



# Sequential Analysis

## Design Methods and Applications

ISSN: 0747-4946 (Print) 1532-4176 (Online) Journal homepage: <http://www.tandfonline.com/loi/lsqla20>

## Sequential testing of a Wiener process with costly observations

Hannah Dyrssen & Erik Ekström

To cite this article: Hannah Dyrssen & Erik Ekström (2018) Sequential testing of a Wiener process with costly observations, *Sequential Analysis*, 37:1, 47-58, DOI: [10.1080/07474946.2018.1427973](https://doi.org/10.1080/07474946.2018.1427973)

To link to this article: <https://doi.org/10.1080/07474946.2018.1427973>



Published with license by Taylor & Francis©  
2018 Hannah Dyrssen and Erik Ekström



Published online: 08 Mar 2018.



Submit your article to this journal [↗](#)



Article views: 43



View related articles [↗](#)



View Crossmark data [↗](#)

# Sequential testing of a Wiener process with costly observations

Hannah Dyrssen  and Erik Ekström

Department of Mathematics, Uppsala University, Uppsala, Sweden

## ABSTRACT

We consider the sequential testing of two simple hypotheses for the drift of a Brownian motion when each observation of the underlying process is associated with a positive cost. In this setting where continuous monitoring of the underlying process is not feasible, the question is not only whether to stop or to continue at a given observation time but also, if continuing, how to distribute the next observation time. Adopting a Bayesian methodology, we show that the value function can be characterized as the unique fixed point of an associated operator and that it can be constructed using an iterative scheme. Moreover, the optimal sequential distribution of observation times can be described in terms of the fixed point.

## ARTICLE HISTORY

Received 20 March 2017  
Revised 28 July 2017  
Accepted 20 December 2017

## KEYWORDS



Brownian motion; hypothesis testing; optimal stopping; sequential analysis

## SUBJECT CLASSIFICATIONS

62L10; 60G40; 62C10; 62L05

## 1. Introduction

In the hypothesis testing problem of a Wiener process, one seeks to determine the value of the drift of a Wiener process. Solving the problem amounts to determining a decision rule that minimizes the total expected cost, which in a Bayesian formulation of the problem is typically defined as the sum of the cost of a faulty decision and the cost of lengthy observations. Early papers in the area, including Bather (1962), Chernoff (1961, 1965) and Breakwell and Chernoff (1964) study hypothesis testing problems with normal prior distributions of the drift for various loss functions, corresponding to different costs of a faulty decision. In the absence of closed form solutions of such problems, the main focus in these references is on determining asymptotic properties of the optimal decision rule. Utilizing the connection between optimal stopping problems and free-boundary problems, Shiryaev (1969, 1978) provides an explicit solution of the hypothesis testing problem when the drift can take only two different values. Notable recent contributions include the extension to the finite horizon hypothesis testing problem (Gapeev and Peskir, 2004), the characterization of the solution to the original Chernoff problem in terms of an associated integral equation (Zhitlukhin and Muravlev, 2013), a study of the case with three hypotheses (Zhitlukhin and Shiryaev, 2011), and a study of the case with general prior distributions (Ekström and Vaicenavicius, 2015). Along a related line of research, various authors have extended the problem to include more general underlying processes. For example, a study of testing two hypotheses on the intensity of a Poisson process was provided in Peskir and Shiryaev (2000), hypotheses testing on the intensity and the jump distribution of a compound Poisson process was investigated

**CONTACT** Erik Ekström  [ekstrom@math.uu.se](mailto:ekstrom@math.uu.se)  Department of Mathematics, Uppsala University, Box 480, Uppsala 75106, Sweden.

Recommended by Allan Gut.

Published with license by Taylor & Francis © 2018 Hannah Dyrssen and Erik Ekström.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.

in Dayanik and Sezer (2006), and results on the testing of two hypotheses for some Lévy processes can be found in Buonaguidi and Muliere (2013, 2016). Furthermore, techniques similar to those employed in the statistical literature have been used to study financial problems involving simultaneous learning about the drift and financial optimization. For example, Lakner (1995) studies a classical problem of utility maximization but with incomplete information about the drift of the underlying asset, Décamps et al. (2005) investigate a timing problem for investing in a real option under incomplete information, and Ekström and Vaicenavicius (2016) consider a liquidation problem for general prior distributions of an unknown drift.

In the current article we study a version of the classical sequential hypothesis testing problem for the drift of a Wiener process where, additionally, each observation is associated with a positive cost. With this assumption, continuous observation of the underlying process is impossible, and a strategy thus consists of a decision whether to stop or not, together with a rule specifying how long to wait for the next observation if continuation is preferred. Imposing a positive cost for each observation gives a discrete structure to the sequential hypothesis testing problem, and we hence analyze it using a certain operator closely associated with the discrete structure of the setup. Our main result states that the value function of the problem can be characterized as the unique fixed point of this operator and that the value function can be determined by an iterative procedure involving the operator. In the iterative construction of the value function, each element in the sequence has a natural interpretation as the value function of a problem with only finitely many observation rights. Moreover, we show that the optimal strategy can be described in terms of the value function. As expected, the optimal strategy consists of a decision rule whether to stop or not at a given observation time, together with a rule that specifies when to make the next observation. The distribution of the next observation time is described by a function of the current posteriori probability process. A numerical study suggests that in the iterative procedure, the sequence of optimal strategies is convergent, but we have not been able to verify this analytically.

The formulation of the problem with fixed observation costs has direct applications in experimental design, where the cost of setting up an experiment is proportional to the number of trials (with coefficient  $c$  in the notation below), and the cost of analyzing an experiment ( $d$  in the notation below) is independent of the number of trials performed. However, while formulated for the hypothesis testing problem, the general methodology of the current article should be applicable in other optimal stopping problems where each observation is costly. To the best of our knowledge, no such optimal stopping problem has been studied in the literature.

The current article is organized as follows. In Section 2, we formulate the sequential hypothesis testing problem for a Wiener process with costly observations under consideration. In Section 3, we introduce a closely associated operator and we study its properties. In particular, we show that the value function is characterized as its unique fixed point. Finally, in Section 4, we show that an optimal decision rule can be described in terms of the value function.

## 2. Problem formulation

Let

$$X_t = \mu t + \sigma W_t$$

be a stochastic process, where  $W$  is a standard Brownian motion,  $\sigma \neq 0$  is a known constant, and the drift  $\mu$  is an unknown constant. Consider a situation in which one wants to determine  $\mu$  from observations of  $X$  as *accurately* as possible and at the same time as *quickly* as possible. In a Bayesian setting, the uncertainty about the drift is captured by modeling  $\mu$  as a random variable with a given prior distribution, and the Bayes risk is defined as the sum of the risk of a large error in the estimate for the drift and the cost of time. In a classical version of the sequential testing problem, the unknown drift can only take values in the set  $\{\mu_1, \mu_2\}$  where  $\mu_1 \neq \mu_2$  are two given constants, and the Bayes risk associated with a strategy  $(\tau, d)$  is specified as

$$R(\tau, d) = \mathbb{E} \left[ a \mathbb{1}_{\{d=\mu_2, \mu=\mu_1\}} + b \mathbb{1}_{\{d=\mu_1, \mu=\mu_2\}} + c\tau \right].$$

Here  $\tau$  is an  $\mathcal{F}^X$ -stopping time, where  $\mathcal{F}^X = \{\mathcal{F}_t^X, t \geq 0\}$  is the filtration generated by the process  $X$ ,  $d$  is an  $\mathcal{F}_\tau^X$ -measurable random variable,  $a > 0$  and  $b > 0$  are the costs for the two possible kinds of faulty decisions, and  $c > 0$  is the observation cost per unit of time. Introducing the *a posteriori probability process*

$$\Pi_t := \mathbb{P}(\mu = \mu_2 | \mathcal{F}_t^X), \quad (2.1)$$

following standard lines of argument, gives that the minimal Bayes risk is given by

$$U(\pi) = \inf_{\tau} \mathbb{E}_{\pi} \left[ g(\Pi_{\tau}) + c\tau \right], \quad (2.2)$$

where  $g(\pi) := a\pi \wedge b(1 - \pi)$ . It is well known that the a posteriori probability process satisfies

$$d\Pi_t = \omega \Pi_t (1 - \Pi_t) d\hat{W}_t,$$

where  $\omega = (\mu_2 - \mu_1)/\sigma$  denotes the *signal-to-noise ratio* and the innovation process

$$\hat{W}_t := \frac{X_t}{\sigma} - \omega \int_0^t \Pi_s ds - \frac{\mu_1}{\sigma} t$$

is a standard Brownian motion. Moreover,  $\Pi$  is a (time-homogeneous) strong Markov process with respect to its natural filtration, which coincides with  $\{\mathcal{F}_t^X, t \geq 0\}$ . It is well known that the function  $U$  defined in (2.2) can be determined as the solution of an associated free-boundary problem; see, for example, Shiryaev (1969, 1978).

We consider a similar hypothesis testing problem but with the added constraint that each observation is associated with a fixed cost. To formulate the problem, let  $\hat{\tau} = \{\tau_k\}_{k=0}^{\infty}$  be an increasing sequence of random times with  $\tau_0 = 0$ , and let

$$\mathcal{F}_t^{\hat{\tau}} = \sigma \left\{ (\tau_1, X_{\tau_1}), (\tau_2, X_{\tau_2}), \dots, (\tau_k, X_{\tau_k}) \text{ where } k = \sup\{j : \tau_j \leq t\} \right\}.$$

We only consider sequences  $\hat{\tau} = \{\tau_k\}_{k=0}^{\infty}$  such that  $\tau_k$  is a predictable  $\mathcal{F}^{\hat{\tau}}$ -stopping time. Note that, due to the discrete structure,  $\tau_k$  is a predictable  $\mathcal{F}^{\hat{\tau}}$ -stopping time precisely if  $\tau_k$  is  $\mathcal{F}_{\tau_{k-1}}^{\hat{\tau}}$ -measurable, so, in particular,  $\tau_1$  is deterministic. A pair  $(\hat{\tau}, \tau)$ , where  $\hat{\tau} = \{\tau_k\}_{k=0}^{\infty}$  is as described above and  $\tau$  is an  $\mathcal{F}^{\hat{\tau}}$ -stopping time with  $\tau(\omega) \in \{\tau_0(\omega), \tau_1(\omega), \tau_2(\omega), \dots\}$  a.s. is called an *admissible strategy*, and the set of admissible strategies is denoted  $\mathcal{T}$ .

Define the value function of the sequential hypothesis testing problem with costly observations to be

$$V(\pi) = \inf_{(\hat{\tau}, \tau) \in \mathcal{T}} \mathbb{E}_{\pi} \left[ g(\Pi_{\tau}) + c\tau + d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \leq \tau\}} \right]. \quad (2.3)$$

Here the constant  $d > 0$  represents the cost of each observation.

**Remark 2.1.** Note that  $U \leq V \leq g$  is immediate from the definition. Also note that an implicit consequence of the definition of  $\mathcal{T}$  is that stopping is only allowed at observation times. This is without loss of generality, since stopping between observation times would necessarily be suboptimal as no more information is obtained in such intervals.

### 3. Analysis of the value function

In this section, we introduce an operator that is closely associated with the sequential hypothesis testing problem (2.3), and we study its properties. Let

$$\mathbb{F} := \{f : [0, 1] \rightarrow [0, \max\{a, b\}] : f \text{ concave}, U \leq f \leq g\},$$

where  $U$  is the value function of the classical hypothesis testing problem defined in (2.2) above. Consider the operator  $\mathcal{J}$  defined by

$$(\mathcal{J}f)(\pi) = \min \left\{ g(\pi), d + \inf_t \{ct + \mathbb{E}_\pi [f(\Pi_t)]\} \right\}$$

for any given function  $f \in \mathbb{F}$ .

**Lemma 3.1.** *Let  $f \in \mathbb{F}$  and  $\pi \in [0, 1]$ . Then the function  $t \mapsto ct + \mathbb{E}_\pi [f(\Pi_t)]$  attains its minimum for some point  $t \in [0, \infty)$ .*

*Proof.* For fixed  $\pi \in [0, 1]$  and  $f \in \mathbb{F}$ , the function  $F(t) := ct + \mathbb{E}_\pi [f(\Pi_t)]$  satisfies  $F(0) = f(\pi)$  and  $\lim_{t \rightarrow \infty} F(t) = \infty$ . Thus, since  $F$  is continuous, its infimum is attained at some  $t \geq 0$ .

In view of Lemma 3.1, we define the function  $t(\cdot; f) : [0, 1] \rightarrow [0, \infty)$  for any  $f \in \mathbb{F}$  by

$$t(\pi; f) = \inf \left\{ t \geq 0 : \inf_s \{cs + \mathbb{E}_\pi [f(\Pi_s)]\} = ct + \mathbb{E}_\pi [f(\Pi_t)] \right\}, \quad (3.1)$$

for  $\pi \in [0, 1]$ . In other words,  $t(\pi; f)$  is the first time at which the function  $s \mapsto cs + \mathbb{E}_\pi [f(\Pi_s)]$  attains its minimum.

**Lemma 3.2.** (a) *If  $f_1, f_2 \in \mathbb{F}$  satisfy  $f_1 \leq f_2$ , then  $\mathcal{J}f_1 \leq \mathcal{J}f_2$ .*

(b) *If  $f \in \mathbb{F}$ , then  $\mathcal{J}f \in \mathbb{F}$ .*

*Proof.* For  $\pi \in [0, 1]$ , we have that

$$\begin{aligned} \mathcal{J}f_1(\pi) &= \min \left\{ g(\pi), \inf_t \{d + ct + \mathbb{E}_\pi [f_1(\Pi_t)]\} \right\} \\ &\leq \min \left\{ g(\pi), \inf_t \{d + ct + \mathbb{E}_\pi [f_2(\Pi_t)]\} \right\} = \mathcal{J}f_2(\pi), \end{aligned}$$

which proves (a).

For (b), note that by definition,  $\mathcal{J}f(\pi) \leq g(\pi)$ . Moreover, for a fixed  $t$ , the function  $\pi \mapsto d + ct + \mathbb{E}_\pi [f(\Pi_t)]$  is concave (for results on preservation of convexity for martingale diffusions, see, for example, Hobson (1998) or Janson and Tysk (2003)), so therefore  $\mathcal{J}f$  is also concave since it is the pointwise minimum of concave functions. It remains to check that  $U \leq \mathcal{J}f$ .

For this, note that  $U \leq f$ , so

$$\mathcal{J}U \leq \mathcal{J}f \quad (3.2)$$

by (a). Moreover, by standard results in optimal stopping theory, we know that the process  $ct + U(\Pi_t)$  is a submartingale, so  $U(\pi) \leq ct + \mathbb{E}_\pi[U(\Pi_t)]$  for any  $t \geq 0$ . Therefore,

$$U(\pi) = \min\{g(\pi), U(\pi)\} \leq \min\{g(\pi), \inf_t \{ct + \mathbb{E}_\pi[U(\Pi_t)]\}\} \leq \mathcal{J}U(\pi),$$

which together with (3.2) gives (b).

Define the sequence  $f_n$  recursively by  $f_0 = g$  and

$$f_{n+1} = \mathcal{J}f_n, \quad n \geq 1.$$

By Proposition 3.2, the sequence  $\{f_n\}$  is decreasing in  $n$  and thus its limit  $f_\infty := \lim_{n \rightarrow \infty} f_n$  exists. Since the pointwise limit of a sequence of concave functions is concave, we have that  $f_\infty \in \mathbb{F}$ .

**Lemma 3.3.** *The function  $f_\infty \in \mathbb{F}$  is a fixed point of the operator  $\mathcal{J}$ . Moreover, it is the largest fixed point in  $\mathbb{F}$ .*

*Proof.* Since  $f_n \geq f_\infty$ , we have  $f_{n+1} = \mathcal{J}f_n \geq \mathcal{J}f_\infty$  by (a) in Proposition 3.2. Consequently,

$$f_\infty \geq \mathcal{J}f_\infty. \quad (3.3)$$

For the opposite inequality, fix  $\pi \in [0, 1]$  and let  $t_\infty = t(\pi; f_\infty)$ , where  $t(\pi; f_\infty)$  is defined as in (3.1). Then

$$f_{n+1}(\pi) = \mathcal{J}f_n(\pi) \leq \min\{g(\pi), d + ct_\infty + \mathbb{E}_\pi[f_n(\Pi_{t_\infty})]\},$$

so letting  $n \rightarrow \infty$  yields

$$f_\infty(\pi) \leq \min\{g(\pi), d + ct_\infty + \mathbb{E}_\pi[f_\infty(\Pi_{t_\infty})]\} = \mathcal{J}f_\infty(\pi)$$

by monotone convergence. Together with (3.3), this shows that  $f_\infty$  is a fixed point.

Finally, assume that  $h \in \mathbb{F}$  is another fixed point of  $\mathcal{J}$ . Then  $f_0 = g \geq h$ , and using (a) in Proposition 3.2, an easy induction argument shows that  $f_n \geq h$ . Consequently,  $f_\infty \geq h$ , which finishes the proof.

Define the function  $V_n : [0, 1] \rightarrow [0, \infty)$  by

$$\begin{aligned} V_n(\pi) &= \inf_{(\hat{\tau}, \tau) \in \mathcal{T}: \tau \leq \tau_n} \mathbb{E}_\pi \left[ g(\Pi_\tau) + c\tau + d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \leq \tau\}} \right] \\ &= \inf_{(\hat{\tau}, \tau) \in \mathcal{T}} \mathbb{E}_\pi \left[ g(\Pi_{\tau \wedge \tau_n}) + c\tau \wedge \tau_n + d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \leq \tau \wedge \tau_n\}} \right] \end{aligned} \quad (3.4)$$

and note that  $V_n$  then is the value function of a version of our hypothesis testing problem where the underlying process may be observed at most  $n$  times.

**Theorem 3.1.** *We have  $V_n = f_n$ ,  $n \geq 0$ .*

*Proof.* First note that  $V_0 = f_0 = g$  by definition. Assume that  $V_{n-1} = f_{n-1}$  for some  $n \geq 1$ , and fix  $\pi \in (0, 1)$  and  $(\hat{\tau}, \tau) \in \mathcal{T}$ . Let  $\tau'_k := \tau_{k+1} - \tau_1$  and set  $\tau' := \tau - \tau_1$  on the set where  $\tau \geq \tau_1$ . By the Markov property, the definition of  $V_{n-1}$ , and the induction hypothesis, we have

$$\begin{aligned}
& \mathbb{E}_\pi \left[ g(\Pi_{\tau \wedge \tau_n}) + c(\tau \wedge \tau_n) + d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \leq \tau \wedge \tau_n\}} \right] \\
&= \mathbb{E}_\pi \left[ \mathbb{1}_{\{\tau=0\}} \left( g(\Pi_{\tau \wedge \tau_n}) + c(\tau \wedge \tau_n) + d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \leq \tau \wedge \tau_n\}} \right) \right] \\
&\quad + \mathbb{E}_\pi \left[ \mathbb{1}_{\{\tau \geq \tau_1\}} (c\tau_1 + d) \right] \\
&\quad + \mathbb{E}_\pi \left[ \mathbb{1}_{\{\tau \geq \tau_1\}} \mathbb{E}_{\Pi_{\tau_1}} \left[ g(\Pi_{\tau' \wedge \tau'_{n-1}}) + c(\tau' \wedge \tau'_{n-1}) + d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau'_k \leq \tau' \wedge \tau'_{n-1}\}} \right] \right] \\
&\geq \mathbb{1}_{\{\tau=0\}} g(\pi) + \mathbb{1}_{\{\tau \geq \tau_1\}} \mathbb{E}_\pi \left[ c\tau_1 + d + V_{n-1}(\Pi_{\tau_1}) \right] \\
&= \mathbb{1}_{\{\tau=0\}} g(\pi) + \mathbb{1}_{\{\tau \geq \tau_1\}} \mathbb{E}_\pi \left[ c\tau_1 + d + f_{n-1}(\Pi_{\tau_1}) \right] \\
&\geq \min \left\{ g(\pi), \inf_{t \geq 0} \{ ct + d + \mathbb{E}_\pi [f_{n-1}(\Pi_t)] \} \right\} = f_n(\pi).
\end{aligned}$$

Taking the infimum over strategies yields  $V_n \geq f_n$ .

For the reverse inequality, fix  $\pi \in (0, 1)$  and let  $t_n := t(\pi; f_{n-1})$ . If  $ct_n + d + \mathbb{E}_\pi[f_{n-1}(\Pi_{t_n})] \geq g(\pi)$ , then  $f_n(\pi) = \mathcal{J}f_{n-1}(\pi) = g(\pi) \geq V_n(\pi)$ . Thus, we may assume that  $ct_n + d + \mathbb{E}_\pi[f_{n-1}(\Pi_{t_n})] < g(\pi)$  so that  $(\mathcal{J}f_{n-1})(\pi) = ct_n + d + \mathbb{E}_\pi[f_{n-1}(\Pi_{t_n})]$ . For a given  $\epsilon > 0$ , let  $\hat{\tau} = \{\tau_k\}_{k=0}^{\infty}$  and  $\tau$  be  $\epsilon$ -optimal in  $V_{n-1}(\Pi_{t_n})$  so that  $\tau \leq \tau_{n-1}$  and

$$V_{n-1}(\Pi_{t_n}) \geq \mathbb{E}_{\Pi_{t_n}} \left[ g(\Pi_\tau) + c\tau + d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \leq \tau\}} \right] - \epsilon.$$

Now, with  $\tau' = t_n + \tau$  and  $\tau'_{k+1} = t_n + \tau_k$ ,  $k \geq 0$ , we have that

$$\begin{aligned}
f_n(\pi) + \epsilon &= ct_n + d + \mathbb{E}_\pi \left[ V_{n-1}(\Pi_{t_n}) \right] + \epsilon \\
&\geq ct_n + d + \mathbb{E}_\pi \left[ \mathbb{E}_{\Pi_{t_n}} \left[ g(\Pi_\tau) + c\tau + d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \leq \tau\}} \right] \right] \\
&= \mathbb{E}_\pi \left[ g(\Pi_{\tau'}) + c\tau' + d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau'_k \leq \tau'\}} \right] \geq V_n(\pi).
\end{aligned}$$

Since  $\pi$  and  $\epsilon > 0$  are arbitrary, we find that  $V_n \leq f_n$ , which completes the proof.

**Theorem 3.2.** *The value function  $V$  satisfies  $V = f_\infty$ . Consequently,  $V$  is the largest fixed point in  $\mathbb{F}$  of the operator  $\mathcal{J}$ .*

*Proof.* In view of Lemma 3.3 and Theorem 3.1, it suffices to prove that

$$\lim_{n \rightarrow \infty} V_n(\pi) = V(\pi).$$

To do that, first note that  $V(\pi) \leq V_{n+1}(\pi) \leq V_n(\pi)$  for any  $\pi$  and all  $n \geq 0$ . Consequently, it suffices to prove that  $\lim_{n \rightarrow \infty} V_n(\pi) \leq V(\pi)$ . Fix  $\pi \in (0, 1)$ , take  $\epsilon > 0$ , and let  $(\hat{\tau}, \tau) \in \mathcal{T}$  be  $\epsilon$ -optimal in  $V(\pi)$ ; that is

$$V(\pi) \geq \mathbb{E}_\pi \left[ g(\Pi_\tau) + c\tau + d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \leq \tau\}} \right] - \epsilon. \quad (3.5)$$

Then

$$\begin{aligned} V_n(\pi) &\leq \mathbb{E}_\pi \left[ g(\Pi_{\tau \wedge \tau_n}) + c(\tau \wedge \tau_n) + d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \leq \tau \wedge \tau_n\}} \right] \\ &\leq \mathbb{E}_\pi \left[ g(\Pi_{\tau \wedge \tau_n}) + c\tau + d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \leq \tau\}} \right]. \end{aligned} \quad (3.6)$$

Since

$$\begin{aligned} V(\pi) + \epsilon &\geq \mathbb{E}_\pi \left[ d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \leq \tau\}} \right] = \mathbb{E}_\pi \left[ d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \leq \tau\}} | \tau \leq \tau_n \right] \mathbb{P}(\tau \leq \tau_n) \\ &\quad + \mathbb{E}_\pi \left[ d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k \leq \tau\}} | \tau > \tau_n \right] \mathbb{P}(\tau > \tau_n), \end{aligned}$$

we have

$$\mathbb{P}(\tau > \tau_n) \leq \frac{V(\pi) + \epsilon}{nd} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Consequently, since  $g$  is bounded,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}_\pi [g(\Pi_{\tau \wedge \tau_n})] &= \lim_{n \rightarrow \infty} (\mathbb{E}_\pi [g(\Pi_\tau) \mathbb{1}_{\{\tau \leq \tau_n\}}] + \mathbb{E}_\pi [g(\Pi_{\tau_n}) \mathbb{1}_{\{\tau > \tau_n\}}]) \\ &= \mathbb{E}_\pi [g(\Pi_\tau)] \end{aligned}$$

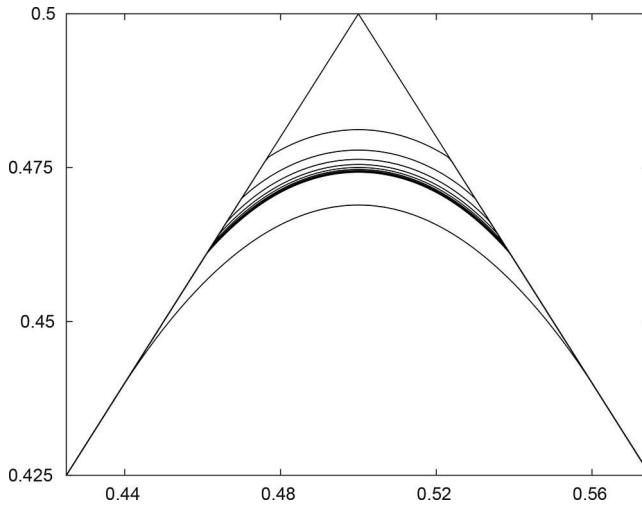
by dominated convergence. Thus, by (3.5) and (3.6) we get

$$\lim_{n \rightarrow \infty} V_n(\pi) \leq V(\pi) + \epsilon,$$

and since  $\epsilon > 0$  is arbitrary, this completes the proof.

**Remark 3.1.** For a graphical illustration of the convergence of the sequence  $\{V_n\}_{n=0}^{\infty}$ , see Figure 1. We point out that while it is well known that the value function  $U$  from the classical sequential hypothesis testing problem with continuous observations satisfies the smooth-fit condition at the boundary points of the continuation region, there is no reason to expect smooth fit for the value functions  $V_n$  or  $V$ . In fact, Figure 1 suggests that smooth fit fails in the case of discrete observation costs.





**Figure 1.** The value  $V_n$  for  $n = 0, \dots, 10$  (in decreasing order) and  $U$  (lowest one), for  $a = b = 1, c = 1, d = 0.001, \mu_2 - \mu_1 = 1$ , and  $\sigma = \sqrt{2}/2$ .

**Remark 3.2.** It follows from Theorem 3.2 that the value function  $V$  is decreasing in the signal-to-noise ratio  $\omega = (\mu_2 - \mu_1)/\sigma$ . Indeed, for given signal-to-noise ratios  $\omega$  and  $\tilde{\omega}$  satisfying  $\omega \leq \tilde{\omega}$ , denote by  $V, \tilde{V}, V_n$ , and  $\tilde{V}_n$  the corresponding value functions. Then  $V_0 = g = \tilde{V}_0$ . Moreover, if  $V_n \geq \tilde{V}_n$  for some  $n \geq 0$ , then by general monotonicity results with respect to the diffusion coefficient (see Hobson (1998) and Janson and Tysk (2003)) one has

$$V_{n+1} = \mathcal{J}V_n \geq \tilde{\mathcal{J}}V_n \geq \tilde{\mathcal{J}}\tilde{V}_n = \tilde{V}_{n+1},$$

where  $\mathcal{J}$  and  $\tilde{\mathcal{J}}$  are the corresponding operators. By induction, it follows that  $V_n \geq \tilde{V}_n$  for all  $n \geq 0$ , so

$$V = \lim_{n \rightarrow \infty} V_n \geq \lim_{n \rightarrow \infty} \tilde{V}_n = \tilde{V}.$$

One can show that the operator  $\mathcal{J}$  fails to be a contraction on  $\mathbb{F}$  (equipped with the sup-norm), so we cannot use the Banach fix-point theorem to establish uniqueness of fix-points or deduce convergence rates for the convergence  $V_n \rightarrow V$ . Instead, we end this section by showing that  $V$  is the unique fix-point using a more direct method.

**Theorem 3.3.**  $V$  is the unique fixed point of  $\mathcal{J}$ .

*Proof.* Define a second sequence  $\{\tilde{f}_n\}_{n=0}^\infty$  in  $\mathbb{F}$  recursively by  $\tilde{f}_0 = U$  and

$$\tilde{f}_{n+1} = \mathcal{J}\tilde{f}_n, \quad n \geq 0.$$

By Proposition 3.2 (b),  $\tilde{f}_1 \geq U = \tilde{f}_0$ , so an induction argument using Proposition 3.2 (a) shows that  $\tilde{f}_{n+1} \geq \tilde{f}_n$  for all  $n \geq 0$ . Also, define the function  $\tilde{V}_n$  by

$$\tilde{V}_n(\pi) = \inf_{\substack{(\hat{\tau}, \tau) \in \mathcal{T} \\ \tau \leq \tau_n}} \mathbb{E}_\pi \left[ g(\Pi_\tau) \mathbb{1}_{\{\tau < \tau_n\}} + U(\Pi_\tau) \mathbb{1}_{\{\tau \geq \tau_n\}} + c\tau + d \sum_{k=1}^\infty \mathbb{1}_{\{\tau_k \leq \tau\}} \right]$$

and note that  $\tilde{V}_n$  then is the value when the underlying process may be observed at most  $n$  times given that if no stopping has occurred then one receives the function  $U$  at the  $n$ th observation time. Using similar arguments as in the proofs of Theorems 3.1–3.2 above, we find that

$$\tilde{f}_n = \tilde{V}_n$$

and

$$\lim_{n \rightarrow \infty} \tilde{V}_n(\pi) = V(\pi).$$

Consequently,

$$\lim_{n \rightarrow \infty} \tilde{f}_n(\pi) = V(\pi). \quad (3.7)$$

Now, assume that  $\hat{V} \in \mathbb{F}$  is a fixed point of  $\mathcal{J}$ . Then, by definition of  $\mathbb{F}$ ,  $\hat{V} \geq U = \tilde{f}_0$ . Consequently,  $\hat{V} = \mathcal{J}\hat{V} \geq \mathcal{J}\tilde{f}_0 = \tilde{f}_1$ , and an induction argument gives that  $\hat{V} \geq \tilde{f}_n$  for all  $n \geq 0$ . By (3.7), this implies that  $\hat{V} \geq V$ , which, in view of Theorem 3.2, completes the proof.

**Remark 3.3.** It follows from the analysis above that even though  $\mathcal{J}$  is not a contraction, the sequence  $\{\hat{f}_n\}_{n=0}^{\infty}$  defined by  $f_0 = f$  and  $f_{n+1} = \mathcal{J}f_n$ ,  $n \geq 0$ , converges to  $V$  for any starting point  $f \in \mathbb{F}$ .

#### 4. The optimal strategy

In Section 3, we characterized the value function  $V$  as the unique fixed point of the operator  $\mathcal{J}$  (Theorem 3.3). Moreover, this fixed point can be determined using an iterative procedure; see Theorem 3.2. Given the value function  $V$ , there is a natural way to define a corresponding strategy. In the current section, we show that this strategy is indeed optimal.

Since the value function  $V$  is concave, the set  $I := \{\pi \in [0, 1] : V(\pi) < g(\pi)\}$  is an open interval; in the case when  $I \neq \emptyset$ , we thus have  $I = (A, B)$ , where

$$A := \inf\{\pi : V(\pi) < g(\pi)\}$$

and

$$B := \sup\{\pi : V(\pi) < g(\pi)\}$$

denote the end-points of  $I$ . For  $\pi \in I$ , we have

$$V(\pi) = \inf_{t \geq 0} \{ct + d + \mathbb{E}_{\pi}[V(\Pi_t)]\},$$

and the infimum is attained for the first time at  $t(\pi; V)$  defined as in (3.1). Let  $t(\pi) := t(\pi; V)$  for  $\pi \in I$  and set  $t(\pi) = 0$  for  $\pi \notin I$ . Since

$$ct + d + \mathbb{E}_{\pi}[V(\Pi_t)]|_{t=0} = d + V(\pi) > V(\pi),$$

we have  $t(\pi) > 0$  for  $\pi \in I$ . Now define the sequence  $\hat{\tau}^* = \{\tau_k^*\}_{k=0}^{\infty}$  recursively by setting  $\tau_0^* = 0$  and

$$\tau_{k+1}^* = \tau_k^* + t(\Pi_{\tau_k^*})$$

for  $k = 0, \dots, n^* - 1$ , where  $n^* = \min\{k : \Pi_{\tau_k^*} \notin I\}$ , and  $\tau_k^* = \infty$  for  $k \geq n^* + 1$ , and let  $\tau^* = \tau_{n^*}^*$ . We say that the strategy  $(\hat{\tau}^*, \tau^*)$  is the *strategy associated with  $V$* .

**Theorem 4.1.** *The strategy  $(\hat{\tau}^*, \tau^*)$  associated with  $V$  is optimal in (2.3).*

*Proof.* Let  $(\hat{\tau}^*, \tau^*)$  be the strategy associated with  $V$  and define the function  $\hat{V}$  by

$$\hat{V}(\pi) = \mathbb{E}_\pi \left[ g(\Pi_{\tau^*}) + c\tau^* + d \sum_{k=1}^{\infty} \mathbb{1}_{\{\tau_k^* \leq \tau^*\}} \right] = \mathbb{E}_\pi [g(\Pi_{\tau^*}) + c\tau^* + dn^*].$$

By definition,  $V \leq \hat{V}$ .

To prove the reverse inequality, we first claim that the dynamic programming principle relation

$$V(\pi) = \mathbb{E}_\pi \left[ (g(\Pi_{\tau^*}) + c\tau^* + dn^*) \mathbb{1}_{\{n^* \leq n\}} + (c\tau_n^* + dn + V(\Pi_{\tau_n^*})) \mathbb{1}_{\{n^* > n\}} \right] \quad (4.1)$$

holds for any  $n \geq 0$ . To see this, first note that for  $n = 0$ , the right-hand side equals  $g(\pi) \mathbb{1}_{\{n^*=0\}} + V(\pi) \mathbb{1}_{\{n^*>0\}} = V(\pi)$  by the definition of  $n^*$ , so (4.1) holds for  $n = 0$ . Next, note that by the Markov property of  $\Pi$  and the definition of  $n^*$ , we have

$$\mathbb{E}_\pi [g(\Pi_{\tau^*}) \mathbb{1}_{\{n^*=n+1\}}] = \mathbb{E}_\pi \left[ \mathbb{E}_{\Pi_{\tau_n^*}} \left[ V(\Pi_{\tau_1^*}) \mathbb{1}_{\{n^*=1\}} \right] \mathbb{1}_{\{n^*>n\}} \right]$$

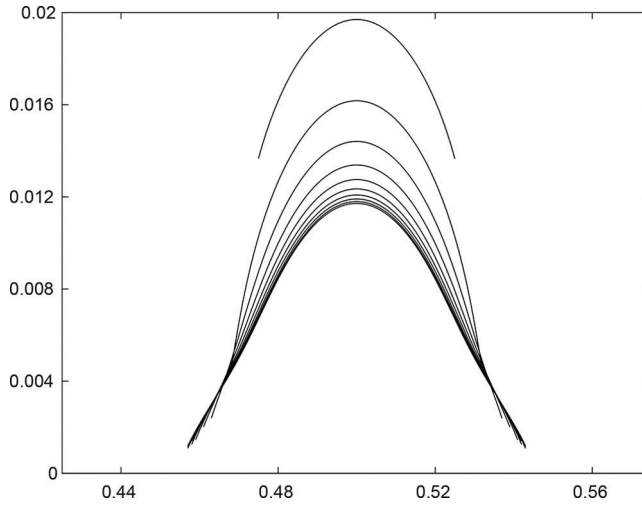
and

$$\mathbb{E}_\pi \left[ V(\Pi_{\tau_{n+1}^*}) \mathbb{1}_{\{n^*>n+1\}} \right] = \mathbb{E}_\pi \left[ \mathbb{E}_{\Pi_{\tau_n^*}} \left[ V(\Pi_{\tau_1^*}) \mathbb{1}_{\{n^*>1\}} \right] \mathbb{1}_{\{n^*>n\}} \right].$$

Using the equations above in the first step, the definition of  $(\hat{\tau}^*, \tau^*)$  in the second and third, and Theorem 3.2 in the final step, the right-hand side of (4.1) for  $n + 1$  satisfies

$$\begin{aligned} & \mathbb{E}_\pi \left[ (c\tau^* + dn^* + g(\Pi_{\tau^*})) \mathbb{1}_{\{n^* \leq n+1\}} \right] \\ & \quad + \mathbb{E}_\pi \left[ (c\tau_{n+1}^* + d(n+1) + V(\Pi_{\tau_{n+1}^*})) \mathbb{1}_{\{n^* > n+1\}} \right] \\ & = \mathbb{E}_\pi \left[ (c\tau^* + dn^* + g(\Pi_{\tau^*})) \mathbb{1}_{\{n^* \leq n\}} \right] \\ & \quad + \mathbb{E}_\pi \left[ (c\tau^* + dn^*) \mathbb{1}_{\{n^*=n+1\}} + (c\tau_{n+1}^* + d(n+1)) \mathbb{1}_{\{n^* > n+1\}} \right] \\ & \quad + \mathbb{E}_\pi \left[ \mathbb{E}_{\Pi_{\tau_n^*}} \left[ V(\Pi_{\tau_1^*}) \mathbb{1}_{\{n^*=1\}} \right] \mathbb{1}_{\{n^* > n\}} \right] \\ & \quad + \mathbb{E}_\pi \left[ \mathbb{E}_{\Pi_{\tau_n^*}} \left[ V(\Pi_{\tau_1^*}) \mathbb{1}_{\{n^*>1\}} \right] \mathbb{1}_{\{n^* > n\}} \right] \\ & = \mathbb{E}_\pi \left[ (c\tau^* + dn^* + g(\Pi_{\tau^*})) \mathbb{1}_{\{n^* \leq n\}} \right] \\ & \quad + \mathbb{E}_\pi \left[ (c\tau_n^* + ct(\Pi_{\tau_n^*}) + d(n+1) + \mathbb{E}_{\Pi_{\tau_n^*}} \left[ V(\Pi_{\tau_1^*}) \right]) \mathbb{1}_{\{n^* > n\}} \right] \\ & = \mathbb{E}_\pi \left[ (c\tau^* + dn^* + g(\Pi_{\tau^*})) \mathbb{1}_{\{n^* \leq n\}} + (c\tau_n^* + dn + \mathcal{J}V(\Pi_{\tau_n^*})) \mathbb{1}_{\{n^* > n\}} \right] \\ & = \mathbb{E}_\pi \left[ (c\tau^* + dn^* + g(\Pi_{\tau^*})) \mathbb{1}_{\{n^* \leq n\}} + (c\tau_n^* + dn + V(\Pi_{\tau_n^*})) \mathbb{1}_{\{n^* > n\}} \right], \end{aligned}$$

which is the right-hand side of (4.1) for  $n$ . Thus, (4.1) holds for all  $n \geq 0$  by induction.



**Figure 2.** The waiting times  $t_n(\pi) := t(\pi; V_{n-1})$  for  $n = 1, \dots, 10$ , for  $a = b = 1$ ,  $c = 1$ ,  $d = 0.001$ ,  $\mu_2 - \mu_1 = 1$ , and  $\sigma = \sqrt{2}/2$ .

Now, by (4.1) we have that

$$\|g\|_\infty \geq V(\pi) \geq \mathbb{E}_\pi \left[ (c\tau^* + dn^* + g(\Pi_{\tau^*})) \mathbb{1}_{\{n^* < n\}} + dn \mathbb{1}_{\{n^* \geq n\}} \right]$$

and thus

$$\mathbb{P}(n^* \geq n) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Consequently,  $n^* < \infty$  a.s. By (4.1) and monotone convergence, we have

$$V(\pi) \geq \mathbb{E}_\pi \left[ (g(\Pi_{\tau^*}) + c\tau^* + dn^*) \mathbb{1}_{\{n^* \leq n\}} \right] \rightarrow \hat{V}(\pi)$$

as  $n \rightarrow \infty$ . Thus,  $\hat{V} \leq V$ , which completes the proof.

**Remark 4.1.** Given  $n \geq 0$ , consider the strategy defined recursively by  $\tau_0^n = 0$  and

$$\tau_{k+1}^n = \tau_k^n + t(\Pi_{\tau_k^n}; V_{n-k-1})$$

for  $k = 0, \dots, n^* - 1$ , where  $n^* = \min\{k : V_{n-k}(\Pi_{\tau_k^n}) = g(\Pi_{\tau_k^n})\}$ , and  $\tau_k^n = \infty$  for  $k \geq n^* + 1$ , and let  $\tau^n = \tau_{n^*}^n$ . Employing similar methods as the ones used in the proof of Theorem 4.1 shows that the strategy  $(\hat{\tau}^n, \tau^n)$ , where  $\hat{\tau}^n = \{\tau_k^n\}_{k=0}^\infty$ , is optimal for the problem  $V_n$  defined in (3.4). Since  $V_n \rightarrow V$  by Theorem 3.2, it seems reasonable to expect that  $t(\cdot; V_n)$  would converge to  $t(\cdot, V)$  and thus that the optimal strategy  $(\hat{\tau}^n, \tau^n)$  would tend to  $(\hat{\tau}^*, \tau^*)$ . While numerical evidence supports this, compare Figure 2, we have not been able to confirm it analytically.

## Acknowledgments

We thank the Associate Editor and two anonymous referees for their suggestions to improve the article.

## Funding

The article was funded by Vetenskapsrådet Grant #622-2008-447.

## ORCID

Hannah Dyrssen  <http://orcid.org/0000-0002-6162-906X>

## References

- Bather, J. (1962). Bayes Procedures for Deciding the Sign of a Normal Mean, *Proceedings of Cambridge Philosophical Society* 58: 599–620.
- Breakwell, J. and Chernoff, H. (1964). Sequential Tests for the Mean of a Normal Distribution. II. (Large T), *Annals of Mathematical Statistics* 35: 162–173.
- Buonaguidi, B. and Muliere, P. (2013). Sequential Testing Problems for Lévy Processes, *Sequential Analysis. Design Methods & Applications* 32: 47–70.
- Buonaguidi, B. and Muliere, P. (2016). Optimal Sequential Testing for an Inverse Gaussian Process, *Sequential Analysis* 35: 69–83.
- Chernoff, H. (1961). Sequential Tests for the Mean of a Normal Distribution, in *1961 Proceedings of 4th Berkeley Symposium in Mathematical Statistics and Probability*, vol. 1: pp. 79–91, Berkeley: University of California Press.
- Chernoff, H. (1965). Sequential Tests for the Mean of a Normal Distribution. III. (Small T), *Annals of Mathematical Statistics* 36: 28–54.
- Dayanik, S. and Sezer, S. (2006). Sequential Testing of Simple Hypotheses about Compound Poisson Processes, *Stochastic Processes and Their Applications* 116: 1892–1919.
- Décamps, J.-P., Mariotti, T., and Villeneuve, S. (2005). Investment Timing under Incomplete Information, *Mathematics of Operations Research* 30: 472–500.
- Ekström, E. and Vaicenavicius, J. (2015). Bayesian Sequential Testing of the Drift of a Brownian Motion, *ESAIM. Probability and Statistics* 19: 626–648.
- Ekström, E. and Vaicenavicius, J. (2016). Optimal Liquidation of an Asset under Drift Uncertainty, *SIAM Journal of Financial Mathematics* 7: 357–381.
- Gapeev, P. and Peskir, G. (2004). The Wiener Sequential Testing Problem with Finite Horizon, *Stochastics and Stochastics Reports* 76: 59–75.
- Hobson, D. (1998). Volatility Misspecification, Option Pricing and Superreplication Via Coupling, *Annals of Applied Probability* 8: 193–205.
- Janson, S. and Tysk, J. (2003). Volatility Time and Properties of Option Prices, *Annals of Applied Probability* 13: 890–913.
- Lakner, P. (1995). Utility Maximization with Partial Information, *Stochastic Processes and Their Applications* 56: 247–273.
- Peskir, G. and Shiryaev, A. (2000). Sequential Testing Problems for Poisson Processes, *Annals of Statistics* 28: 837–859.
- Shiryaev, A. (1969). Two Problems of Sequential Analysis, *Cybernetics* 3: 63–69.
- Shiryaev, A. (1978). *Optimal Stopping Rules, Applications of Mathematics*, vol. 8, New York: Springer.
- Zhitlukhin, M. and Muravlev, A. (2013). On Chernoff's Hypothesis Testing Problem for the Drift of a Brownian Motion, *Theory of Probability and Its Applications* 57: 708–717.
- Zhitlukhin, M. and Shiryaev, A. (2011). A Bayesian Sequential Testing Problem of Three Hypotheses for Brownian Motion, *Statistics & Risk Modeling with Applications in Finance and Insurance* 28: 227–249.