i, j \leq M$,

$$\Xi_M h(i,j) := \binom{M}{i}\binom{M}{j}(-1)^{M-i}(-1)^{M-j}\Delta_2^{M-j}\Delta_1^{M-i}h(i,j)\,,$$

where $\Delta_z^r, 0 \leq r \leq M, z = 1, 2$, denotes the r-th difference operator applied to the z-th component of 2-dimensional arrays, e.g.

$$\Delta_1^2 h(i,j) = \Delta_1(\Delta_1 h)(i,j) \quad \text{and} \quad \Delta_1 h(i,j) = h(i+1,j) - h(i,j), \quad \text{etc.}$$

As in the 1-dimensional case, cf. Feller (1965, pp 222), it follows by straightforward but somewhat lengthy computation that

$$\Xi_M \vec{e}_{k,l}(\cdot,\cdot) = \delta_{(k/M),(l/M)}(\cdot,\cdot)\,,$$

where $\delta_{(x,y)}(\cdot,\cdot)$ denotes the Dirac measure for a point (x, y) of the unit square. Thus, when applied to Beibel's problem, Method II can be summarized as follows:

Step 1. Solve L_M^{II} where $v_1 = 0, v_2 = v_3 = v_4 = K$; apply Ξ_M to the optimal program $\vec{\mu}^{T,*}$. Typically, $\{(i/M, Kj/M)|\Xi_M\vec{\mu}^{T,*}(i,j) > \delta_1\}$, $\delta_1 \approx 10^{-3}$, is a narrow strip parallel to the π-axis in $\bar{G}$, cf. Figures 1 and 5; the choice of the tuning parameter δ somewhat depends on the parameters of the problem.

Step 2. Choose $v_1 = 0, v_4 = K$ and $v_2 < v_3$ such that I_2 covers the strip identified by Step 1. Apply Ξ_M to all $\vec{\mu}_s^{T,*}, 1 \leq s \leq 3$; typically $\Xi_M\vec{\mu}_1^{T,*} \equiv \Xi_M\vec{\mu}_3^{T,*} \equiv 0$, and $\{(i/M, v_2 + j(v_3 - v_2)/M)|\Xi_M\vec{\mu}_2^{T,*}(i,j) > \delta_1\}$ is a narrow strip in $[0, 1] \times [v_2, v_3]$.

Step 3. Repeat Step 2, but increase v_2 and decrease v_3 to just cover the strip found in Step 2, cf. Figure 1. Stop if the total "mass" of $\Xi_M\vec{\mu}^{T,*}$ cannot be captured within a strip of size $v_3 - v_2 < \delta_2$ or less; typically, $\delta_2 \approx 10^{-2}$.
For "large" M, i.e. $M \approx 14$, it will turn out that very often the optimal threshold v^* is close to the midpoint of v_3 and v_2.

Table 1: Approximate values of the optimal stopping point v^* and the optimal value R^* based on Method I and repeated iterations of Method II for different parameters.

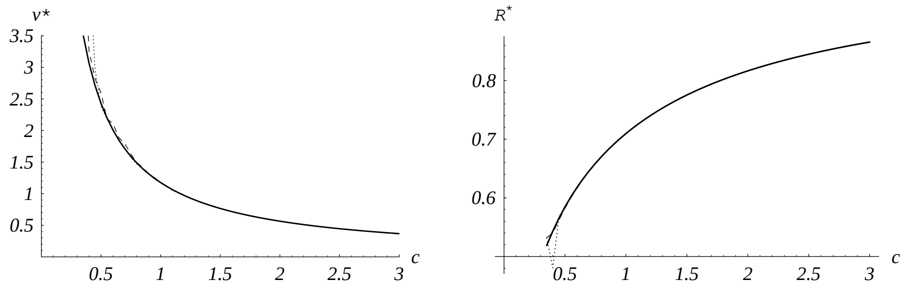
c	α		analytical	Method I	- - - - - M e t h o d II - - - - -			
λ	θ		solution		Step 1	1st refin.	2nd refin.	3rd refin.
1	1	v^*	1.17553	1.165697	1.08552	1.16659	1.17431	1.17598
1	1	R^*	0.70954	0.709694	0.70829	0.70919	0.70911	0.70918
3	1	v^*	0.36737	0.367652	0.337643	0.36666	0.367668	0.367389
1	1	R^*	0.86552	0.865522	0.86316	0.8654	0.86551	0.86551
1	3	v^*	0.35787	0.356857	0.336934	0.358333	0.358	0.357842
1	1	R^*	0.88288	0.88289	0.88074	0.8827	0.88279	0.88279
1	1	v^*	3.30573	3.361972	3.02402	3.2554	3.30619	3.298
3	1	R^*	0.47876	0.479021	0.47642	0.47794	0.477995	0.47796
1	1	v^*	3.1309	2.747672	no reasonable/reliable numerical solution found			
1	3	R^*	0.55263	0.549324				
6	1	v^*	0.17718	0.178098	0.169664	0.178333	0.177167	0.177166
1	1	R^*	0.92542	0.925418	0.92346	0.92533	0.92541	0.92542

2.3 Numerical results

Table 1 summarizes numerical results for several combinations of parameters. One can observe that for most parameters the numerical values are very close to the analytical ones, especially for repeated application of the refinement method. It is, however, obvious that at least for one parameter constellation the LP method failed to yield resasonable approximations.

We also investigated the behaviour of the stopping level and objective value with respect to variations of c; a graphical representation of the dependence can be found in Figure 2. The left-hand graph shows the stopping level v^* as a function of c for the analytical solution (bold line), the results of a line search technique applied to $\overline{L}$ (dotted line) and the search-method applied to $\underline{L}$ (dashed line). The right-hand graph shows the same graphs for the value R^* as a function of c. For a large range of parameters c, numerical results are very close to the values derived from the analytical expressions, cf. Section 1. If, however, c becomes small ($c \leq 0.5$), the numerical results deviate and display an irregular behaviour indicating that for such c-values the LP method cannot be relied upon.

Figure 2: The optimal stopping point v^* and the optimal value R^* as functions of c ($\alpha = \lambda = \theta = 1$). The bold line depicts the analytical solution. The dashed lines show the *argmin* and the optimal value ($\underline{L}$) of problem (15) while the dotted lines show the *argmax* and the optimal value ($\overline{L}$) of problem (16).



It is also interesting to see, cf. Lemmata 3.5 and 3.6 in Section 3, how the objective function value divides into its components *"error probability"* and *"expected exponential delay"* (reduced by one and multiplied by c). Figure 3 shows for c-values between 0.5 and 4 the error probability (dashed line), the exponential delay reduced by one and multiplied by c (dotted line), and the objective function value as the sum of the two components (bold line).

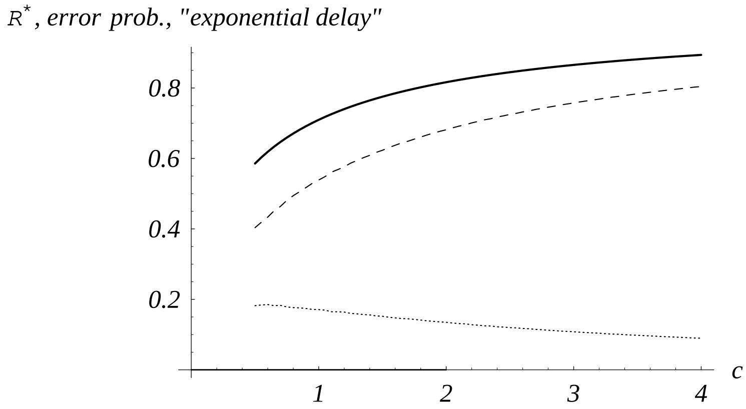
3 Constrained Problems

In this section we consider two modifications of Beibel's problem which we both analyse numerically using Method II. Method II is employed since we have no prior information on the optimal stopping region of the constrained problems. The numerical results suggest characterization theorems which will be proven afterwards. The following problem is a straightforward modification of the original problem:

$$\hat{R}(\beta) := \min_T E\left[e^{\alpha(T-\tau)^+} - 1\right] \quad \text{subject to} \quad P[T < \tau] \leq \beta\,, \qquad (\hat{\mathcal{P}}_\beta)$$

i.e. find a stopping T that minimizes the expected exponential delay (reduced by 1 according to the statement of the original problem) while the error probability must not exceed a prespecified level β.

Figure 3: The optimal value R^* (bold line), error probability $P[S_{v^*} < \tau]$ (dashed line) and expected exponential delay (dotted line) of S_{v^*} as functions of c ($\alpha = \lambda = \theta = 1$).



This modification can be easily accomodated in the original LP formulation by modifying the objective function and by adding the linear constraint $1 - \mu^T[1,0] \leq \beta$. Plots of the optimal value $\hat{R}(\beta)$ and the optimal stopping level $v^*(\beta)$ are depicted in Figure 4. The graphs suggest a reciprocal dependence of values $\hat{R}(\beta)$ and $v^*(\beta)$ on β.

A second straightforward modification of the original problem is to minimize the error probability while the expected exponential delay (reduced by 1)

Figure 4: The optimal value $\hat{R}$ and the optimal stopping point v^*as functions of β ($\alpha = \lambda = \theta = 1$).

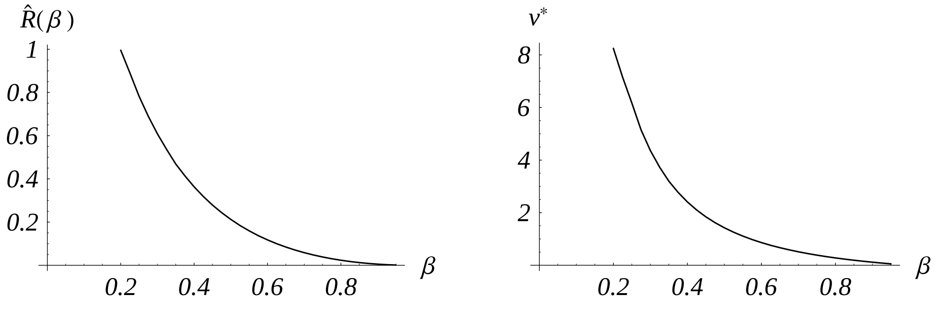
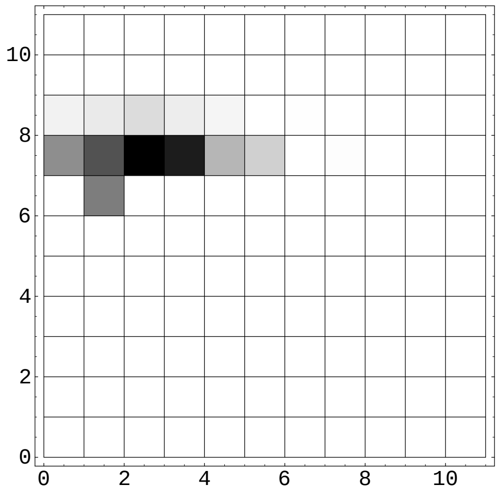


Figure 5: The grey-scale representation of the values $\Xi_M \vec{\mu}^{T,*}(i,j)$ after Step 1 of Method II applied to $\hat{\mathcal{P}}_\beta$, for $M = 10, \beta = 0.8, K = 0.4, \alpha = \lambda = \theta = 1$; i is marked on the horizontal axis, j is marked on the vertical axis; larger values are associsted with darker squares.



must not exceed a prespecified level γ:

$$\tilde{R}(\gamma) := \min_T P[T < \tau] \quad \text{subject to} \quad E\left[e^{\alpha(T-\tau)^+} - 1\right] \leq \gamma\,. \qquad (\tilde{\mathcal{P}}_\gamma)$$

Again, by adding the appropriate linear constraint and by changing the objective function the LP method will provide estimates of the optimal value and the optimal stopping time. In both cases, Method II suggests that the optimal stopping time for a constrained problem is again determined by a stopping line parallel to the π-axis, see Table 2 and Figure 5.

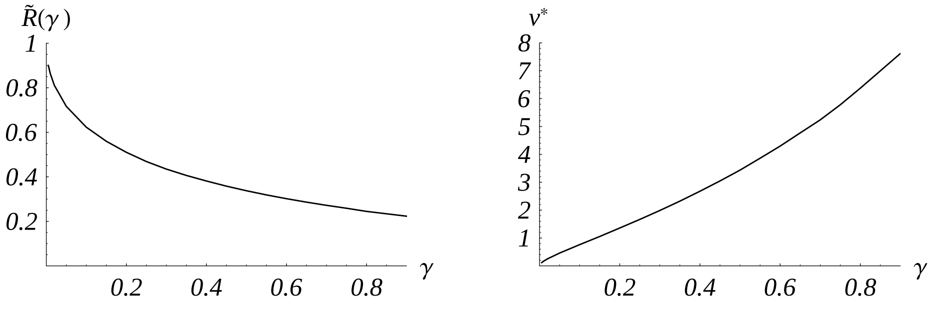
Figure 6 shows the graph of the optimal value $\tilde{R}(\gamma)$ and the stopping level as functions of γ. Moreover, a closer look at the numerical results suggests that for the same α, λ and θ-values there is a one-to-one correspondence between constrained problems and unconstrained ones.

To formulate this correspondence we shall use the following notation which

Table 2: The values $\Xi_M \vec{\mu}^{T,*}(i,j)$ after Step 1 of Method II applied to $\hat{\mathcal{P}}_\beta$, for $M = 10, \beta = 0.8, K = 0.4, \alpha = \lambda = \theta = 1$. The alignment of the numbers corresponds to the representation in Figure 5; $\Xi_M \vec{\mu}^{T,*}(0,0)$ is in the lower left corner of the Table.

0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.012	0.0193	0.0313	0.0168	0.0096	0.0008	0.0009	0.	0.	0.	0.
0.1007	0.1538	0.227	0.2017	0.0656	0.0421	0.	0.0027	0.	0.	0.
0.	0.1157	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.
0.	0.	0.	0.	0.	0.	0.	0.	0.	0.	0.

Figure 6: The optimal value $\tilde{R}$ and the optimal stopping point v^* as functions of γ ($\alpha = \lambda = \theta = 1$).



will be used throughout Section 3:

v_c^*	...	the optimal stopping level of the process $\tilde{\Psi}_t$ in Beibel's original problem with parameter c
T_c	...	the optimal stopping time for Beibel's original problem with parameter c, i.e. $T_c = S_{v_c^*}$, cf. Theorem 1.2
R_c	...	the optimal value of the original problem, i.e. $R_c = R(T_c) = R(S_{v_c^*})$
EP_c	...	the error probability corresponding to T_c, i.e. $P[T_c < \tau]$
EX_c	...	the expected exponential delay reduced by 1 corresponding to T_c, i.e. $E\left[e^{\alpha(T_c-\tau)^+} - 1\right]$

The following identity is an immediate consequence of these definitions:

$$R_c = EP_c + c \cdot EX_c \,. \tag{19}$$

The following characterization theorem precisely formulates the correspondence between the constrained problem $\hat{\mathcal{P}}_\beta$ and Beibel's unconstrained problem.

Theorem 3.1 *Let α, λ and θ be given. For each $\beta \in (0,1)$ there is a number $c = c(\beta) > 0$ such that*

(i) T_c is an optimal stopping time of $\hat{\mathcal{P}}_\beta$,

(ii) $EX_c = \hat{R}(\beta)$ and $EP_c = \beta$; $R_c = \beta + c\hat{R}(\beta)$.

The proof of the theorem is based on several lemmas.

Lemma 3.2 *$c \mapsto v_c^*$ (the optimal stoping level) is a continuous function in c.*

Proof The optimal stopping level v_c^* is the solution of the equation (see Beibel, p. 1699)

$$h(x) := \frac{g(x)}{g'(x)} - x - 1 = \frac{1}{c} \,. \tag{20}$$

Since

$$h'(x) = \left(\frac{g(x)}{g'(x)} - x - 1\right)' = -\frac{g(x)g''(x)}{[g'(x)]^2} \neq 0 \,, \tag{21}$$

h has a differentiable inverse, and thus $c \mapsto v_c^*$ is well defined and continuous. $\diamond$

Lemma 3.3 *$c \mapsto R_c$ is a concave function on $[0, \infty)$, and is continuous on $(0, \infty)$.*

Proof Let $0 < \lambda < 1$ and $c_1, c_2 \in [0, \infty)$ be given, put $c_{12} := \lambda c_1 + (1-\lambda)c_2$. By definition,

$$\begin{aligned}
R_{c_{12}} &= R_{\lambda c_1 + (1-\lambda)c_2} \\
&= EP_{c_{12}} + (\lambda c_1 + (1-\lambda)c_2)EX_{c_{12}} \\
&= \lambda EP_{c_{12}} + (1-\lambda)EP_{c_{12}} + \lambda c_1 EX_{c_{12}} + (1-\lambda)c_2 EX_{c_{12}} \\
&= \lambda[EP_{c_{12}} + c_1 EX_{c_{12}}] + (1-\lambda)[EP_{c_{12}} + c_2 EX_{c_{12}}] \\
&\geq \lambda[EP_{c_1} + c_1 EX_{c_1}] + (1-\lambda)[EP_{c_2} + c_2 EX_{c_2}] \\
&= \lambda R_{c_1} + (1-\lambda)R_{c_2} .
\end{aligned}$$

Since $c \mapsto R_c$ is concave, the function is continuous in the interior of its domain. $\diamond$

Remark 3.4 *An alternative proof of Lemma 3.3 runs as follows: The optimal stopping time for Beibel's problem with parameter c, when used for a problem with parameter $\hat{c}$, has the objective value $EP_c + \hat{c}EX_c$. The optimal value $R_{\hat{c}}$ for parameter $\hat{c}$ will be at least as good as, or even better (smaller), than this sum,*

$$R_{\hat{c}} \leq EP_c + \hat{c}EX_c = R_c + (\hat{c} - c)EX_c \, ,$$

i.e. (for fixed c) $R_c + (\hat{c} - c)EX_c$ is a linear function of $\hat{c}$ which lies above the value function and is equal to the value function at point c. Such linear functions exist for every c. Hence the value function R_c is concave, and therefore continuous in the interior.

Lemma 3.5 *$c \mapsto EP_c$ and $c \mapsto EX_c$ are continuous functions in c.*

Proof The process $\tilde{\Psi}_t$ has been defined as the quotient of the expected exponential delay conditioned on information up to time t and the probability of false alarm at time t, minus 1:

$$\tilde{\Psi}_t = \frac{E\left[e^{\alpha(t-\tau)^+} \mid \mathcal{F}_t\right]}{1 - \pi_t} - 1 \, .$$

By definition, the process $\tilde{\Psi}_t$ equals v_c^* when it hits the stopping level v_c^*. Therefore, no matter what the individual values of $(1-\pi_t)$ and $E\Big[e^{\alpha(t-\tau)^+} \mid \mathcal{F}_t\Big]$ might be, we know that at time T_c the ratio of these two values equals (for almost all $\omega \in \Omega$)

$$v_c^* + 1 = \tilde{\Psi}_{T_c} + 1 = \frac{E\Big[e^{\alpha(T_c-\tau)^+} \mid \mathcal{F}_{T_c}\Big]}{1-\pi_{T_c}}\,;$$

this identity is equivalent to

$$E\Big[e^{\alpha(T_c-\tau)^+} \mid \mathcal{F}_{T_c}\Big] = (1+v_c^*)(1-\pi_{T_c})\,.$$

Taking expectations on both sides we obtain

$$EX_c + 1 = (1+v_c^*)EP_c\,. \tag{22}$$

Equation (22), which relates EP_c, EX_c and v_c^*, together with the definition of R_c imply

$$R_c = EP_c + c\big((1+v_c^*)EP_c - 1\big) \quad \Leftrightarrow \quad EP_c = \frac{c+R_c}{1+c(1+v_c^*)}\,, \tag{23}$$

and

$$R_c = (1+v_c^*)^{-1}(EX_c+1) + cEX_c \quad \Leftrightarrow \quad EX_c = \frac{(1+v_c^*)R_c - 1}{1+c(1+v_c^*)}\,. \tag{24}$$

Since both, $c \mapsto v_c^*$ and $c \mapsto R_c$, are continuous functions in c, Lemma 3.5 follows. $\diamond$

Lemma 3.6 *(i)* $\lim_{c\searrow 0} EP_c = 0$ *and (ii)* $\lim_{c\to\infty} EP_c = 1$.

Proof (i) The asymptotic expansion, cf. Beibel (2000, p. 1699),

$$v_c^* = \frac{1}{c}\frac{\gamma_2 - 1}{2-\gamma_2}(1+o(1)) \quad \text{as} \quad c \to 0\,, \tag{25}$$

which follows from the definition of v_c^*, cf. (20), and the integral representation of g, cf. (6), implies $v_c^* \to \infty$ as $c \to 0$.

Moreover, see Theorem 1.2, the value function R_c equals $f(v_c^*) - c$ and, by definition of g, this expression can be written as

$$R_c = \frac{1 + c(v_c^* + 1)}{g(v_c^*)} - c\,. \tag{26}$$

Equations (23) and (26) together imply

$$EP_c = 1/g(v_c^*)\,. \tag{27}$$

While the numerator of the fraction is a constant, the denominator grows to infinity as c tends to 0 since v_c^* grows to infinity,,see above, and so does $g(v_c^*)$, cf. Remark 1.3, *Fact 3*. Thus $EP_c \to 0$ as $c \to 0$.

(ii) v_c^* is the solution of the equation, cf. (20), $h(x) = g(x)/g'(x) - x - 1 = 1/c$. Since $g(0) = 1$ and $g'(0) = 1$ it follows that $h(0) = 0$. We have shown in the proof of Lemma 3.2 that $h'(x) \neq 0$ for $x > 0$. Hence h has a continuous inverse with $h^{-1}(0) = 0$. If $c \to \infty$, i.e. $1/c \to 0$, the continuity of h^{-1} implies

$$v_c^* = h^{-1}(1/c) \to 0 \quad \text{as} \quad c \to \infty\,;$$

this convergence property, together with $g(0) = 1$ and the identity $EP_c = 1/g(v_c^*)$, implies $\lim_{c\to\infty} EP_c = 1$.

Now we are able to complete the proof of Theorem 3.1.

Proof The two Lemmata 3.5 and 3.6, combined with the intermediate value theorem imply that for each $\beta \in (0,1)$ there exists a value $c := c(\beta)$ such that $EP_c = \beta$. Use this particular value c in the original problem; thus $R_c = \beta + c \cdot EX_c$. Let $T_c = T_{c(\beta)}$ be the corresponding optimal stopping time. By construction, T_c is a feasible stopping time for the modified problem $\hat{P}(\beta)$ which yields the value $EX_{c(\beta)}$. Now let us assume $\hat{T}_\beta$ were a better stopping time for the modified problem $\hat{P}(\beta)$ than T_c, i.e. $\hat{T}_\beta$ satisfies the constraint $P[\hat{T}_\beta < \tau] \leq \beta$ and $\hat{R}(\beta) < EX_{c(\beta)}$. Then $R(\hat{T}_\beta) < R_c$, i.e. $\hat{T}_\beta$ would also yield a smaller value in the original problem, which contradicts the assumption that $T_{c(\beta)}$ is an optimal stopping time in Beibel's problem. This concludes the proof of Theorem 3.1. $\diamond$

The next result, which complements Theorem 3.1, is the characterization theorem for $\tilde{\mathcal{P}}_\gamma$.

Theorem 3.7 *Let α, λ and θ be given. For each $\gamma \in (0, \infty)$ there is a number $c = c(\gamma) > 0$ such that:*

(i) T_c is an optimal stopping time of $\tilde{\mathcal{P}}_\gamma$,

(ii) $EP_c = \tilde{R}(\gamma)$ and $EX_c = \gamma$; $R_c = \tilde{R}(\gamma) + c\gamma$.

Besides Lemma 3.5 the proof of Theorem 3.7 requires the following addition to Lemma 3.6.

Lemma 3.8 *(i)* $\lim_{c \searrow 0} EX_c = \infty$ *and (ii)* $\lim_{c \to \infty} EX_c = 0$

Proof (i) Equation (26), cf. proof of Lemma 3.6, together with equation (24) imply

$$EX_c = \frac{1}{g(v_c^*)} + \frac{v_c^*}{g(v_c^*)} - 1. \tag{28}$$

We know, cf. proof of Lemma 3.6, that $v_c^* \to \infty$ as $c \to 0$. Moreover, cf. Remark 1.3, *Fact 4*, $g(.)$ grows more slowly than the identity. Hence, $\lim_{v \to \infty} v/g(v) = \infty$, which implies $EX_c \to \infty$ as $c \to 0$.

(ii) The value function R_c is bounded from above by 1 since immediate stopping is always a feasible policy; this particular policy has error probability 1 and no delay. The identity $R_c = EP_c + c \cdot EX_c$ is equivalent to

$$EX_c = 1/c \cdot (R_c - EP_c) \,. \tag{29}$$

Since $\lim_{c \to \infty} EP_c = 1$ (see Lemma 3.6) and $R_c \leq 1$, for all $c \geq 0$, it follows from equation (29) that $\lim_{c \to \infty} EX_c = 0$. $\diamond$

Remark 3.9 *The following argument shows that the function $c \mapsto EX_c$ is actually non-increasing on $(0, \infty)$.*

Proof We first show that the value function R_c is a non-decreasing function of c. Let $c, \delta > 0$; by definition,

$$\begin{aligned} R_c &= EP_c + cEX_c \\ &\leq EP_{c+\delta} + cEX_{c+\delta} \\ &= R_{c+\delta} - \delta EX_{c+\delta} \,; \end{aligned} \tag{30}$$

actually, since $EX_{c+\delta} > 0$, the strict inequality $R_c < R_{c+\delta}$ holds. By the same argument we have:

$$\begin{aligned} R_{c+\delta} &= EP_{c+\delta} + (c+\delta)EX_{c+\delta} \\ &\leq EP_c + (c+\delta)EX_c \\ &= R_c + \delta EX_c \,. \end{aligned} \tag{31}$$

Combining the last inequality with inequality (30) we obtain:

$$\begin{aligned} R_{c+\delta} &\leq R_{c+\delta} - \delta EX_{c+\delta} + \delta EX_c \\ \Rightarrow \quad EX_{c+\delta} &\leq EX_c \,. \end{aligned} \qquad \diamond$$

Remark 3.10 *Unless $EX_c = EX_{c+\delta} = 0$, the inequality in the last line of the proof will be strict.*

Now we are able to prove Theorem 3.7.

Proof We are going to use the same kind of arguments as in the proof of Theorem 3.1. Lemmas 3.5 and 3.8 together with the intermediate value theorem imply that for each $\gamma \in (0, \infty)$ there exists $c := c(\gamma)$ such that $EX_c = \gamma$. Use this particular value c in the original problem, thus $R_c = EP_c + c \cdot \gamma$. Let $T_c := T_{c(\gamma)}$ be the corresponding optimal stopping time. By construction, T_c is a feasible stopping time for problem $\tilde{P}(\gamma)$ with value $EP_{c(\gamma)}$. Assume $\tilde{T}_\gamma$ were a better stopping than T_c time for problem $\tilde{P}(\gamma)$, i.e. $\tilde{T}_\gamma$ satisfies the constraint $E\left[e^{\alpha(T-\tau)^+} - 1\right] \leq \gamma$ and $\tilde{R}(\gamma) < P[T_c \leq \tau]$. Then $R(\tilde{T}_\gamma) < R_c$, i.e. $\tilde{T}_\gamma$ would also yield a smaller value in the original problem, which contradicts the assumption that T_c is an optimal stopping time in Beibel's problem. $\diamond$

References

Beibel, M. (2000). A note on sequential detection with exponential delay penalty. *Ann. Statist.* **28**, 1696–1701.

Cho, M. J. (2000). Linear programming formulation for optimal stopping. *PhD thesis,* The Graduate School, University of Kentucky, Lexington.

Cho, J. C. and Stockbridge, R. H. (2002) Linear programming formulation for optimal stopping Problems. *SIAM J. Control Optim.* **40**, 1965–1982.

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

Helmes, K. (2002). Numerical methods for optimal stopping using linear and non-linear programming. In Series: Lecture Notes in Control and Information Sciences **280**. *Proceedings of a Workshop "'Stochastic Theory and Control"', held in Lawrence, Kansas.* (ed. Pasik-Duncan, B.), Springer–Verlag, Berlin, 185–202.

Helmes, K., Röhl, S. and Stockbridge, R. H. (2001). Computing moments of the exit time distribution for Markov processes by Linear Programming. *Oper. Res.* **49**, 516–530.

Helmes, K. and Stockbridge, R. H. (2000). Numerical comparison of controls and verification of optimality for stochastic control problems. *J. Optim. Th. Appl.* **106**, 107–127.

Helmes, K. and Stockbridge, R. H. (2001). Numerical evaluation of resolvents and Laplace transforms of Markov processes. *Math. Methods Oper. Res.* **53**, 309–331.

Karlin, S. and Shapley, L. (1953). Geometry of moment spaces. *Mem. Amer. Math. Soc.* **12**.

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

Kurtz, T. G. and Stockbridge, R. H. (1999). Martingale problems and linear programs for singular control. *Thirty-Seventh Annual Allerton Conference on Communication, Control, and Computing (Monticello, Ill.),* 11-20, Univ. Illinois, Urbana-Champaign, Ill.

Lasserre, J. B. and Prieto Rumeau, T. (2003). SDP vs. LP relaxations for the moment approach in some performance evaluation

problems. *Rapport LAAS No. 03125.*
POOR, H. V. (1998). Quickest detection with exponential penalty for delay. *Ann. Statist.* **26**, 2179–2205.
RÖHL, S. (2001). Ein linearer Programmierungsansatz zur Lösung von Stopp- und Steuerungsproblemen. *Ph. D. Dissertation* Humboldt-Universität zu Berlin, Berlin, Germany.
SHIRYAYEV, A. N. (1978). *Optimal Stopping Rules,* Springer, New York.
SHOHAT, J. and TAMARKIN, J. (1943). The Problem of Moments (1st ed.). *American Mathematical Society,* Providence, RI.
STOCKBRIDGE, R. H. (1990). Time-average control of martingale problems: A linear programming formulation. *Ann. Probab.* **18**, 206–217.