Statistica Sinica **10**(2000), 715-729

SEQUENTIAL DETECTION OF SIGNALS WITH KNOWN SHAPE AND UNKNOWN MAGNITUDE

M. Beibel

Albert–Ludwigs–Universität Freiburg

Abstract: We study a problem of sequential detection in a continuous time change-point model with a transition period. Let W denote a Brownian motion process which has zero drift during the time interval $[0,\nu)$ and drift $\theta h(t-\nu)$ during the time interval $[\nu,\infty)$. Here h is a known deterministic function and θ and ν are unknown parameters. The goal is to find a stopping time T of W that stops as soon and as reliably as possible after the change-point ν. We consider stopping rules based on mixtures of likelihoods and show that they are approximately Bayes optimal.

Key words and phrases: Bayes problem, change-point, sequential detection, transition period.

1. Introduction

Suppose one continuously monitors a process which initially is "in control". At some future time point this process may go "out of control" and it is then desirable to react. A convenient mathematical model for this situation is as follows. Let $(W_t; 0 \le t < \infty)$ denote a Brownian motion process with drift zero during the time interval $[0,\nu)$ and drift $\theta > 0$ during $[\nu,\infty)$ for some time point ν. Let $P_{(\theta,\nu)}$ denote the corresponding probability and $E_{(\theta,\nu)}$ the expectation with respect to $P_{(\theta,\nu)}$. Let P_∞ denote the probability when $\nu = \infty$ and E_∞ the expectation with respect to P_∞. Note that W is a standard Brownian motion under P_∞. The unknown time point ν is usually referred to as the "change-point" and one wants to detect it as fast as possible without raising a false alarm too frequently. The objectives "quick detection" and "low false alarm rate" are still informal and have to be specified further. Moreover the two goals are conflicting.

Shiryayev (1963) considered an exponential prior ρ for ν and studied the Bayes risk

$$\int_0^\infty \Big\{ P_\infty(T<\nu) + cE_{(\theta,\nu)}((T-\nu)^+) \Big\} \rho(d\nu)$$

for some $c > 0$. He proved that it is then optimal to stop as soon as the posterior probability of a change exceeds a certain level depending on c. Note that $P_\infty(T < \nu) = P_{(\theta,\nu)}(T < \nu)$. Pollak and Siegmund (1985) considered the minimization of $\sup_\nu E_{(\theta,\nu)}(T-\nu|T\geq\nu)$ under the constraint that $E_\infty(T)\geq A$ for some $A>0$. They studied in particular the performance of the so-called CUSUM and Shiryayev-Roberts procedures. See also Lai (1995), Siegmund (1994) and Zacks (1983) for an extensive overview on different optimality concepts for change-point detection problems.

One might think of situations where the assumption of an abrupt change is not appropriate. Discrete time change-point models with a transition period are for example considered by Huang and Chang (1993), Boukai and Zhou (1997) and Huskova (1999) in the context of nonparametric retrospective estimation of the change-point ν. We transfer their approach to continuous time sequential detection. The drift of W then does not change instantaneously but increases gradually from 0 to $\theta>0$. More precisely W is a Brownian motion process which has drift zero during the time interval $[0,\nu)$ and drift $\theta h(t-\nu)$ during $[\nu,\infty)$, where h is some nondecreasing function with $\lim_{t\to\infty}h(t)=1$.

It turns out that the monotonicity of h is not essential for our arguments. We therefore study the following slightly more general setup. Let h denote a given measurable function on $[0,\infty)$ such that $h(t)>0$ for $t\geq0$. Let $P_{(\theta,\nu)}$ denote the probability measure on $\sigma(W_s;0\leq s<\infty)$ under which W is a Brownian motion process with zero drift during the time interval $[0,\nu)$, and drift $\theta h(t-\nu)$ during $[\nu,\infty)$. This means that the process

$$W_t-\theta\int_{\nu\wedge t}^{t}h(u-\nu)du \tag{1}$$

is a standard Brownian motion under $P_{(\theta,\nu)}$. Let $E_{(\theta,\nu)}$ again denote the expectation with respect to $P_{(\theta,\nu)}$. Continuous time change-point models with a translated signal of known shape are suggested in Kolmogorov, Prokhorov and Shiryayev (1990). Note that for $h(t)=1_{[0,\infty)}(t)$ this model reduces to the one considered above.

We study the Bayes risk

$$L(c,T)=P_\infty(T<\infty)+c\int_0^\infty I(\theta)\int_0^\infty\left[E_{(\theta,\nu)}\left(\int_{\nu\wedge T}^{T}h^2(u-\nu)du\right)\right]\rho(d\nu)G(d\theta),$$

where ρ and G are probability distributions on $[0,\infty)$ and $(0,\infty)$ respectively, and $I(\theta)=\theta^2/2$. See Beibel (1997) for a related setup where $h(t)=1_{[0,\infty)}(t)$ and G is a normal distribution. The idea to put a constraint on $P_\infty(T<\infty)$ in order to control the rate of false alarms goes back to Pollak and Siegmund (1975). See also the discussion in Assaf, Pollak, Ritov and Yakir (1993) and Yakir (1996). Our costs for stopping late are proportional to the square of the current drift. The Kullback-Leibler information of $P_{(\theta,\nu)}$ with respect to P_∞ on $\sigma(W_s;0\leq s\leq t)$ is given by $I(\theta)\int_{\nu\wedge t}^{t}h^2(u-\nu)du$. This links our delay costs to the expected amount of incoming information. See also the discussion in Tartakovsky (1995). The basic idea lying behind this approach is that it should be easier to detect the change-point ν when the signal $\theta h(t-\nu)$ is large. This means that we standardize problems according to their difficulty. Our loss structure is also convenient from a mathematical point of view since it is closely related to the observed likelihood, see Proposition 3 below. It is not clear how to treat other delay costs in a similar way. See however Corollary 1 below for a partial result concerning the weighted delay $I(\theta)(T-\nu)^+$. For $h(t) = 1_{[0,\infty)}(t)$ our delay costs are the average weighted delay. The more general loss structure

$$P_\infty(T < \infty) + c\int_0^\infty C(\theta)\int_0^\infty \left[E_{(\theta,\nu)}\left(\int_{\nu\wedge T}^T h^2(u-\nu)du\right)\right]\rho(d\nu)G(d\theta)$$

for some positive weights $C(\theta)$ with $\int_\Omega [C(\theta)/I(\theta)]G(d\theta) < \infty$ may also be treated by introducing a prior distribution $\tilde{G}(d\theta)$ proportional to $[C(\theta)/I(\theta)]G(d\theta)$. All conditions on G then become conditions on $\tilde{G}$.

The goal is to minimize $L(c,T)$ over all stopping times T of W. We study the case $c \to 0$. When the costs for stopping late become small, one allows for more post-change observations. The asymptotic analysis of $L(c,.)$ for $c \to 0$ therefore is in essence a large sample asymptotic. We provide an asymptotic expansion of the minimal Bayes risk when the costs c become small. Moreover we show that certain mixture stopping rules are asymptotically optimal. This is the content of Theorem 1 and Theorem 2 below. We also examine the performance of non-optimal mixture stopping rules, see Theorem 3 below. This covers in particular the case of a (slightly) misspecified signal h. Our results show that the effect of the transition period does not appear in the leading terms $c\log(1/c)$ and $(c/2)\log\log(1/c)$. The effect of the transition period only appears in the constants $K(\rho, G, h)$ and $\Delta((h,\rho,G),(\tilde{h},\tilde{\rho},\tilde{G}))$ of Theorem 2 and Theorem 3 respectively. Hence this effect is of order $O(c)$ as $c \to 0$. Such a difference seems to be negligible from a practical point of view. Of particular interest is the case $\tilde{h}(t) = 1_{[0,\infty)}(t)$, $\tilde{\rho} = \rho$ and $\tilde{G} = G$ in Theorem 3.

Let $L_c^* = \inf_T L(c,T)$, where the infimum is taken over all stopping times T of W. Let $S_b = \inf\{t > 0 | \int_0^\infty\int_0^\infty e^{y\int_{s\wedge t}^t h(u-s)dW_u - \frac{y^2}{2}\int_{s\wedge t}^t h^2(u-s)du}\rho(ds)g(y)dy > b\}$ and $\beta(c) = 1/c$.

We impose the following conditions on h, ρ, and G:

(A1) $\int_0^t h^2(s)ds < \infty$ for all $t > 0$.

(A2) $\int_0^\infty h^2(s)ds = +\infty$.

(B) $\int_0^\infty\int_0^\infty [h(t) - h(t-s)1_{\{s\le t\}}]^2 dt\rho(ds) < \infty$.

(C1) The distribution G has an absolutely continuous Lebesgue density g on $[0,\infty)$ with $g(0) < \infty$.

(C2) $\int_0^\infty y^{2+\delta} g(y)dy < \infty$ for some $\delta > 0$.
(C3) $\int_0^\infty |H(y) \log |H(y)||g(y)dy < \infty$, where $H(y) = g'(y)/g(y)$.

A few words on our conditions are in order. Let B denote a standard Brownian motion. Then the distribution of $B_s + \theta \int_0^s h(u)du$ is equivalent to the distribution of B_s for $0 \le s \le t$ if and only if $\int_0^t h^2(u)du < \infty$. Condition (A1) therefore implies that the measures $P_{(\theta,0)}$ and P_∞ are equivalent on $\sigma(W_s; 0 \le s \le t)$ for all $0 \le t < \infty$. Condition (A2) implies that $P_{(\theta,0)}$ and P_∞ are orthogonal on $\sigma(W_s; 0 \le s < \infty)$. The Kullback-Leibler information of $P_{(\theta,\nu)}$ with respect to $P_{(\theta,0)}$ on $\sigma(W_s; 0 \le s < \infty)$ is given by

$$E_{(\theta,\nu)} \left(\log \frac{dP_{(\theta,\nu)}}{dP_{(\theta,0)}} \Big|_{\sigma(W_s; 0 \le s < \infty)} \right) = I(\theta) \int_0^\infty [h(t) - h(t-\nu) 1_{\{\nu \le t\}}]^2 dt.$$

So, condition (B) implies that the averaged Kullback-Leibler information

$$\int_0^\infty E_{(\theta,\nu)} \left(\log \frac{dP_{(\theta,\nu)}}{dP_{(\theta,0)}} \Big|_{\sigma(W_s; 0 \le s < \infty)} \right) \rho(d\nu)$$

is finite. If h is nondecreasing, then

$$\int_0^\infty [h(t) - h(t-s) 1_{\{s \le t\}}]^2 dt \le s \lim_{t \to \infty} h^2(t).$$

Condition (B) is therefore satisfied whenever h is nondecreasing with $\lim_{t\to\infty} h(t) < \infty$ and $\int_0^\infty s\rho(ds) < \infty$. Note also that the Conditions (A1) and (A2) are trivially satisfied whenever h is nondecreasing with $\lim_{t\to\infty} h(t) < \infty$. The conditions (C1) to (C3) are of a more technical nature and are needed in an approximation argument below, see Proposition 5 in the Appendix. Note however that condition (C2) implies $\int_0^\infty y^2 g(y)dy < \infty$ and so provides, together with condition (B), the finiteness of

$$\int_0^\infty \int_0^\infty \left[E_{(\theta,\nu)} \left(\log \frac{dP_{(\theta,\nu)}}{dP_{(\theta,0)}} \Big|_{\sigma(W_s; 0 \le s < \infty)} \right) \right] \rho(d\nu) g(\theta) d\theta.$$

Theorem 1. *Let h, ρ, and G satisfy the conditions* (A1), (A2), (B), *and* (C1) *to* (C3). *Then $L_c^* = L(c, S_{\beta(c)}) + o(c)$ when $c \to 0$.*

Theorem 2. *Let h, ρ, and G satisfy the conditions* (A1), (A2), (B), *and* (C1) *to* (C3). *Then $L(c, S_{\beta(c)}) = c[\log(1/c) + \frac{1}{2} \log\log(1/c) + K(\rho, G, h)] + o(c)$ when $c \to 0$ for some constant $K(\rho, G, h)$.*

See Remark 3 in Section 3 for a more explicit expression for K.

Let $\tilde{L}(c,T) = P_\infty(T < \infty) + c \int_0^\infty I(\theta) \int_0^\infty [E_{(\theta,\nu)}(T-\nu)^+] \rho(d\nu) G(d\theta)$. Then obviously $\tilde{L}(c,T) = L(c,T) + c \int_0^\infty I(\theta) \int_0^\infty [E_{(\theta,\nu)}(\int_{\nu\wedge T}^T [1 - h^2(u-\nu)]du)] \rho(d\nu) G(d\theta)$. The asymptotic results of Theorem 1 and Theorem 2 therefore immediately yield

Corollary 1. *Let h, ρ, and G satisfy the conditions* (B) *and* (C1) *to* (C3). *If in addition $0 < h \le 1$ and $\int_0^\infty [1-h(u)^2]du < \infty$, then*

$$\inf_T \tilde{L}(c,T) = L_c^* + c\int_0^\infty \frac{y^2}{2}G(dy)\int_0^\infty [1-h^2(u)]du + o(c)$$
$$= L(c, S_{\beta(c)}) + c\int_0^\infty \frac{y^2}{2}G(dy)\int_0^\infty [1-h^2(u)]du + o(c)$$

as $c \to 0$.

Let $\tilde{h}$ denote a positive function on $[0,\infty)$. Let $\tilde{P}_{(\theta,\nu)}$ for $\theta \in (0,\infty)$ and $\nu \in [0,\infty)$ denote the probability measure on $\sigma(W_s; 0 \le s < \infty)$ under which the process $W_t - \theta\int_0^t \tilde{h}(u-\nu)1_{\{\nu \le u\}}du$ is a standard Brownian motion. Let $\tilde{\rho}$ and $\tilde{G}$ be probability distributions on $[0,\infty)$. Let

$$\tilde{S}_b = \inf\left\{t > 0 | \int_0^\infty \int_0^\infty e^{y\int_{s\wedge t}^t \tilde{h}(u-s)dW_u - \frac{y^2}{2}\int_{s\wedge t}^t \tilde{h}^2(u-s)du}\tilde{\rho}(ds)\tilde{g}(y)dy > b\right\}.$$

Theorem 3. *Let (h,ρ,G) and $(\tilde{h},\tilde{\rho},\tilde{G})$ both satisfy* (A1), (A2), (B), *and* (C1) *to* (C3). *Let $\int_0^\infty [h(u)-\tilde{h}(u)]^2 du < \infty$ and suppose that $\tilde{\rho}$ and $\tilde{G}$ are dominated by ρ and G respectively with $\int_0^\infty \log(d\rho(u)/d\tilde{\rho}(u))\rho(du) < \infty$ and $\int_0^\infty \log(g(y)/\tilde{g}(y))$ $g(y)dy < \infty$. Then $L(c,\tilde{S}_{\beta(c)}) = L_c^* + c\Delta((h,\rho,G),(\tilde{h},\tilde{\rho},\tilde{G})) + o(c)$ when $c \to 0$, where $\Delta((h,\rho,G),(\tilde{h},\tilde{\rho},\tilde{G})) = E_P[\log(dP/d\tilde{P}|\sigma(W_s; 0 \le s < \infty))]$ equals the Kullback-Leibler information of $\tilde{P} = \int_0^\infty\int_0^\infty \tilde{P}_{(\theta,\nu)}\tilde{\rho}(d\nu)\tilde{G}(d\theta)$ relative to $P = \int_0^\infty\int_0^\infty P_{(\theta,\nu)}\rho(d\nu)G(d\theta)$ on $\sigma(W_s; 0 \le s < \infty)$.*

Remark 1. If $\tilde{\rho} = \rho$ and $\tilde{G} = G$, then $\Delta \le \int_0^\infty I(y)g(y)dy \int_0^\infty [\tilde{h}(u) - h(u)]^2 du$.

The rest of this paper is organized as follows. In Section 2 we construct a suitable model and establish some key facts. Section 3 contains the proof of Theorem 2. Theorem 1 is then proved in Section 4 and Theorem 3 is proved in Section 5. Most of our arguments are analogous to those of Beibel (1997). The Appendix contains an approximation argument which we need in the proof of Proposition 3 in Section 3. These arguments extend some of the results of Paulsen (1999) to our setting.

2. Some Results on the Structure

We assume throughout this article that the conditions (A1), (A2), (B), and (C1) to (C3) are satisfied. Let B denote a standard Brownian motion. Let Y be a positive random variable with $P(Y \le y) = \int_0^y g(s)ds$ and τ be a nonnegative random variable with $P(\tau > t) = \int_{(t,\infty)} \rho(ds)$ for all $t \ge 0$. Let B, Y and τ be independent under P. Put $W_t = B_t + \int_0^t R_s ds$, where $R_s = Yh(s-\tau)1_{\{\tau \le s\}}$.

Let $\mathcal{F}_t = \mathcal{F}_t^W = \sigma(W_s; 0 \le s \le t)$ and $\mathcal{F}_t^{B,Y,\tau} = \sigma(B_s; 0 \le s \le t, Y, \tau)$. Let P_∞ denote the probability measure on $\sigma(B_s, 0 \le s < \infty, Y, \tau)$ given by

$$\frac{dP_\infty}{dP}\Big|_{\mathcal{F}_t^{B,Y,\tau}} = \exp\left\{-Y\int_{\tau\wedge t}^t h(s-\tau)dB_s - \frac{Y^2}{2}\int_{\tau\wedge t}^t h^2(s-\tau)ds\right\}$$

for all $0 \le t < \infty$. Obviously the P_∞-distribution of (Y,τ) coincides with the distribution of (Y,τ) under P. Moreover $(W_t, \mathcal{F}_t; 0 \le t < \infty)$ is a standard Brownian motion and independent of (Y,τ) under P_∞. This implies

$$\frac{dP}{dP_\infty}\Big|_{\mathcal{F}_t} = \int_0^\infty \int_0^\infty e^{y\int_{s\wedge t}^t h(u-s)dW_u - \frac{y^2}{2}\int_{s\wedge t}^t h(u-s)^2du}\rho(ds)g(y)dy.$$

Let $\psi_t = \frac{dP}{dP_\infty}|_{\mathcal{F}_t}$. Then

$$L(c,T) = P_\infty(T<\infty) + cE\left(\int_0^T R_s^2 ds\right) \ . \tag{2}$$

We first derive a stochastic differential equation for $\log\psi_t$. Let $\widehat{R}_t = E(R_t|\mathcal{F}_t)$ and $\widehat{R_t^2} = E(R_t^2|\mathcal{F}_t)$. Let $\overline{W}$ denote the innovation process $\overline{W_t} := W_t - \int_0^t \hat{R}_s ds$. This process is a standard Brownian motion under the probability measure P relative to the filtration $\mathcal{F}$ (see Liptser and Shiryayev (1977, pp.297-299)). Theorem 7.13 in Liptser and Shiryayev (1977) yields that

Proposition 1.

$$d\log\psi_t = \frac{1}{2}(\hat{R}_t)^2 dt + \hat{R}_t d\overline{W_t} \ . \tag{3}$$

We need one more probability measure. Let P_0 denote the probability measure on $\sigma(B_s, 0 \le s < \infty, Y, \tau)$ given by

$$\frac{dP_0}{dP}\Big|_{\sigma(B_s; 0\le s<\infty,\tau,Y)} = \exp\Big\{Y\int_0^\infty [h(t) - h(t-\tau)1_{\{\tau\le t\}}]dB_t - \frac{Y^2}{2}\int_0^\infty [h(t)-h(t-\tau)1_{\{\tau\le t\}}]^2 dt\Big\}.$$

Note that $\int_0^\infty [h(t)-h(t-\tau)1_{\{\tau\le t\}}]^2dt < \infty$ holds with P-probability one. The process $(W_t - Y\int_0^t h(s)ds; 0 \le t < \infty)$ is a standard Brownian motion under P_0 with respect to the filtration $\mathcal{F}^{B,Y,\tau}$. Moreover $(W_t - Y\int_0^t h(s)ds; 0 \le t < \infty)$ is independent of (Y,τ) under P_0. Therefore the distribution of W under P_0 is given by $\int_0^\infty P_{(\theta,0)}g(\theta)d\theta$. The probability measures P and P_0 are equivalent on the σ-algebra $\sigma(W_s; 0 \le s < \infty, Y, \tau)$. The following remark summarizes the essential features of our setup.

Remark 2. Let $P^{(W,Y,\tau)}$, $P_\infty^{(W,Y,\tau)}$ and $P_0^{(W,Y,\tau)}$ denote the distribution of (W,Y,τ) under P, P_∞ and P_0, respectively, $C[0,\infty)$ the space of all real-valued continuous functions on $[0,\infty)$ vanishing at zero, μ_W the Wiener measure on $C[0,\infty)$, and $\mathcal{T}_1$ and $\mathcal{T}_2$ denote the transformations from $C[0,\infty)\times(0,\infty)\times[0,\infty)$ to itself given by $\mathcal{T}_1(f(.),\theta,\nu) = (f(.) - \theta\int_0^{\cdot} h(s)ds,\theta,\nu)$ and $\mathcal{T}_2(f(.),\theta,\nu) = (f(.)-\theta\int_0^{\cdot} h(s-\nu)1_{\{\nu\le s\}}ds,\theta,\nu)$. Then $P_\infty^{(W,Y,\tau)} = P_0^{(W,Y,\tau)}\mathcal{T}_1^{-1} = P^{(W,Y,\tau)}\mathcal{T}_2^{-1} = \mu_W\otimes G\otimes\rho$.

Let $N_t = dP/dP_0|\sigma(W_s; 0\le s\le t)$ and $N_\infty = dP/dP_0|\sigma(W_s; 0\le s<\infty)$. We have

$$\lim_{t\to\infty}\Big[\int_0^t h(s)dW_s\Big/\int_0^t h^2(s)ds\Big] = Y \tag{4}$$

with P_0-probability one.

Proposition 2. $P(\lim_{t\to\infty}[\log\psi_t/\int_0^t R_s^2 ds] = 1/2) = 1$.

Proof. We have $\log\psi_t = \log N_t + \log\psi_t^{(0)}$, where $\psi_t^{(0)} = dP_0/dP_\infty|_{\mathcal{F}_t}$. Theorem 7.13 in Liptser and Shiryayev (1977) provides

$$\log\psi_t^{(0)} = \int_0^t E_0(Y|\mathcal{F}_s)h(s)dW_s - \frac{1}{2}\int_0^t [E_0(Y|\mathcal{F}_s)]^2 h^2(s)ds. \tag{5}$$

Clearly (4) implies $P_0(\lim_{t\to\infty} E_0(Y|\mathcal{F}_t) = Y) = 1$ and so $P_0(\lim_{t\to\infty}[\log\psi_t^{(0)}/\int_0^t h^2(s)ds] = Y^2/2) = 1$. The P_0-martingale N_t converges under P_0 with probability one to N_∞. Therefore $P_0(\lim_{t\to\infty}[\log N_t/\int_0^t h^2(s)ds] = 0) = 1$. We have $\int_{t-\tau}^t h^2(u)du \le 2\int_0^t h(u)[h(u)-h(u-\tau)1_{\{\tau\le u\}}]du$ for $t\ge\tau$. Condition (B) therefore yields $\lim_{t\to\infty}[\int_\tau^t h^2(s-\tau)ds/\int_0^t h^2(s)ds] = 1$ with probability one. The assertion now follows since P and P_0 are equivalent on $\sigma(W_s, 0\le s<\infty)$.

We need some further notation. Let $\mathcal{F}_t^{W,\tau}$ denote the σ-algebra $\sigma(W_s; 0\le s\le t,\tau)$. Let $\widehat{R}_t^{(\tau)} = E(R_t|\mathcal{F}_t^{W,\tau})$ and $\widehat{R_t^2}^{(\tau)} = E(R_t^2|\mathcal{F}_t^{W,\tau})$. Theorem 7.13 in Liptser and Shiryayev (1977) yields (by conditioning on τ)

$$\frac{dP}{dP_\infty}\bigg|_{\mathcal{F}_t^{W,\tau}} = \exp\left\{\int_0^t \widehat{R}_s^{(\tau)}dW_s - (1/2)\int_0^t[\widehat{R}_s^{(\tau)}]^2 ds\right\}. \tag{6}$$

Let $N_t^{(\tau)} = dP/dP_0|\mathcal{F}_t^{W,\tau}$ and $N_\infty^{(\tau)} = dP/dP_0|\sigma(W_s; 0\le s<\infty,\tau)$. Then (4) implies

$$N_\infty^{(\tau)} = \frac{dP}{dP_0}\bigg|_{\sigma(W_s;0\le s<\infty,\tau,Y)} \tag{7}$$

with P_0-probability one, and so

$$E(\log N_\infty^{(\tau)}) = (1/2)E(Y^2)E\left(\int_0^\infty [h(t)-h(t-\tau)1_{\{\tau\le t\}}]^2 dt\right) < \infty.$$

Moreover $N_\infty = dP/dP_0|\sigma(W_s; 0 \le s < \infty, Y)$. Since $\sigma(W_s; 0 \le s < \infty) \subset \sigma(W_s; 0 \le s < \infty, \tau)$ we obtain $0 \le E(\log N_\infty) \le E(\log N_\infty^{(\tau)}) < \infty$. Let $\mathbb{V}_t^{(\tau)} := \widehat{R_t^2}^{(\tau)} - [\hat{R}_t^{(\tau)}]^2$ and $\tilde{\mathbb{V}}_t^{(\tau)} := h^2(t-\tau)1_{\{\tau \le t\}}/[\int_{t\wedge\tau}^t h^2(u-\tau)du + 1]$. We have $E(\int_0^T \tilde{\mathbb{V}}_s^{(\tau)} ds) = E[\log(\int_{\tau\wedge T}^T h^2(t-\tau)dt + 1)]$.

Proposition 3. *For all stopping times T with $E(\int_0^T R_s^2 ds) < \infty$, $L(c,T) = E[\psi_T^{-1} + \frac{c}{2}\int_0^T \widehat{R^2}_s ds]$ and*

$$\begin{aligned}\frac{1}{2}E\left(\int_0^T \widehat{R_s^2} ds\right) &= E(\log\psi_T) + \frac{1}{2}E\left(\int_0^T (\widehat{R_s^2} - [\hat{R}_s]^2) ds\right)\\ &= E(\log\psi_T) + \frac{1}{2}E\left[\log\Big(\int_{\tau\wedge T}^T h^2(s-\tau)ds + 1\Big)\right]\\ &\quad + \frac{1}{2}E\left(\int_0^T [\mathbb{V}_s^{(\tau)} - \tilde{\mathbb{V}}_s^{(\tau)}] ds\right) + \frac{1}{2}E\left(\int_0^T \left[\hat{R}_s^{(\tau)} - \hat{R}_s\right]^2 ds\right).\end{aligned} \tag{8}$$

Moreover $\left|E\left(\int_0^T \left[\mathbb{V}_s^{(\tau)} - \tilde{\mathbb{V}}_s^{(\tau)}\right] ds\right)\right| \le E\left(\int_\tau^\infty \left|\mathbb{V}_s^{(\tau)} - \tilde{\mathbb{V}}_s^{(\tau)}\right| ds\right) < \infty$ *and*

$$\begin{aligned}E\left(\int_0^T (\hat{R}_s^{(\tau)} - \hat{R}_s)^2 ds\right) &\le E\left(\int_0^\infty (\hat{R}_s^{(\tau)} - \hat{R}_s)^2 ds\right)\\ &= 2\left[E\left(\log N_\infty^{(\tau)}\right) - E\left(\log N_\infty\right)\right] < \infty.\end{aligned}$$

Proof. Fubini's theorem, $P_\infty(T < \infty) = E(1/\psi_T 1_{\{T<\infty\}})$ and Proposition 1 yield, for all $\mathcal{F}^W$-stopping times T with $L(c,T) < \infty$,

$$L(c,T) = E\left\{\psi_T^{-1} + \frac{c}{2}\int_0^T \widehat{R_s^2} ds\right\} = E\left\{\psi_T^{-1} + c\log\psi_T\right\} + \frac{c}{2}E\left(\int_0^T (\widehat{R_s^2} - [\hat{R}_s]^2) ds\right). \tag{9}$$

Another Fubini type of argument provides

$$E\left(\int_0^T (\widehat{R_s^2} - [\hat{R}_s]^2) ds\right) = E\left(\int_0^T [\widehat{R_s^2}^{(\tau)} - (\hat{R}_s^{(\tau)})^2] ds\right) + E\left(\int_0^T \left(\hat{R}_s^{(\tau)} - \hat{R}_s\right)^2 ds\right). \tag{10}$$

This yields (8). Proposition 5 in the Appendix provides $E(\int_\tau^{+\infty} |\mathbb{V}_s^{(\tau)} - \tilde{\mathbb{V}}_s^{(\tau)}| ds) < +\infty$. A third Fubini type of argument, together with (5) and (6) for $0 \le t < \infty$, yields

$$\begin{aligned}E\left(\int_0^t \left(\hat{R}_s^{(\tau)} - \hat{R}_s\right)^2 ds\right) &= E\left(\int_0^t \left(\hat{R}_s^{(\tau)} - E_0(Y|\mathcal{F}_s)h(s)\right)^2 ds\right)\\ &\quad - E\left(\int_0^t \left(\hat{R}_s - E_0(Y|\mathcal{F}_s)h(s)\right)^2 ds\right)\end{aligned} \tag{11}$$

$$= 2\left[E\Big(\log N_t^{(\tau)}\Big) - E\big(\log N_t\big)\right]. \tag{12}$$

Note that $E(\hat{R}_s^{(\tau)}|\mathcal{F}_s) = \hat{R}_s$. The submartingales $\log N_t^{(\tau)}$ and $\log N_t$ are uniformly integrable. The last assertion follows if we let $t \to \infty$.

3. Proof of Theorem 2

We recall that $S_b = \inf\{t > 0 | \psi_t \geq b\}$. We have $P(S_b < \infty) = 1$ for all $b > 1$ since the probability measures P_∞ and P are orthogonal on $\sigma(W_s, 0 \leq s < \infty)$. Obviously $P\left(\lim_{b\to\infty} S_b = +\infty\right) = 1$. Proposition 2 therefore yields

$$P\left(\lim_{b\to\infty}\left[\frac{1}{2}\int_0^{S_b} R_s^2 ds\Big/\log b\right] = 1\right) = 1. \tag{13}$$

Proposition 3 yields

$$E\left(\int_0^{S_b} R_s^2 ds\right) \leq 2\log b + E\left[\log\left(\int_{\tau\wedge S_b}^{S_b} h^2(u-\tau)du + 1\right)\right] + A,$$

for some constant A. The same arguments as in the proof of Lemma 3 of Beibel (1997) therefore imply

$$E\left(\int_0^{S_b} R_s^2 ds\right) \leq 4\log b + O(1), \tag{14}$$

as $b \to \infty$. Let $\xi(b, y, s) = \inf\{t > 0 | y^2 \int_{s\wedge t}^t h^2(u-s)du \geq \eta_b\}$, where $\eta_b = (1/3)\log b$. Obviously $\{\int_0^{S_b} R_s^2 ds \leq \eta_b\} = \{S_b \leq \xi(b, Y, \tau)\}$. It holds that

$$P\Big(\int_0^{S_b} R_s^2 ds \leq \eta_b\Big) \leq b^{-\frac{2}{3}} E_\infty\bigg(\frac{dP}{dP_\infty}\bigg|_{\mathcal{F}^{B,Y,\tau}_{\xi(b,Y,\tau)}}\bigg)^2 + b^{\frac{2}{3}} P_\infty\Big(\int_0^{S_b} R_s^2 ds \leq \eta_b\Big). \tag{15}$$

The independence of S_b, Y and τ under P_∞ yields

$$P_\infty\Big(\int_0^{S_b} R_s^2 ds \leq \eta_b\Big) = \int_0^\infty\int_0^\infty P_\infty\Big(S_b \leq \xi(b, y, s)\Big) g(y)dy\rho(ds) \leq 1/b. \tag{16}$$

Now $\int_0^{\xi(b,Y,\tau)} R_s^2 ds = \eta_b$. So,

$$E_\infty\bigg(\frac{dP}{dP_\infty}\bigg|_{\mathcal{F}^{B,Y,\tau}_{\xi(b,Y,\tau)}}\bigg)^2 \leq b^{\frac{1}{3}} E_\infty\left(e^{\int_0^{\xi(b,Y,\tau)}[2R_s]dW_s - \frac{1}{2}\int_0^{\xi(b,Y,\tau)}[2R_s]^2 ds}\right) \leq b^{\frac{1}{3}}. \tag{17}$$

It follows from (15), (16), and (17) that

$$P\bigg(\int_0^{S_b} R_s^2 ds \leq \tfrac{1}{3}\log b\bigg) \leq 2b^{-1/3}\ . \tag{18}$$

Inequalities (14) and (18) yield the uniform integrability of $\log(\int_0^{S_b} R_s^2 ds \big/ \log b)$; see the proof of Proposition 4 in Beibel (1997). Combining (13), (14) and (18) we arrive at the following Proposition.

Proposition 4. *$E(\log[\int_{\tau\wedge S_b}^{S_b} h^2(u-\tau)du+1]) = \log\log b - \int_0^\infty \log(I(y))\, g(y)dy + o(1)$, as $b \to \infty$.*

Proof of Theorem 2. Proposition 3 has $L(c, S_{\beta(c)}) = E\{\psi_{S_{\beta(c)}}^{-1} + c\log(\psi_{S_{\beta(c)}}) + \frac{c}{2}\int_0^{S_{\beta(c)}} [\widehat{R^2}_s - (\widehat{R}_s)^2]ds\}$. For sufficiently small c it holds that $E(1/\psi_{S_{\beta(c)}} + c\log\psi_{S_{\beta(c)}}) = c + c\log(1/c)$. Proposition 3 and Proposition 4 yield the assertion.

Remark 3. We have $K(\rho, G, h) = 1 - (1/2)\int_0^\infty \log(I(y))g(y)dy + E(\int_\tau^\infty [\mathbb{V}_s^{(\tau)} - \tilde{\mathbb{V}}_s^{(\tau)}]ds) + E(\log N_\infty^{(\tau)}) - E(\log N_\infty)$ with $E(\log N_\infty^{(\tau)}) = \int_0^\infty I(y)g(dy)\int_0^\infty(\int_0^\infty [h(t) - h(t-s)1_{\{s\le t\}}]^2 dt)\rho(ds)$. If $\lim_{t\to\infty} h(t)$ exists and $\lim_{t\to\infty} h(t) = 1$, then

$$E\Big(\log N_\infty\Big) = \frac{1}{2}\int_0^\infty \int_0^\infty E_{(\theta,\nu)}\Big[\log\Big(\int_0^\infty \exp\Big\{-\theta\int_0^\infty [h(t) - h(t-s)1_{\{s\le t\}}]dW_t + I(\theta)s\Big\}\rho(ds)\Big)\Big] g(\theta)d\theta\rho(d\nu).$$

4. Proof of Theorem 1

We compare the performance of the stopping times $S_{\beta(c)}$ with the performance of c^2-optimal solutions. Let T_c for $0 < c \le 1$ be an c^2-optimal stopping rule, that is a stopping time with $L(c, T_c) \le L_c^* + c^2$. We may assume without loss of generality that $T_c \le S_{\beta(c)}$. Theorem 2 yields $c^{-1}(L_c^* - c\log(1/c)) \le (1/2)\log\log(1/c) + O(1)$ as $c \to 0$. Similar arguments as in the proof of Lemma 5 and Lemma 6 of Beibel (1997) therefore provide, as $c \to 0$,

$$E\left(\log\frac{\int_{\tau\wedge S_{\beta(c)}}^{S_{\beta(c)}} h^2(s-\tau)ds + 1}{\int_{\tau\wedge T_c}^{T_c} h^2(s-\tau)ds + 1}\right) \to 0. \tag{19}$$

Moreover $T_c \to \infty$ in P-probability. The function $g_c(x) = 1/x + c\log x$ assumes its unique minimum over the interval $(0, 1)$ at $x = \beta(c)$. Proposition 3 therefore provides

$$\begin{aligned} 0 \le L(c, S_{\beta(c)}) - L_c^* &\le L(c, S_{\beta(c)}) - L(c, T_c) + c^2 \\ &\le cE\left(\int_{T_c}^\infty \Big[\hat{R}_s^{(\tau)} - \hat{R}_s\Big]^2 ds\right) + cE\left(\int_{T_c}^\infty \Big[\mathbb{V}_s^{(\tau)} - \tilde{\mathbb{V}}_s^{(\tau)}\Big]ds\right) \\ &\quad + cE\left(\log\frac{\int_{\tau\wedge S_{\beta(c)}}^{S_{\beta(c)}} h^2(s-\tau)ds + 1}{\int_{\tau\wedge T_c}^{T_c} h^2(s-\tau)ds + 1}\right) + c^2. \end{aligned}$$

Proposition 3 and (19) now yield Theorem 1.

5. Proof of Theorem 3

Let $\tilde{P}$ denote the probability measure on $\sigma(B_s; 0 \le s < \infty, Y, \tau)$ given by

$$\frac{d\tilde{P}}{dP}\bigg|_{\sigma(B_s;0\le s\le t,Y,\tau)} = \exp\left\{Y\int_{\tau\wedge t}^{t}[\tilde{h}(u-\tau)-h(u-\tau)]dB_u - \frac{Y^2}{2}\int_{\tau\wedge t}^{t}[\tilde{h}(u-\tau)-h(u-\tau)]^2du\right\}\frac{d\tilde{\rho}(\tau)}{d\rho(\tau)}\frac{\tilde{g}(Y)}{g(Y)}$$

for $0 \le t < \infty$. Then

$$\frac{d\tilde{P}}{dP}\bigg|_{\sigma(B_s;0\le s<\infty,Y,\tau)} = \exp\left\{Y\int_{\tau}^{\infty}[\tilde{h}(u-\tau)-h(u-\tau)]dB_u - \frac{Y^2}{2}\int_{\tau}^{\infty}[\tilde{h}(u-\tau)-h(u-\tau)]^2du\right\}\frac{d\tilde{\rho}(\tau)}{d\rho(\tau)}\frac{\tilde{g}(Y)}{g(Y)}.$$

Under $\tilde{P}$ the process $\tilde{B}_t = B_t - Y\int_{\tau\wedge t}^{t}[\tilde{h}(u-\tau)-h(u-\tau)]du = W_t - Y\int_{\tau\wedge t}^{t}\tilde{h}(u-\tau)du$ is a standard Brownian motion. Let $\tilde{P}_\infty$ denote the probability measure on $\sigma(\tilde{B}_s, 0 \le s < \infty, Y, \tau) = \sigma(B_s, 0 \le s < \infty, Y, \tau)$ given by $dP_\infty/dP|\mathcal{F}_t^{B,Y,\tau} = \exp\{-Y\int_{\tau\wedge t}^{t}\tilde{h}(s-\tau)d\tilde{B}_s - (Y^2/2)\int_{\tau\wedge t}^{t}\tilde{h}^2(s-\tau)ds\}$ for all $0 \le t < \infty$. Then $(W_t, \mathcal{F}_t; 0 \le t < \infty)$ is a standard Brownian motion with respect to $\tilde{P}_\infty$. Therefore $\tilde{P}_\infty$ and P_∞ coincide on $\sigma(W_s; 0 \le s < \infty)$. Moreover $d\tilde{P}_\infty/dP_\infty|\sigma(W_s; 0 \le s < \infty, Y, \tau) = (d\tilde{\rho}(\tau)/d\rho(\tau))(\tilde{g}(Y)/g(Y))$. So, W is independent of (Y, τ) under $\tilde{P}_\infty$. Let

$$\tilde{\psi}_t = \frac{d\tilde{P}}{d\tilde{P}_\infty}\bigg|_{\mathcal{F}_t} = \frac{d\tilde{P}}{dP_\infty}\bigg|_{\mathcal{F}_t} = \int_0^\infty\int_0^\infty e^{y\int_{s\wedge t}^{t}\tilde{h}(u-s)dW_u - \frac{y^2}{2}\int_{s\wedge t}^{t}\tilde{h}^2(u-s)du}\tilde{\rho}(ds)\tilde{g}(y)dy.$$

Now

$$\log\tilde{\psi}_t = \log\psi_t - \log L_t, \tag{20}$$

where $L_t = dP/d\tilde{P}|\mathcal{F}_t$. The submartingale $\log L_t$ is uniformly integrable with respect to P. Moreover L_t converges, as $t \to \infty$, to $L_\infty = dP/d\tilde{P}|\sigma(W_s; 0 \le s < \infty)$. Obviously $P_\infty(\tilde{S}_b < \infty) = 1/b$ for $b \ge 1$. Proposition 3 yields, for $b \ge 1$,

$$E\left(\int_0^{\tilde{S}_b}R_s^2ds\right) = 2\log b + E(\log L_{\tilde{S}_b}) + E\left[\log\Big(\int_{\tau\wedge\tilde{S}_b}^{\tilde{S}_b}h^2(s-\tau)ds+1\Big)\right] + E\left(\int_0^{\tilde{S}_b}[\mathbb{V}_s^{(\tau)}-\tilde{\mathbb{V}}_s^{(\tau)}]ds\right) + E\left(\int_0^{\tilde{S}_b}\left[\hat{R}_s^{(\tau)}-\hat{R}_s\right]^2ds\right).$$

We have $P(\lim_{b\to\infty}\tilde{S}_b = \infty) = 1$ and $\lim_{b\to\infty} E(\log L_{\tilde{S}_b}) = E(\log L_\infty) = \Delta((h,\rho,G),(\tilde{h},\tilde{\rho},\tilde{G}))$. Theorem 3 therefore follows if we show that

$$E\left[\log\Big(\int_{\tau\wedge\tilde{S}_b}^{\tilde{S}_b} h^2(u-\tau)du+1\Big)\right] = \log\log b - \int_0^\infty \log\left(I(y)\right) g(y)dy + o(1) \quad (21)$$

when $b\to\infty$. Similar arguments as in the proof of Proposition 4 give (21) if we show that

i) $P(\lim_{b\to\infty}[\int_0^{\tilde{S}_b} R_s^2 ds/2\log b] = 1) = 1$.

ii) $E(\int_0^{\tilde{S}_b} R_s^2 ds) \le 2\log b + O(1)$.

iii) $P(\int_0^{\tilde{S}_b} R_s^2 ds \le (1/3)\log b) \le 2b^{-1/3}$.

Proposition 2 and (20) yield $P(\lim_{t\to\infty}[\log\tilde{\psi}_t/\int_0^t R_s^2 ds] = 1/2) = 1$, and thus i). From (20) with $b \ge 1$, $E(\log \psi_{\tilde{S}_b}) \le b - E(\log L_{\tilde{S}_b})$. Similar arguments as in Section 3 therefore provide ii). The same arguments as in the proof of (18) in Section 3, with S_b replaced by $\tilde{S}_b$, provide iii).

Acknowledgements

The author is grateful to anonymous referees and the associate editor for suggestions which helped to improve the presentation of the paper.

Appendix. The Integrated Posterior Variance

Following the lines of Paulsen (1999) we will now approximate $\mathbb{V}_t^{(\tau)} = \widehat{R_t^2}^{(\tau)} - [\hat{R}_t^{(\tau)}]^2$ by $\tilde{\mathbb{V}}_t^{(\tau)} = h^2(t-\tau)/[\int_{t\wedge\tau}^t h^2(u-\tau)du+1]$. Proposition 5 below transfers the results of Proposition 3.3 and Lemma 4.1 in Paulsen (1999) to our setup. Note that we allow for $g(0) > 0$. Then g is not left-continuous at 0. This leads to the additional term $g(0)(\psi_t^{(\tau)})^{-1}$ in (25) below, which is not a closable martingale. This complicates our arguments further. For $\tau \le t < \infty$, let $\widehat{Y}_t^{(\tau)} = E(Y|\mathcal{F}_t^{W,\tau})$ and $\widehat{Y_t^2}^{(\tau)} = E(Y^2|\mathcal{F}_t^{W,\tau})$. If $\tau \le t < \infty$, $\hat{R}_t^{(\tau)} = h(t-\tau)\widehat{Y}_t^{(\tau)}$ and $\widehat{R_t^2}^{(\tau)} = h^2(t-\tau)\widehat{Y_t^2}^{(\tau)}$. Let $\mathbf{v}_t^{(\tau)} = \widehat{Y^2}_t^{(\tau)} - [\widehat{Y}_t^{(\tau)}]^2$. Then $\mathbb{V}_t^{(\tau)} = h^2(t-\tau)\mathbf{v}_t^{(\tau)}$. Let $I_t^{(\tau)} = \int_{\tau\wedge t}^t h^2(u-\tau)du = \int_0^{(t-\tau)^+} h^2(u)du$.

Proposition 5. $E(\int_\tau^\infty |\mathbb{V}_s^{(\tau)} - \tilde{\mathbb{V}}_s^{(\tau)}|ds) < +\infty$.

Proof. Let $\xi^{(\tau)}(v) = \inf\{t \ge \tau | I_t^{(\tau)} = v\} = \tau + \inf\{t \ge 0 | \int_0^t h^2(u)du = v\}$. Then

$$\int_\tau^\infty \left|\mathbb{V}_s^{(\tau)} - \tilde{\mathbb{V}}_s^{(\tau)}\right| ds = \int_\tau^\infty \left|\bar{\mathbf{v}}_s - \frac{1}{1+s}\right| ds,$$

where $\bar{\mathbf{v}}_s = \mathbf{v}^{(\tau)}_{\xi^{(\tau)}(s)}$. Let $\alpha \in (0,2)$. Then

$$\int_0^\infty \left|\bar{\mathbf{v}}_s - \frac{1}{s+1}\right| ds \le \left(\sup_{0\le s<\infty}\left|\bar{\mathbf{v}}_s - \frac{1}{s+1}\right|^{1-\alpha}\right)\int_0^\infty \left|\bar{\mathbf{v}}_s - \frac{1}{s+1}\right|^\alpha ds. \quad (22)$$

The Hölder inequality with $p = 2/(2-\alpha)$ and $q = 2/\alpha$ now yields

$$\int_0^\infty \left|\bar{\mathbf{v}}_s - \frac{1}{1+s}\right|^\alpha ds \le C(\alpha)^{(2-\alpha)/2}\left(\int_0^\infty \left[(s+1)\bar{\mathbf{v}}_s - 1\right]^2 ds\right)^{\frac{\alpha}{2}}, \quad (23)$$

where $C(\alpha) = \int_0^\infty (s+1)^{-2\alpha/(2-\alpha)} ds$. Note that $2\alpha/(2-\alpha) > 1$ and so $C(\alpha) < \infty$ for $\alpha > 2/3$. Inequalities (22) and (23) yield

$$E\left(\int_0^\infty \left|\bar{\mathbf{v}}_s - \frac{1}{s+1}\right| ds\right) \le C(\alpha) E\left\{M^{1-\alpha}\left(\int_0^\infty \left|(s+1)\bar{\mathbf{v}}_s - 1\right|^2 ds\right)^{\alpha/2}\right\},$$

where $M = \sup_{0\le s<\infty}\left|\bar{\mathbf{v}}_s - \frac{1}{s+1}\right|$. The Hölder inequality implies

$$E\left(\int_0^\infty \left|\bar{\mathbf{v}}_s - \frac{1}{s+1}\right| ds\right)$$
$$\le C(\alpha)\left\{E\left(M^{p(1-\alpha)}\right)\right\}^{1/p}\left\{E\left(\int_0^\infty \left|(s+1)\bar{\mathbf{v}}_s - 1\right|^2 ds\right)^{q\alpha/2}\right\}^{1/q}.$$

Let $\alpha = 2(1+\eta)/3$, $q = 3(1-\eta)/[2(1+\eta)]$ and $p = q/(q-1) = 3(1-\eta)/(1-5\eta)$ for some $\eta \in (0, 1/5)$. Then $q\alpha/2 < 1/2$, and so Lemma 1 below provides

$$E\left[\int_0^\infty \left[(s+1)\bar{\mathbf{v}}_s - 1\right]^2 ds\right]^{q\alpha/2} = E\left[\int_\tau^\infty \left((I_s^{(\tau)}+1)\mathbf{v}_s^{(\tau)} - 1\right)^2 h^2(s-\tau) ds\right]^{q\alpha/2} < \infty.$$

Since $0 \le \mathbf{v}_t^{(\tau)} \le E(Y^2|\mathcal{F}_t^{W,\tau})$ and $E(Y^{2+\delta}) < \infty$ we have $E[\sup_{0\le t<\infty}|\bar{\mathbf{v}}_t^{(\tau)}|^\beta] < \infty$ for $0 \le \beta < 1 + \delta/2$. Now $p(1-\alpha) = (1-\eta)(1-2\eta)/(1-5\eta)$. Hence $p(1-\alpha) < 1 + \delta/2$ for η sufficiently small. Then $E(M^{p(1-\alpha)}) < \infty$ and the assertion follows.

Lemma 1. *$E[\int_\tau^\infty ((I_s^{(\tau)}+1)\mathbf{v}_s^{(\tau)} - 1)^2 h^2(s-\tau) ds]^\alpha < \infty$ for all $0 \le \alpha < 1/2$.*

Proof. We first derive stochastic differential equations for $\widehat{Y}_t^{(\tau)}$ and $[I_t^{(\tau)}+1]\widehat{Y}_t^{(\tau)}$. It holds that $\widehat{Y}_t^{(\tau)} = \mu(X_t^{(\tau)}, I_t^{(\tau)})$ and $\widehat{Y_t^2}^{(\tau)} = \rho(X_t^{(\tau)}, I_t^{(\tau)})$, where

$$\mu(x,t) = \int_0^\infty y e^{yx - \frac{y^2}{2}t} g(y) dy \Big/ \int_0^\infty e^{yx - \frac{y^2}{2}t} g(y) dy,$$

$$\rho(x,t) = \int_0^\infty y^2 e^{yx - \frac{y^2}{2}t} g(y) dy \Big/ \int_0^\infty e^{yx - \frac{y^2}{2}t} g(y) dy,$$

and $X_t^{(\tau)} = \int_{\tau\wedge t}^t h(u-\tau)dW_u$. Note that after the random time change $\xi^{(\tau)}$ the process $X^{(\tau)}$ becomes a Brownian motion with linear drift. See Novikov (1971) and Chapter 5 in Küchler and Sørensen (1997) for this type of transformation. We have (see Paulsen (1999)) $\partial_x\mu(x,t) = v(x,t)$ and $(1/2)\partial_x\partial_x\mu(x,t) + \partial_t\mu(x,t) = -\mu(x,t)v(x,t)$, where $v(x,t) = \rho(x,t) - \mu(x,t)^2$. Therefore Itô's formula implies $d\widehat{Y}_t^{(\tau)} = v(X_t^{(\tau)}, I_t^{(\tau)})d\overline{X}_t^{(\tau)} = \mathbf{v}_t^{(\tau)} d\overline{X}_t^{(\tau)}$ and further

$$d\left([I_t^{(\tau)}+1]\widehat{Y}_t^{(\tau)} - X_t^{(\tau)}\right) = \left[\left(I_t^{(\tau)}+1\right)\mathbf{v}_t^{(\tau)} - 1\right] d\overline{X}_t^{(\tau)}, \tag{24}$$

with $\overline{X}_t^{(\tau)} = X_t^{(\tau)} - \int_{\tau\wedge t}^t \widehat{Y}_u^{(\tau)} h(u-\tau)du$. Lemma 1 on p. 21 in Stein (1986) yields for all $B > 0$ and $A \in (-\infty, +\infty)$,

$$\int_0^\infty (By - A)e^{Ay - B\frac{y^2}{2}} g(y)dy = g(0) + \int_0^\infty \left[\frac{g'(y)}{g(y)}\right] e^{Ay - B\frac{y^2}{2}} g(y)dy.$$

Then, with $A = X_t^{(\tau)}$ and $B = I_t^{(\tau)}$, for $\tau \le t < \infty$,

$$I_t^{(\tau)}\widehat{Y}_t^{(\tau)} - X_t^{(\tau)} = g(0)\frac{1}{\psi_t^{(\tau)}} + E\left(H(Y)|\mathcal{F}_t^{W,\tau}\right), \tag{25}$$

where $\psi_t^{(\tau)} = \int_0^\infty e^{y\int_{\tau\wedge t}^t h(u-\tau)dW_u - \frac{y^2}{2}\int_{\tau\wedge t}^t h^2(u-\tau)du} g(y)dy = dP/dP_\infty|\mathcal{F}_t^{W,\tau}$. See Proposition 4 in Woodroofe (1992) for a related identity. We note two useful facts. We have $E(|H(Y)\log|H(Y)||) < \infty$. Therefore Doob's inequality and $EY^2 < \infty$ provide $E\left[\sup_{0\le t<\infty} E(H(Y)|\mathcal{F}_t^{W,\tau})\right] < \infty$ and $E\left[\sup_{0\le t<\infty} E(Y|\mathcal{F}_t^{W,\tau})\right] < \infty$. Moreover for $x \ge 1$,

$$P\left(\sup_{0\le t<\infty}[1/\psi_t^{(\tau)}] > x\right) = P(\sigma_x^{(\tau)} < \infty) = E_\infty\left([\psi_{\sigma_x^{(\tau)}}^{(\tau)}]1_{\{\sigma_x^{(\tau)}<\infty\}}\right) = 1/x, \tag{26}$$

where $\sigma_x^{(\tau)} = \inf\{t > 0|[\psi_t^{(\tau)}]^{-1} > x\}$. Note that $P_\infty(\sigma_x^{(\tau)} < \infty) = 1$ for $x \ge 1$. We immediately obtain $E(\sup_{0\le t<\infty}[1/\psi_t^{(\tau)}])^\beta < \infty$ for $0 \le \beta < 1$. So $E(\sup_{0\le t<\infty}|(I_t^{(\tau)}+1)Y_t^{(\tau)} - X_t^{(\tau)}|^\beta) < \infty$ for all $0 \le \beta < 1$. Representation (24) and the Burkholder-Davis-Gundy inequalities now yield the assertion. Note that $dX_t^{(\tau)} = h(t-\tau)dW_t$ for $\tau \le t < \infty$.

References

Assaf, D., Pollak, M., Ritov, Y. and Yakir, B. (1993). Detecting a change of a normal mean by dynamic sampling with a probability bound on a false alarm. *Ann. Statist.* **21**, 1155-1165.

Beibel, M. (1997). Sequential change-point detection in continuous time when the post-change drift is unknown. *Bernoulli* **3**, 457-478.

Boukai, B. and Zhou, H. (1997). Nonparametric estimation in a two change-point model. *Nonparametr. Statist.* **8**, 275-292.

Huang, W. T. and Chang, Y. P. (1993). Nonparametric estimation in change-point models. *J. Statist. Plann. Inference* **35**, 335-347.

Huskova, M. (1999). Gradual changes versus abrupt changes. *J. Statist. Plann. Inference* **76**, 109-125.

Kolmogorov, A. N., Prokhorov, Y. V. and Shiryayev, A. N. (1990). Methods of detecting spontaneously occuring effects. *Proc. Steklov Inst. Math.* **1**, 1-21.

Küchler, U. and Sørensen, M. (1997). *Exponential Families of Stochastic Processes.* Springer-Verlag, Berlin.

Lai, T. L. (1995). Sequential changepoint detection in quality control and dynamical systems. *J. Roy. Statist. Soc. Ser. B* **57**, 613-658.

Lerche, H. R. (1986). The shape of Bayes tests of power one. *Ann. Statist.* **14**, 1030-1048.

Liptser, R. S. and Shiryayev, A. N. (1977). *Statistics of Random Processes*, Volume 1. Springer-Verlag, Berlin.

Novikov, A. A. (1971). Sequential estimation of parameters of diffusion processes. *Theory Probab. Appl.* **16**, 344-349.

Paulsen, V. (1999). A martingale approach for detecting the drift of a Wiener process. *Stochastic Process. Appl.* **80**, 177-191.

Pollak, M. and Siegmund, D. (1975). Approximations to the expected sample size of certain sequential tests. *Ann. Statist.* **3**, 1267-1282.

Pollak, M. and Siegmund, D. (1985). A diffusion process and its application to detecting a change in the drift of Brownian motion. *Biometrika* **72**, 267-280.

Siegmund, D. (1994). A retrospective of Wald's sequential analysis — its relation to change-point detection and sequential clinical trials. In *Statistical Decision Theory and Related Topics V* (Edited by S. S. Gupta and J. O. Berger), 19-33. Springer-Verlag, Berlin.

Shiryayev, A. N. (1963). On optimum methods in quickest detection problems. *Theory Probab. Appl.* **8**, 22-41.

Stein, C. (1986). *Approximate Computation of Expectations.* IMS Lecture Notes-Monograph Series, **7**. IMS, Hayward, CA.

Tartakovsky, A. G. (1995). Asymptotic properties of CUSUM and Shiryaev's procedures for detecting a change in a nonhomogeneous Gaussian process. *Math. Methods Statist.* **4**, 389-404.

Woodroofe, M. (1992). Integrable expansions for posterior distributions for one-parameter exponential families. *Statist. Sinica* **2**, 91-111.

Yakir, B. (1996). A lower bound on the ARL to detection of a change with a probability constraint on false alarm. *Ann. Statist.* **24**, 431-435.

Zacks, S. (1983). Survey of classical and Bayesian approaches to the change–point problem: Fixed sample and sequential procedures of testing and estimation. In *Recent Advances in Statistics, Papers in Honor of Herman Chernoff* (Edited by M. H. Rizvi, J. Rustagi and D. Siegmund), 245-269. Academic Press, New York.

Institut für Mathematische Stochastik, Albert–Ludwigs–Universität Freiburg, Eckerstr. 1, 79104 Freiburg, Germany.

E-mail: beibel@stochastik.uni-freiburg.de

(Received March 1998; accepted November 1999)