

Monotonicity and robustness in Wiener disorder detection

Erik Ekström*

Juozas Vaicenavicius †

Abstract

We study the problem of detecting a drift change of a Brownian motion under various extensions of the classical case. Specifically, we consider the case of a random post-change drift and examine monotonicity properties of the solution with respect to different model parameters. Moreover, robustness properties – effects of misspecification of the underlying model – are explored.

1 Introduction

In the classical version of the quickest disorder detection (QDD) problem [9], one observes a one-dimensional process Y which satisfies

$$Y_t = b(t - \Theta)^+ + \sigma W_t,$$

where b and σ are non-zero constants, W is a standard Brownian motion and the disorder time Θ is an exponentially distributed random variable (with intensity $\lambda > 0$) such that W and Θ are independent. The associated Bayes' risk (expected cost) corresponding to a stopping rule τ is defined as

$$\mathbb{P}(\Theta > \tau) + c\mathbb{E}[(\tau - \Theta)^+], \tag{1.1}$$

where $c > 0$ is the cost of one unit of detection delay. It is well-known (see [10, Chapter 4]) that to minimise the Bayes risk one should stop the first time the conditional probability process $\Pi_t := \mathbb{P}(\Theta \leq t | \mathcal{F}_t^Y)$ reaches a certain level a . Moreover, the level a is characterized as the unique solution of a transcendental equation.

In many situations, however, it is natural *not* to know the exact value of the disorder magnitude b , but merely its *distribution*. This is the case for example when a specific machine is monitored continuously, and the machine can break down in several possible ways. To study such a situation, we allow for the new drift to be a random variable B with distribution μ such that B is independent of the other sources of randomness. In this setting we study **monotonicity properties** of the QDD problem, i.e. whether the

*Department of Mathematics, Uppsala University, Box 480, 751 06 Uppsala, Sweden (ekstrom@math.uu.se).

†Department of Information Technology, Uppsala University, Box 337, 751 05 Uppsala, Sweden (juozas.vaicenavicius@it.uu.se).

(minimal) expected cost is monotone with respect to various model parameters. In particular, we study the dependence of the expected cost on the volatility σ , the distribution μ , and the disorder intensity λ . We also study **robustness** in the QDD problem, i.e. what happens if one misspecifies various model parameters. More specifically, we aim at estimates for the increased cost associated with the use of suboptimal strategies. Clearly, such estimates are helpful in situations where the model is badly calibrated as well as when one chooses to use a simpler suboptimal strategy rather than a computationally more demanding optimal strategy.

As mentioned above, the classical version of the QDD problem was studied in [9], see also [10, Chapter 4] and [8, Section 22]; for extensions to the case of detecting a change in the intensity of a Poisson process, see [7], [3], [4]. Regarding the case of a random disorder magnitude, [2] obtains asymptotic results of a problem with normally distributed drift. Concavity of the value function in a related hypothesis testing problem with two possible post-change drift values in a time-homogeneous case was obtained in [6]. Finally, practical significance of the disorder detection problem in modern engineering applications is explained in [11].

2 General model formulation

We model a signal-processing activity on a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where the filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ satisfies the usual conditions. We are interested in the signal process X , which is not directly observable, but we can continuously observe the noisy process

$$Y_t = \int_0^t X_u du + \int_0^t \sigma(u) dW_u, \quad t \geq 0. \quad (2.1)$$

Here W is a Brownian motion independent of X , the dispersion σ is deterministic and strictly positive, and the signal process follows

$$X_t = B^0 \mathbb{1}_{\{\Theta=0\}} + B^1 \mathbb{1}_{\{0 < \Theta \leq t\}}, \quad (2.2)$$

where Θ is a $[0, \infty)$ -valued random variable representing the disorder occurrence time. Moreover, B^0, B^1 are real-valued random variables corresponding to disorder magnitudes in the cases ‘disorder occurs before we start observing Y ’ and ‘disorder occurs while we observe Y ’, respectively. Also, Θ, B^0 , and B^1 are independent. Let Θ have the distribution $\tilde{\pi} \delta_0 + (1 - \tilde{\pi}) \nu$, where ν is a probability measure on $(0, \infty)$ with a continuously differentiable distribution function F_ν . In addition, denote the distributions of B^0 and B^1 by μ^0 and μ^1 , respectively. When referring to μ^0 and μ^1 collectively, we will simply say that the prior is μ . Let us introduce the notation

$$D^n := \{\pi \in [0, \infty)^n : \|\pi\|_1 \leq 1\}$$

and

$$\Delta^n := \{\pi \in [0, \infty)^n : \|\pi\|_1 = 1\},$$

where $\|\pi\|_1 = \sum_{i=1}^n \pi_i$. We assume that

$$\mu^0 = \sum_{i=1}^n \check{p}_i \delta_{b_i}, \quad \mu^1 = \sum_{i=1}^n p_i \delta_{b_i},$$

where $b_1, \dots, b_n \in \mathbb{R} \setminus \{0\}$ and $(\check{p}_1, \dots, \check{p}_n), (p_1, \dots, p_n) \in \Delta^n$.

We are interested in a disorder detection strategy τ incorporating two objectives: short detection delay and a small portion of false alarms. As noted in the introduction, a classical choice of Bayes' risk (see [10]) for a detection strategy to minimize is given by (1.1). In the present paper, we consider a slightly more flexible risk structure by allowing a time-dependent cost for the detection delay. More precisely, we consider the Bayes risk

$$R(\tau) = \mathbb{E} \left[\mathbb{1}_{\{\tau < \Theta\}} + \int_{\Theta}^{\tau} c(u) du \right]$$

where $\mathbb{1}_{\{\tau < \Theta\}}$ is a fixed penalty for a false alarm, and the term $\int_{\Theta}^{\tau} c(u) du$ is a penalty for detection delay. Here $t \mapsto c(t)$ is a deterministic function with $c(t) > 0$ for all $t \geq 0$. Writing $\mathcal{F}^Y = \{\mathcal{F}_t^Y\}_{t \geq 0}$ for the filtration generated by Y (which is our observation filtration), let us introduce $\tilde{\Pi}_t := \mathbb{E}[\mathbb{1}_{\mathbb{R} \setminus \{0\}}(X_t) | \mathcal{F}_t^Y]$. Then

$$\begin{aligned} R(\tau) &= \mathbb{E} \left[1 - \mathbb{E}[\mathbb{1}_{\{\Theta \leq \tau\}} | \mathcal{F}_\tau^Y] \right] + \int_0^\infty c(t) \mathbb{E}[\mathbb{1}_{\{t \leq \tau\}} \mathbb{E}[\mathbb{1}_{\{\Theta \leq t\}} | \mathcal{F}_t^Y]] dt \\ &= \mathbb{E} \left[1 - \tilde{\Pi}_\tau + \int_0^\tau c(t) \tilde{\Pi}_t dt \right]. \end{aligned}$$

Hence the optimal stopping problem to solve is

$$V = \inf_{\tau \in \mathcal{T}^Y} \mathbb{E} \left[1 - \tilde{\Pi}_\tau + \int_0^\tau c(t) \tilde{\Pi}_t dt \right], \quad (2.3)$$

where \mathcal{T}^Y denotes the set of \mathcal{F}^Y -stopping times.

2.1 Filtering equations

Let us define $\Pi_t^{(i)} := \mathbb{E}[\mathbb{1}_{\{X_t = b_i\}} | \mathcal{F}_t^Y]$, where $i = 1, \dots, n$. By the Kallianpur-Striebel formula (see [5, Theorem 2.9 on p. 39]),

$$\begin{aligned} \Pi_t^{(i)} &= \frac{\tilde{\pi} \check{p}_i e^{\int_0^t \frac{b_i}{\sigma(u)^2} dY_u - \int_0^t \frac{b_i^2}{2\sigma(u)^2} du} + (1 - \tilde{\pi}) p_i \int_{[0,t]} e^{\int_\theta^t \frac{b_i}{\sigma(u)^2} dY_u - \int_\theta^t \frac{b_i^2}{2\sigma(u)^2} du} \nu(d\theta)}{\tilde{\pi} \sum_j \check{p}_j e^{\int_0^t \frac{b_j}{\sigma(u)^2} dY_u - \int_0^t \frac{b_j^2}{2\sigma(u)^2} du} + (1 - \tilde{\pi}) \left(\sum_j p_j \int_{[0,t]} e^{\int_\theta^t \frac{b_j}{\sigma(u)^2} dY_u - \int_\theta^t \frac{b_j^2}{2\sigma(u)^2} du} \nu(d\theta) + \nu((t, \infty)) \right)} \end{aligned} \quad (2.4)$$

for $i = 1, \dots, n$. Moreover, from the Kushner-Stratonovich equation (see [5, Theorem 3.1 on p. 58]), we know that $\Pi^{(i)}$ satisfies

$$d\Pi_t^{(i)} = p_i \lambda(t) \left(1 - \sum_{j=1}^n \Pi_t^{(j)}\right) dt + \frac{\Pi_t^{(i)}}{\sigma(t)} \left(b_i - \sum_{j=1}^n b_j \Pi_t^{(j)}\right) d\hat{W}_t, \quad i = 1, \dots, n. \quad (2.5)$$

Here $\lambda(t) = F'_\nu(t)/(1 - F_\nu(t))$ is the intensity of the disorder occurring at time $t > 0$ (conditional on not having occurred yet), and

$$\hat{W}_t = \int_0^t \frac{1}{\sigma(u)} (dY_u - \mathbb{E}[X_u | \mathcal{F}_u^Y]) du$$

is a standard Brownian motion with respect to $\{\mathcal{F}_t^Y\}_{t \geq 0}$, see [1] (the process \hat{W}_t is referred to as *the innovation process*). Note that $\tilde{\Pi}_t = \sum_{i=1}^n \Pi_t^{(i)}$ yields

$$d\tilde{\Pi}_t = \lambda(t)(1 - \tilde{\Pi}_t) dt + \frac{\hat{X}_t}{\sigma(t)} (1 - \tilde{\Pi}_t) d\hat{W}_t, \quad (2.6)$$

where $\hat{X}_t = \mathbb{E}[X_t | \mathcal{F}_t^Y]$.

The posterior distribution $\mathbb{P}(X_t \in \cdot | \mathcal{F}_t^Y) = \sum_{i=1}^n \Pi_t^{(i)} \delta_{b_i}(\cdot)$, so the n -tuple $\Pi_t = (\Pi_t^{(1)}, \dots, \Pi_t^{(n)})$ fully describes the posterior. As a result, (2.4) and (2.5) provide two different representations of the posterior distribution.

2.2 Markovian embedding

Following standard lines in optimal stopping theory, we embed our optimal stopping problem into a Markovian framework. To do that, define a Markovian value function V by

$$V(t, \pi) := \inf_{\tau \in \mathcal{T}_t^\Pi} \mathbb{E}^{t, \pi} \left[1 - \tilde{\Pi}_{t+\tau} + \int_t^{t+\tau} c(u) \tilde{\Pi}_u du \right], \quad (t, \pi) \in [0, \infty) \times D^n, \quad (2.7)$$

where \mathcal{T}_t^Π denotes the stopping times with respect to the n -dimensional process $\{\Pi_{t+s}^{t, \pi}\}_{s \geq 0}$ starting from π at time t and satisfying (2.5). It is worth noting that $V(t, \pi)$ corresponds to the value of the problem (2.3) in which the initial time is t and $\mu_0 = \sum_{i=1}^n \pi_i \delta_{b_i}$.

Remark 2.1. The value function $V(t, \cdot)$ in (2.7) is concave for any $t \geq 0$. Indeed, the concavity proof in [6] extends to the current setting. Since concavity is not used in the monotonicity results below, however, we omit the details.

2.2.1 The classical Shiryaev solution

In this subsection we recall the solution in the classical case where the cost c , the intensity λ and the post-change drift b are constants. In that case, we have the optimal stopping

problem

$$U(\pi) = \sup_{\tau \in \mathcal{T}^\Pi} \mathbb{E} \left[1 - \Pi_\tau + c \int_0^\tau \Pi_t dt \right] \quad (2.8)$$

with an underlying diffusion process

$$d\Pi_t = \lambda(1 - \Pi_t) dt + \frac{b}{\sigma} \Pi_t(1 - \Pi_t) d\hat{W}_t.$$

It is well-known (see [10, Chapter 4] or [8, Sections 22]) that U solves the free-boundary problem

$$\begin{cases} \frac{b^2 \pi^2 (1-\pi)^2}{2\sigma^2} \partial_\pi^2 U + \lambda(1-\pi) \partial_\pi U + c\pi = 0 & \pi \in (0, a) \\ U(\pi) = 1 - \pi & \pi \in [a, 1] \\ \partial_\pi U(a) = -1. \end{cases} \quad (2.9)$$

Here $a \in (0, 1)$ is the free-boundary, and it can be determined as the solution of a certain transcendental equation. Moreover, the stopping time $\tau := \inf\{t \geq 0 : \Pi_t \geq a\}$ is optimal in (2.8), and one can check that the value function U is decreasing and concave.

3 Value dependencies and robustness

3.1 Monotonicity properties of the value function

In this section, we study parameter dependence of the optimal stopping problem (2.7). In particular, we investigate how the value function changes when we alter parameters of the probabilistic model, which include the prior for the drift magnitude and the prior for the disorder time.

The effects of adding more noise, stretching out the prior by scaling, and increasing the observation cost are explained by the following theorem.

Theorem 3.1 (General monotonicity properties of the value function v).

1. V is increasing in the volatility $\sigma(\cdot)$.
2. Let V_k denote the Markovian value function (2.7) in the case when the drift magnitude prior is $\mu(\frac{\cdot}{k})$. Then the map $k \mapsto V_k(t, \pi)$ is decreasing on $(0, \infty)$ for any (t, π) .
3. V is increasing in the cost function $c(\cdot)$.

Proof. For simplicity of notation, and without loss of generality, we consider the case $t = 0$ in the proofs below.

1. For the volatility, let $t \mapsto \sigma_1(t)$ and $t \mapsto \sigma_2(t)$ be two time-dependent volatility functions satisfying $\sigma_1(t) \leq \sigma_2(t)$ for all $t \geq 0$. Also, let

$$Y_t^i := \int_0^t X_u du + \int_0^t \sigma_i(u) dW_u, \quad i = 1, 2,$$

and let V_i , $i = 1, 2$, be the corresponding value functions. In addition, let W^\perp be a standard Brownian motion independent of W and X . Then, clearly,

$$V_1 = \inf_{\tau \in \mathcal{T}^{Y^1}} \mathbb{E} \left[\mathbb{1}_{\{\tau < \Theta\}} + \int_{\Theta}^{\tau} c(u) du \right] = \inf_{\tau \in \mathcal{T}^{Y^1, W^\perp}} \mathbb{E} \left[\mathbb{1}_{\{\tau < \Theta\}} + \int_{\Theta}^{\tau} c(u) du \right].$$

Moreover, the process

$$\tilde{Y}_t^2 := Y_t^1 + \int_0^t \sqrt{\sigma_2^2(u) - \sigma_1^2(u)} dW_u^\perp$$

coincides in law with Y^2 and $\mathcal{T}^{\tilde{Y}^2} \subseteq \mathcal{T}^{Y^1, W^\perp}$. Hence it follows that

$$V_1 = \inf_{\tau \in \mathcal{T}^{Y^1, W^\perp}} \mathbb{E} \left[\mathbb{1}_{\{\tau < \Theta\}} + \int_{\Theta}^{\tau} c(u) du \right] \leq \inf_{\tau \in \mathcal{T}^{\tilde{Y}^2}} \mathbb{E} \left[\mathbb{1}_{\{\tau < \Theta\}} + \int_{\Theta}^{\tau} c(u) du \right] = V_2,$$

which finishes the proof of the claim.

2. Note that for $k > 0$, the process

$$Y_t^k := \int_0^t k X_u du + \int_0^t \sigma(u) dW_u$$

satisfies $Y_t^k = k \tilde{Y}_t$, where

$$\tilde{Y}_t := \int_0^t X_u du + \int_0^t \frac{\sigma(u)}{k} dW_u.$$

Moreover, the set of \mathcal{F}^{Y^k} -stopping times coincides with the set of $\mathcal{F}^{\tilde{Y}}$ -stopping times, so monotonicity in k is implied by monotonicity in the volatility. Thus claim 2 follows from claim 1.

3. The fact that the value is increasing in c is obvious from the definition (2.7) of the value function. □

It is not clear how a change in the disorder intensity λ affects the value function under a general disorder magnitude distribution. However, we have the following comparison result for the case of constant parameters.

Theorem 3.2 (Monotonicity in the intensity for constant parameters). *Assume that the disorder magnitude can only take one value $b \in \mathbb{R} \setminus \{0\}$. Let the cost c , the volatility σ and the intensity λ be constants, and assume that $\lambda \geq \lambda'(\cdot)$. Let U be the value function for Shiryayev's problem with parameters (b, σ, λ, c) , and let V denote the value function for the problem specification (b, σ, λ', c) . Then $U(\pi) \leq V(t, \pi)$ for all $\pi \in [0, 1]$ and $t \geq 0$.*

Proof. Without loss of generality, we only consider the case $t = 0$. Let $\pi \in [0, 1]$ and $\tau \in \mathcal{T}^Y$ be a bounded stopping time. Also, denote by Π' the process Π started from π at

time 0 when the model specification is (b, σ, λ', c) . Then, applying (a generalised version of) Ito's formula and taking expectations at the stopping time τ , we get

$$\begin{aligned}
U(\pi) &= \mathbb{E} [U(\Pi'(\tau))] - \mathbb{E} \left[\int_0^\tau \left(\lambda'(s) (1 - \Pi'(s)) \partial_\pi U(\Pi'(s)) \right. \right. \\
&\quad \left. \left. + \frac{b^2}{2\sigma^2} (\Pi'(s))^2 (1 - \Pi'(s))^2 \partial_\pi^2 U(\Pi'(s)) \right) ds \right] \\
&\leq \mathbb{E} [U(\Pi'(\tau))] - \mathbb{E} \left[\int_0^\tau \left(\lambda (1 - \Pi'(s)) \partial_\pi U(\Pi'(s)) \right. \right. \\
&\quad \left. \left. + \frac{b^2}{2\sigma^2} (\Pi'(s))^2 (1 - \Pi'(s))^2 \partial_\pi^2 U(\Pi'(s)) \right) ds \right] \\
&\leq \mathbb{E} [U(\Pi'(\tau))] + \mathbb{E} \left[c \int_0^\tau \Pi'(s) ds \right] \\
&\leq \mathbb{E} [1 - \Pi'(\tau)] + \mathbb{E} \left[c \int_0^\tau \Pi'(s) ds \right],
\end{aligned}$$

where we used the monotonicity of U and the fact that

$$\lambda(1 - \pi) \partial_\pi U(\pi) + \frac{b^2}{2\sigma^2} \pi^2 (1 - \pi)^2 \partial_\pi^2 U(\pi) + c\pi \geq 0 \quad (3.1)$$

at all points away from the optimal stopping boundary of Shiryaev's classical problem, compare (2.9). Taking the infimum over bounded stopping times τ , we get $U(\pi) \leq V(0, \pi)$, which finishes the proof. \square

Remark 3.3. The monotonicity in intensity result does not easily extend to cases with unknown post-change drift by the same argument. It is not difficult to convince oneself that in higher dimensions the partial derivatives $\frac{\partial V}{\partial \pi_i}$ are not necessarily all negative.

3.2 Robustness

Robustness concerns how a possible misspecification of the model parameters affects the performance of the detection strategy when evaluated under the real physical measure. In this section, we use coupling arguments to study robustness properties with respect to the disorder magnitude and disorder time. For simplicity, we assume that the parameters λ , c and σ are constant so that we have a time-independent case; generalizations to the time-dependent case are straightforward but notationally more involved.

Thus we assume that the signal process follows

$$X_t = B^0 \mathbb{1}_{\{\Theta=0\}} + B^1 \mathbb{1}_{\{0 < \Theta \leq t\}}, \quad (3.2)$$

where B^0, B^1 are random variables with distributions μ^0, μ^1 respectively, and Θ has the distribution $\nu_{\tilde{\pi}} := \tilde{\pi} \delta_0 + (1 - \tilde{\pi}) \nu$, where ν is an exponential distribution with intensity λ . Let us simply write $\mu := (\mu^0, \mu^1)$.

For a given $l \in \mathbb{R} \setminus \{0\}$, let Θ_l satisfy $\Theta_l \geq \Theta$ with distribution $\tilde{\pi}\delta_0 + (1 - \tilde{\pi})\nu_l$, where ν_l is an exponential distribution with intensity $\lambda_l \leq \lambda$. Let

$$g_l(t, \tilde{\pi}, Y) := \frac{\tilde{\pi}e^{\frac{l}{\sigma^2}Y_t - \frac{l^2}{2\sigma^2}t} + (1 - \tilde{\pi})\lambda_l \int_0^t e^{\frac{l}{\sigma^2}(Y_t - Y_\theta) - \frac{l^2}{2\sigma^2}(t-\theta)} e^{-\theta/\lambda_l} d\theta}{\tilde{\pi}e^{\frac{l}{\sigma^2}Y_t - \frac{l^2}{2\sigma^2}t} + (1 - \tilde{\pi}) \left(\lambda_l \int_0^t e^{\frac{l}{\sigma^2}(Y_t - Y_\theta) - \frac{l^2}{2\sigma^2}(t-\theta)} e^{-\theta/\lambda_l} d\theta + 1 - e^{-t/\lambda_l} \right)}, \quad (3.3)$$

compare (2.4). Also, we introduce the notation

$$\begin{aligned} Y_t^\mu &:= \int_0^t X_u du + \sigma W_t, \\ Y_t^{\delta_l} &:= l(t - \Theta_l)^+ + \sigma W_t, \\ \tilde{\Pi}_{\delta_l}^{\delta_l}(t) &:= g_l(t, \tilde{\pi}, Y^{\delta_l}) \end{aligned}$$

and

$$\tilde{\Pi}_{\delta_l}^\mu(t) := g_l(t, \tilde{\pi}, Y^\mu).$$

Here Y^μ is the observation process for a setting in which the post-change drift has distribution μ and the disorder happens at Θ . The process Y^{δ_l} is the observation process and $\tilde{\Pi}_{\delta_l}^{\delta_l}$ is the corresponding conditional probability process in the situation of a post-change drift l that occurs at Θ_l . Moreover, the process $\tilde{\Pi}_{\delta_l}^\mu$ represents the conditional probability process calculated *as if* the drift change is described by (l, Θ_l) in the scenario where the true drift-change is given by (μ, Θ) .

Now, let $a := a_l$ denote the optimal stopping boundary for the classical Shiryaev one-dimensional problem in the scenario (l, Θ_l) , and define

$$\tau_{\delta_l}^{\delta_l} := \inf\{t \geq 0 : \tilde{\Pi}_{\delta_l}^{\delta_l}(t) \geq a\},$$

$$\tau_{\delta_l}^\mu := \inf\{t \geq 0 : \tilde{\Pi}_{\delta_l}^\mu(t) \geq a\},$$

and

$$V_{\delta_l}^\mu := \mathbb{E}[\mathbb{1}_{\{\tau_{\delta_l}^\mu < \Theta\}} + c(\tau_{\delta_l}^\mu - \Theta)^+].$$

Here $\tau_{\delta_l}^{\delta_l}$ is the optimal stopping time in the scenario (l, Θ_l) , and $\tau_{\delta_l}^\mu$ is the (sub-optimal) stopping time and $V_{\delta_l}^\mu$ is the corresponding cost for someone who believes in (l, Θ_l) , whereas the true model has (μ, Θ) .

Finally, let

$$\tilde{\Pi}_t^\mu := \mathbb{P}(\mathbb{1}_{\mathbb{R} \setminus \{0\}}(X_t) | \mathcal{F}_t^{Y^\mu}) = \Pi_t^{(1)} + \dots + \Pi_t^{(n)}$$

as in Section 2, and define

$$\gamma_{\delta_l}^\mu := \inf\{t \geq 0 : \tilde{\Pi}_t^\mu \geq a\}.$$

Theorem 3.4 (Robustness).

1. Suppose that $\inf(\text{supp } \mu) > 0$ or $\sup(\text{supp } \mu) < 0$, and let $l := \arg \min_{x \in \text{supp}(\mu)} |x|$.

(a) Then

$$V^\mu \leq V_{\delta_l}^\mu \leq V^{\delta_l} + c \frac{\lambda - \lambda_l}{\lambda \lambda_l} (1 - \tilde{\pi}), \quad (3.4)$$

where V^μ and V^{δ_l} denote the minimal associated Bayes risks for the prior distributions μ and δ_l , respectively.

(b) Also,

$$V^\mu \leq \mathbb{P}(\Theta > \gamma_{\delta_l}^\mu) + c \mathbb{E}[(\gamma_{\delta_l}^\mu - \Theta)^+] \leq V^{\delta_l}. \quad (3.5)$$

2. Suppose $r := \arg \max_{x \in \text{supp}(\mu)} |x|$, and define $V_{\delta_r}^\mu$ like $V_{\delta_l}^\mu$ for $l = r$. If $\lambda_r \geq \lambda$, then

$$V^{\delta_r} \leq V^\mu \leq V_{\delta_r}^\mu. \quad (3.6)$$

Remark 3.5. Note that (3.4) and (3.6) correspond to situations where the tester mis-specifies the model. More precisely, filtering and stopping are performed as if the underlying model had the one-point distribution for the prior (the classical Shiryaev case). Equation (3.4) gives a bound for the loss in performance for a tester who underestimates the disorder intensity and the disorder magnitude (believing that it is a constant). In (3.5), on the other hand, filtering is performed according to the correct model but Shiryaev's threshold strategy (suboptimal) is used for stopping.

Proof. 1. (a) For definiteness, we consider the case $\inf(\text{supp } \mu) > 0$ so that $l > 0$; the other case is completely analogous. First note that the suboptimality of $\tau_{\delta_l}^\mu$ yields $V^\mu \leq V_{\delta_l}^\mu$. Next, observe that we have $Y_t^{\delta_l} = Y_t^\mu$ for all $0 \leq t \leq \Theta$ and $Y_t^{\delta_l} \leq Y_t^\mu$ for all $t \geq 0$, and therefore

$$\tilde{\Pi}_{\delta_l}^{\delta_l}(t) = \tilde{\Pi}_{\delta_l}^\mu(t) \quad \text{for } t \in [0, \Theta]$$

and

$$\tilde{\Pi}_{\delta_l}^{\delta_l}(t) \leq \tilde{\Pi}_{\delta_l}^\mu(t) \quad \text{for all } t \geq 0$$

by the filtering equation (3.3). Consequently,

$$\tau_{\delta_l}^{\delta_l} \geq \tau_{\delta_l}^\mu,$$

so

$$\begin{aligned} \mathbb{E}[(\tau_{\delta_l}^{\delta_l} - \Theta_l)^+] &\geq \mathbb{E}[(\tau_{\delta_l}^\mu - \Theta)^+] - \mathbb{E}[(\Theta_l - \Theta)^+] \\ &= \mathbb{E}[(\tau_{\delta_l}^\mu - \Theta)^+] - \frac{\lambda - \lambda_l}{\lambda \lambda_l} (1 - \tilde{\pi}). \end{aligned} \quad (3.7)$$

Moreover, since $\tilde{\Pi}_{\delta_l}^{\delta_l}(t) = \tilde{\Pi}_{\delta_l}^\mu(t)$ on the time interval $[0, \Theta]$, we have

$$\mathbb{P}(\tau_{\delta_l}^{\delta_l} < \Theta_l) \geq \mathbb{P}(\tau_{\delta_l}^{\delta_l} < \Theta) = \mathbb{P}(\tau_{\delta_l}^\mu < \Theta),$$

which together with (3.7) yields

$$\begin{aligned} V^{\delta_l} &= \mathbb{E}[\mathbb{1}_{\{\tau_{\delta_l}^{\delta_l} < \Theta_l\}} + c(\tau_{\delta_l}^{\delta_l} - \Theta_l)^+] \\ &\geq \mathbb{E}[\mathbb{1}_{\{\tau_{\delta_l}^\mu < \Theta\}} + c(\tau_{\delta_l}^\mu - \Theta)^+] - c \frac{\lambda - \lambda_l}{\lambda \lambda_l} (1 - \tilde{\pi}) \\ &= V_{\delta_l}^\mu - c \frac{\lambda - \lambda_l}{\lambda \lambda_l} (1 - \tilde{\pi}). \end{aligned}$$

- (b) The first inequality is immediate by suboptimality of $\gamma_{\delta_l}^\mu$. For the second one, let U be the value function of the classical Shiryaev problem so that $U(\tilde{\pi}) = V^{\delta_l}$. Then U is C^2 on $[0, a_l) \cup (a_l, 1]$ and C^1 on $[0, 1]$, so applying Itô's formula to $U(\tilde{\Pi}_t)$ and taking expectations at the bounded stopping time $\gamma_{\delta_l}^\mu \wedge k$, we get

$$\begin{aligned} U(\tilde{\pi}) &= \mathbb{E}[U(\tilde{\Pi}_{\gamma_{\delta_l}^\mu \wedge k})] - \mathbb{E} \left[\int_0^{\gamma_{\delta_l}^\mu \wedge k} \lambda(1 - \tilde{\Pi}_u)U'(\tilde{\Pi}_u) + \frac{\hat{X}_u^2}{2\sigma^2}(1 - \tilde{\Pi}_u)^2U''(\tilde{\Pi}_u) du \right] \\ &\geq \mathbb{E}[U(\tilde{\Pi}_{\gamma_{\delta_l}^\mu \wedge k})] - \mathbb{E} \left[\int_0^{\gamma_{\delta_l}^\mu \wedge k} \lambda_l(1 - \tilde{\Pi}_u)U'(\tilde{\Pi}_u) + \frac{l^2}{2\sigma^2}\tilde{\Pi}_u^2(1 - \tilde{\Pi}_u)^2U''(\tilde{\Pi}_u) du \right] \\ &= \mathbb{E} \left[U(\tilde{\Pi}_{\gamma_{\delta_l}^\mu \wedge k}) \right] + \mathbb{E} \left[c \int_0^{\gamma_{\delta_l}^\mu \wedge k} \tilde{\Pi}_u du \right], \end{aligned}$$

where monotonicity and concavity of U were used in the inequality. Letting $k \rightarrow \infty$ gives

$$U(\tilde{\pi}) \geq \mathbb{E} \left[1 - \tilde{\Pi}_{\gamma_{\delta_l}^\mu} \right] + \mathbb{E} \left[c \int_0^{\gamma_{\delta_l}^\mu} \tilde{\Pi}_u du \right],$$

which finishes the proof of the claim.

2. Recall that

$$d\tilde{\Pi}_t = \lambda(1 - \tilde{\Pi}_t) dt + \frac{\hat{X}_t}{\sigma}(1 - \tilde{\Pi}_t) d\hat{W}_t.$$

Let $U(\tilde{\pi}) = V^{\delta_r}(\tilde{\pi})$. Since U is C^1 on $[0, 1]$ and C^2 on $[0, a) \cup (a, 1]$, where $a = a_r$ is the boundary in Shiryaev's problem with drift r and intensity λ_r , applying Itô's

formula to $U(\tilde{\Pi}_t)$ and taking expectations at a bounded stopping time τ yields

$$\begin{aligned} U(\tilde{\pi}) &= \mathbb{E}[U(\tilde{\Pi}_\tau)] - \mathbb{E} \left[\int_0^\tau \lambda(1 - \tilde{\Pi}_u)U'(\tilde{\Pi}_u) + \frac{\hat{X}_u^2}{2\sigma^2}(1 - \tilde{\Pi}_u)^2U''(\tilde{\Pi}_u) du \right] \\ &\leq \mathbb{E}[U(\tilde{\Pi}_\tau)] - \mathbb{E} \left[\int_0^\tau \lambda_r(1 - \tilde{\Pi}_u)U'(\tilde{\Pi}_u) + \frac{r^2}{2\sigma^2}\tilde{\Pi}_u(1 - \tilde{\Pi}_u)^2U''(\tilde{\Pi}_u) du \right] \\ &\leq \mathbb{E} \left[U(\tilde{\Pi}_\tau) \right] + \mathbb{E} \left[c \int_0^\tau \tilde{\Pi}_u du \right] \end{aligned} \quad (3.8)$$

$$\leq \mathbb{E} \left[1 - \tilde{\Pi}_\tau \right] + \mathbb{E} \left[c \int_0^\tau \tilde{\Pi}_u du \right]. \quad (3.9)$$

Here concavity was used for the first inequality, (3.8) follows from the fact that

$$\lambda_r(1 - \tilde{\pi})U'(\tilde{\pi}) + \frac{r^2}{2\sigma^2}\tilde{\pi}(1 - \tilde{\pi})^2U''(\tilde{\pi}) + c\tilde{\pi} \geq 0, \quad \tilde{\pi} \in [0, a) \cup (a, 1],$$

and the inequality (3.9) because $U(\tilde{\pi}) \leq 1 - \tilde{\pi}$. Hence, since the same value V^μ is obtained if one in (2.3) restricts the infimum to only bounded stopping times,

$$V^{\delta_r} = U \leq V^\mu.$$

Lastly, since τ_l^μ is a suboptimal strategy, we also have

$$V^\mu \leq V_{\delta_r}^\mu,$$

which finishes the claim. □

Corollary 3.6. *In the notation above, assume that $\lambda = \lambda_l$ so that there is no misspecification of the intensity. Moreover, assume that $\text{supp}(\mu) \subseteq [l, r]$, where $0 < l < r$. Then*

$$V^{\delta_r} \leq V^\mu \leq V^{\delta_l},$$

so monotonicity in the disorder magnitude holds when comparing with deterministic magnitudes. Furthermore,

$$0 \leq V_{\delta_l}^\mu - V^\mu \leq V^{\delta_l} - V^{\delta_r},$$

so the increase in the Bayes risk due to underestimation (with a constant) of the disorder magnitude is bounded by the difference of two value functions of the classical Shiryaev problem.

We finish with some implications concerning the stopping strategy $\tau_{\mathcal{D}} := \inf\{t \geq 0 : \Pi_t \in \mathcal{D}\}$, where $\mathcal{D} = \{\pi \in \Delta^n : V(\pi) = 1 - \pi\}$ is a standard abstractly defined optimal stopping set, see [8] (we now assume that we are in the case of time-independent coefficients so that the value function is merely a function of $\pi \in D^n$). The concavity of V , compare Remark 2.1, yields the existence of a boundary $\gamma \subset \Delta^n$ separating \mathcal{D} from

its complement $\Delta^n \setminus \mathcal{D}$. The following result provides a more accurate location of the boundary γ .

Corollary 3.7 (Confined stopping boundary). *Assume that the coefficients c , σ and λ are constant and that $\text{supp}(\mu) \subseteq [l, r]$, where $0 < l < r$. Let a_l and a_r denote the boundaries in the classical Shiryaev problem with disorder magnitude l and r , respectively. Then*

$$a_l \leq \inf\{\|\pi\|_1 : \pi \in \gamma\} \leq \sup\{\|\pi\|_1 : \pi \in \gamma\} \leq a_r,$$

i.e. the stopping boundary is contained in a strip. Moreover, the optimal strategy $\tau_{\mathcal{D}}$ satisfies

$$1 - a_r \leq \mathbb{P}(\tau_{\mathcal{D}} < \Theta | \mathcal{F}_{\tau_{\mathcal{D}}}^Y) \leq 1 - a_l.$$

References

- [1] Bain, A., Crisan, D. *Fundamentals of stochastic filtering*. Stochastic Modelling and Applied Probability, 60. Springer, New York, 2009.
- [2] Beibel, M. Sequential change-point detection in continuous time when the post-change drift is unknown. *Bernoulli*, vol. 3, no. 4, 1997, 457-478.
- [3] Bayraktar, E., Dayanik, S., Karatzas, K. The standard Poisson disorder problem revisited. *Stochastic Processes and their Applications*, vol. 115, no. 9, 2005, 1437-1450.
- [4] Bayraktar, E., Dayanik, S., Karatzas, I. *The Annals of Applied Probability*, vol. 16, no. 3, 2006, 1190-1261
- [5] Crisan, Dan, and Boris Rozovskii. *The Oxford handbook of nonlinear filtering*. Oxford University Press, 2011.
- [6] Muravlev, A., Shiryaev, A. Two-Sided Disorder Problem for a Brownian Motion in a Bayesian Setting. *Proceedings of the Steklov Institute of Mathematics*, vol. 287, 2014, 202-224.
- [7] Peskir, G. and Shiryaev, A. Solving the Poisson disorder problem. *Advances in finance and stochastics*, 295-312, Springer, Berlin, 2002.
- [8] Peskir, G., Shiryaev, A. *Optimal stopping and free-boundary problems*. Lectures in Mathematics, ETH Zürich. Birkhäuser Verlag, Basel, 2006.
- [9] Shiryaev, A. N. Two problems of sequential analysis. *Cybernetics*, vol. 3, no. 2, 1967, 63-69.
- [10] Shiryaev, A. N. *Optimal Stopping Rules*. Springer, New York, 1978.
- [11] Zucca, C., Tavella, P. and Peskir, G. Detecting atomic clock frequency trends using an optimal stopping method. *Metrologia*, vol. 53, 2016, 89-95.