



## Compound Poisson Disorder Problems with Nonlinear Detection Delay Penalty Cost Functions

Savas Dayanik

To cite this article: Savas Dayanik (2010) Compound Poisson Disorder Problems with Nonlinear Detection Delay Penalty Cost Functions, Sequential Analysis, 29:2, 193-216, DOI: [10.1080/07474941003741227](https://doi.org/10.1080/07474941003741227)

To link to this article: <https://doi.org/10.1080/07474941003741227>



Published online: 17 May 2010.



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



Article views: 44



Citing articles: 6 View citing articles [↗](#)

---

## Compound Poisson Disorder Problems with Nonlinear Detection Delay Penalty Cost Functions

**Savas Dayanik**

Department of Operations Research and Financial Engineering,  
Princeton University, New Jersey, USA and  
Departments of Industrial Engineering and Mathematics,  
Bilkent University, Ankara, Turkey

**Abstract:** The quickest detection of the unknown and unobservable disorder time, when the arrival rate and mark distribution of a compound Poisson process suddenly changes, is formulated in a Bayesian setting, where the detection delay penalty is a general smooth function of the detection delay time. Under suitable conditions, the problem is shown to be equivalent to the optimal stopping of a finite-dimensional piecewise-deterministic strongly Markov sufficient statistic. The solution of the optimal stopping problem is described in detail for the compound Poisson disorder problem with polynomial detection delay penalty function of arbitrary but fixed degree. The results are illustrated for the case of the quadratic detection delay penalty function.

**Keywords:** Bayesian sequential change detection; Compound Poisson disorder problem; Optimal stopping; Piecewise-deterministic Markov processes.

**Subject Classifications:** 62L10; 62L15; 62C10; 60G40.

### 1. INTRODUCTION

Suppose that the arrival rate and mark distribution of a compound Poisson process changes at some unknown and unobservable disorder time. We would like to detect the disorder time by a stopping rule that depends only on the observations of the point process and that minimizes the total risk arising from frequent false alarms and long detection delay times.

The disorder time is assumed to follow a zero-modified exponential distribution. The formulation of the problem is Bayesian, and for each stopping time of point

Received March 29, 2009, Revised June 20, 2009, Accepted June 25, 2009

Recommended by A. G. Tartakovsky

Address correspondence to Savas Dayanik, Department of Industrial Engineering, Bilkent University, Bilkent, Ankara 06800, Turkey; E-mail: sdayanik@bilkent.edu.tr

process observations, the Bayes risk is the expected sum the false alarm frequency and detection delay penalty, which is a general smooth function of the detection delay time.

The compound Poisson disorder problems arise in homeland security to detect and analyze the abnormal flow of passengers and commodities at the ports of entries, in computer network security to identify attempts to gain unauthorized control of services from incoming packet flows to various communication ports of Web servers, and in public health to determine the onset of an epidemic in a geographical area from the fluctuations in the emergency room visits to the hospitals.

Several non-Bayesian formulations and solutions of the quickest change-detection problems have been studied by Baron and Tartakovsky (2006), Hadjiliadis (2005), Hadjiliadis and Moustakides (2005), Moustakides (2004, 2008), and Shiryaev (1996) in continuous time, and by Lorden (1971), Moustakides (1986), Pollak (1985), Tartakovsky (2008), and Tartakovsky and Veeravalli (2004) in discrete time. Shiryaev (1963) introduced and solved the Bayesian formulation of quickest detection problem for general distributions in discrete time and for a change in the drift of a Brownian motion in continuous time. Galchuk and Rozovskii (1971) formulated simple Poisson disorder problem and provided partial solution, which has been completed by Peskir and Shiryaev (2002). Gapeev (2005) solved compound Poisson disorder problem with exponentially distributed jumps. The solution for the general case was provided by Dayanik and Sezer (2006). Bayesian sequential detection of a change in the local characteristics of a finite-activity Lévy process has been formulated and solved by Dayanik et al. (2008). Basseville and Nikiforov (1993), Peskir and Shiryaev (2006), and Poor and Hadjiliadis (2008) give a detailed review of the literature on both non-Bayesian and Bayesian sequential change detection problems.

Higher moments of detection delay time were shown by Baron and Tartakovsky (2006) and Tartakovsky and Veeravalli (2004) to be asymptotically minimized by the Shiryaev's procedure in Bayesian setting. The solution of Bayesian sequential change detection problems with exponential detection delay penalties were found by Poor (1998) in discrete time, by Beibel (2000) in detecting a change in the drift of a Brownian motion, by Bayraktar and Dayanik (2006) and Bayraktar et al. (2005) in simple Poisson disorder problem, and by Dayanik and Sezer (2006) in compound Poisson disorder problem. Shiryaev (1964, 2008) and Shiryaev and Zryumov (2009) derived the sufficient statistics for sequential change detection problems with nonlinear detection delay penalty costs, which include as special cases the higher moments and exponential functions of detection delay time. Gapeev (2006) studied nonadditive Bayesian problems of detecting a change in the drift rate of an observed diffusion.

We give the precise description of the compound Poisson disorder problem in Section 2, where we show that for infinitely continuously differentiable detection delay penalty functions, there are countably infinitely many piecewise deterministic strongly Markov sufficient statistic for the problem. Our derivation is different from that of Shiryaev (2008) in that we use a suitable reference probability measure, change of measure, and change-of-variable formula to systematically "complete" minimal sufficient statistic to a Markov sufficient statistic. The detection delay penalty functions, which are the solutions of homogeneous  $m + 1$ st order constant coefficient ordinary differential equations, are shown to lead to

an  $m$ -dimensional sufficient statistic, which is a piecewise deterministic strong Markov process. Therefore, any penalty function that is a linear combinations of products of exponential, polynomial, and sinusoidal functions is a solution of some homogeneous constant-coefficient ordinary differential equation and leads to a finite-dimensional sufficient statistic that is a piecewise deterministic strongly Markov process. In the meantime, the disorder problem can be reduced to an optimal stopping problem, and with a finite-dimensional piecewise deterministic strongly Markov process, one can solve it by using dynamic programming and successive approximations.

In Section 3, we explain the solution methodology in detail by specializing to polynomial disorder detection penalty function with arbitrary but fixed degree. By means of suitable dynamic programming operator, the continuous-time optimal stopping problem is reduced to an essentially discrete-time optimal stopping problem. This approach is based on the stochastic dynamic optimization theory for piecewise deterministic Markov processes; see, for example, Gugerli (1986) and Davis (1993). The dynamic programming operator maps every bounded function to another bounded function, whose value at every point in the domain is obtained as the solution of a straightforward deterministic optimization problem. The repeated applications of the dynamic programming operator to constant zero mapping result in successive approximations of the key optimal stopping problem's value function, which turns out to be unique bounded fixed point of the dynamic programming operator. In the meantime, the solutions of deterministic optimization problems naturally lead to nearly-optimal detection alarm times. We show that optimal alarm time exists and can be characterized as the first hitting time of the Markov sufficient statistic to a closed convex subset, which can be approximated arbitrarily well by the zero sets of the successive approximations of the value function. We also show that successive approximations converge to the value function over state space uniformly and exponentially fast, and explicit error bound allows one to set the accuracy of nearly-optimal alarm times to any desired level.

In Section 4, we illustrate some of the findings on the compound Poisson problem with quadratic detection delay penalty cost function. We described qualitatively, but quite explicitly, the form of optimal stopping time of the auxiliary optimal stopping problem, which is also the optimal alarm time for the compound Poisson disorder problem. Finally, the long proofs of selected results are deferred to the Appendix.

## 2. PROBLEM DESCRIPTION

Let  $(T_n, Z_n)$ ,  $n \geq 1$  be a compound Poisson process whose arrival rate  $\lambda$  and mark distribution  $\nu$  on some measurable space  $(E, \varepsilon)$  changes from  $(\lambda_0, \nu_0)$  to  $(\lambda_1, \nu_1)$  at some unobservable disorder time  $\Theta$ , which has zero-modified exponential distribution

$$\mathbb{P}\{\Theta = 0\} = p \quad \text{and} \quad \mathbb{P}\{\Theta > t\} = (1 - p)e^{-\lambda t}, \quad t \geq 0$$

for some known constants  $\lambda_0 > 0$ ,  $\lambda_1 > 0$ ,  $\lambda > 0$ ,  $0 \leq p < 1$ , and known probability distributions  $\nu_0$  and  $\nu_1$  on  $(E, \varepsilon)$ . We want to detect the disorder time  $\Theta$  by means of a stopping time  $\tau$  of the observation filtration  $\mathcal{F}_t = \sigma\{(T_n, Z_n); n \geq 1$

such that  $T_n \leq t$ ,  $t \geq 0$  so as to minimize the expected total risk of false alarms and detection delay time. For every  $(\mathcal{F}_t)_{t \geq 0}$ -stopping time we define the Bayes risk as

$$R_\tau(p) := \mathbb{E}[1_{\{\tau < \Theta\}} + f(\tau - \Theta)1_{\{\tau \geq \Theta\}}]$$

for some general sufficiently smooth penalty function  $f: \mathbb{R}_+ \mapsto \mathbb{R}$  of detection delay time  $(\tau - \Theta)^+$ . We would like to (i) calculate the smallest Bayes risk

$$\inf_{\tau \in \mathcal{S}} R_\tau(p) \quad \text{for every } 0 \leq p < 1,$$

where the infimum is taken over the collection  $\mathcal{S}$  of all  $(\mathcal{F}_t)_{t \geq 0}$ -stopping times; and (ii) find a stopping time in  $\mathcal{S}$  which attains the infimum, if one such stopping time exists.

It is always possible to construct a probability space  $(\Omega, \mathcal{F}, \mathbb{P}_\infty)$  with a reference probability measure  $\mathbb{P}_\infty$  under which (i)  $(T_n, Z_n)_{n \geq 1}$  is a compound Poisson process with arrival rate  $\lambda_0$  and mark distribution  $v_0$  on  $(E, \varepsilon)$ , and  $\Theta$  is an independent random variable with zero-modified exponential distribution. Suppose that  $\lambda_1$  is a positive constant and  $v_1$  is a probability distribution on  $(E, \varepsilon)$  absolutely continuous with respect to  $v_0$ , and either  $\lambda_0 \neq \lambda_1$  or  $v_0 \neq v_1$ . Let  $\mathcal{G}_t := \sigma(\Theta) \vee \mathcal{F}_t$ ,  $t \geq 0$  be the filtration obtained by augmenting  $(\mathcal{F}_t)_{t \geq 0}$  with the information about  $\Theta$  and define the probability measure  $\mathbb{P}$  locally on  $(\Omega, \mathcal{G}_\infty)$  through the Radon–Nikodym derivatives

$$\left. \frac{d\mathbb{P}}{d\mathbb{P}_\infty} \right|_{\mathcal{G}_t} = Z_t := 1_{\{t < \Theta\}} + 1_{\{t \geq \Theta\}} \frac{L_t}{L_\Theta}, \quad t \geq 0,$$

where

$$\begin{aligned} L_t &:= e^{-(\lambda_1 - \lambda_0)t} \prod_{n: T_n \leq t} \left( \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(Z_n) \right) \\ &= \exp \left\{ -(\lambda_1 - \lambda_0)t + \int_0^t \int_E \left( \log \frac{\lambda_1}{\lambda_0} \frac{v_1}{v_0}(z) \right) N(ds, dz) \right\}, \end{aligned}$$

and  $N(ds, dz)$  is the Poisson random measure on  $(\mathbb{R}_+ \times E, \mathcal{B}(\mathbb{R}_+) \times \varepsilon)$  with mean measure  $\lambda_0 ds v_0(dz)$  under  $\mathbb{P}_\infty$ . Girsanov’s change-of-measure theorem guarantees that  $(T_n, Z_n)_{n \geq 1}$  and  $\Theta$  have jointly the same statistical law under  $\mathbb{P}$  as they are described in the introduction. Therefore, we will work in the remainder with  $\mathbb{P}$  obtained by a change-of-measure from the reference probability measure  $\mathbb{P}_\infty$  on  $(\Omega, \mathcal{G}_\infty)$ . The change-of-variable formula gives the dynamics of process  $L = \{L_t, \mathcal{F}_t; t \geq 0\}$  as

$$L_0 = 1 \quad \text{and} \quad dL_t = L_{t-} \int_E \left( \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(z) - 1 \right) [N(dt, dz) - \lambda_0 dt v_0(dz)], \quad t \geq 0. \quad (2.1)$$

For every stopping time  $\tau \in \mathcal{S}$ , we have

$$\begin{aligned} &\mathbb{E}[f(\tau - \Theta)1_{\{\tau \geq \Theta\}}] - f(0)\mathbb{P}\{\tau \geq \Theta\} \\ &= \mathbb{E}[(f(\tau - \Theta) - f(0))1_{\{\tau \geq \Theta\}}] = \mathbb{E} \left[ 1_{\{\tau \geq \Theta\}} \int_0^{\tau - \Theta} f'(t) dt \right] \\ &= \mathbb{E} \left[ 1_{\{\tau \geq \Theta\}} \int_\Theta^\tau f'(t - \Theta) dt \right] = \mathbb{E} \left[ \int_0^\infty f'(t - \Theta) 1_{\{\Theta \leq t\}} 1_{\{\tau > t\}} dt \right]. \end{aligned}$$

Because  $\tau \wedge \Theta$  is  $\mathbb{P}_\infty$ -a.s. finite stopping time of  $(\mathcal{G}_t)_{t \geq 0}$  and  $Z_{\Theta \wedge \tau} = Z_\Theta = 1$  on  $\{\Theta \leq \tau\}$ , we have

$$\mathbb{P}\{\tau \geq \Theta\} = \mathbb{E}_\infty[Z_{\tau \wedge \Theta} 1_{\{\tau \geq \Theta\}}] = \mathbb{P}_\infty\{\tau \geq \Theta\} = p + (1-p)\mathbb{E}_\infty\left[\int_0^\tau \lambda e^{-\lambda t} dt\right].$$

Since  $Z_t 1_{\{\Theta > t\}} = 1_{\{\Theta > t\}}$  for every  $t \geq 0$ , the independence of  $\Theta$  and  $\mathcal{F}_t$  under  $\mathbb{P}_\infty$  implies that we have  $\mathbb{E}_\infty[Z_t 1_{\{\Theta > t\}} | \mathcal{F}_t] = \mathbb{P}_\infty\{\Theta > t\} = (1-p)e^{-\lambda t}$  for every  $t \geq 0$ , and

$$\begin{aligned} & \mathbb{E}\left[\int_0^\infty f'(t - \Theta) 1_{\{\Theta \leq t\}} 1_{\{\tau > t\}} dt\right] \\ &= \mathbb{E}_\infty\left[\int_0^\infty \mathbb{E}_\infty[Z_t f'(t - \Theta) 1_{\{\Theta \leq t\}} | \mathcal{F}_t] 1_{\{\tau > t\}} dt\right] \\ &= \mathbb{E}_\infty\left(\int_0^\tau \mathbb{E}_\infty[Z_t 1_{\{\Theta > t\}} | \mathcal{F}_t] \frac{\mathbb{E}_\infty[Z_t f'(t - \Theta) 1_{\{\Theta \leq t\}} | \mathcal{F}_t]}{\mathbb{E}_\infty[Z_t 1_{\{\Theta > t\}} | \mathcal{F}_t]} dt\right) \\ &= (1-p)\mathbb{E}_\infty\left[\int_0^\tau e^{-\lambda t} \Phi_t^{(1)} dt\right] \end{aligned}$$

in terms of the first element of the sequence of processes

$$\Phi_t^{(n)} := \frac{\mathbb{E}_\infty[Z_t f^{(n)}(t - \Theta) 1_{\{\Theta \leq t\}} | \mathcal{F}_t]}{\mathbb{E}_\infty[Z_t 1_{\{\Theta > t\}} | \mathcal{F}_t]} = \frac{\mathbb{E}[f^{(n)}(t - \Theta) 1_{\{\Theta \leq t\}} | \mathcal{F}_t]}{\mathbb{P}\{\Theta > t | \mathcal{F}_t\}}, \quad t \geq 0, \quad n \geq 1,$$

where we denote by  $f^{(n)}$  the  $n$ th derivative of  $f(\cdot)$ , and the last equality follows from Bayes formula. Therefore, the Bayes risk of every stopping time  $\tau \in S$  can be written as

$$\begin{aligned} R_\tau(p) &= \mathbb{P}\{\tau < \Theta\} + f(0)\mathbb{P}\{\tau \geq \Theta\} + \mathbb{E}\left[\int_0^\infty f'(t - \Theta) 1_{\{\Theta \leq t\}} 1_{\{\tau > t\}} dt\right] \\ &= 1 - p + pf(0) + (1-p)\mathbb{E}_\infty\left[\int_0^\tau e^{-\lambda t} (\Phi_t^{(1)} + \lambda f(0) - \lambda) dt\right]. \end{aligned}$$

**Proposition 2.1.** *If the detection delay penalty function  $f(\cdot)$  is continuously differentiable, then the Bayesian sequential quickest detection problem is equivalent to solving*

$$\inf_{\tau \in \mathcal{S}} R_\tau(p) = 1 - p + pf(0) + (1-p) \inf_{\tau \in \mathcal{S}} \mathbb{E}_\infty\left[\int_0^\tau e^{-\lambda t} (\Phi_t^{(1)} + \lambda f(0) - \lambda) dt\right], \quad 0 \leq p < 1.$$

*If the optimal stopping problem on the righthand side admits an optima  $(\mathcal{F}_t)_{t \geq 0}$ -stopping time, then it is also a Bayes-optimal change-detection alarm time for the compound Poisson disorder problem.*

In the remainder, we will develop and use methods to solve the optimal stopping problem of Proposition 2.1 and identify optimal and nearly-optimal stopping times. The process  $\Phi^{(1)} = \{\Phi_t^{(1)}, \mathcal{F}_t; t \geq 0\}$ , is a sufficient statistic for the quickest detection problem in the sense that  $\{\Phi_s^{(1)}; 0 \leq s \leq t\}$  summarizes all of the information

contained in the observations  $\mathcal{F}_t$  up to and including time  $t$  for a decision to be made at time  $t$  between raising an alarm and waiting for at least some more infinitesimal amount of time. However,  $\Phi^{(1)}$  is in general not a Markov process under  $\mathbb{P}_\infty$ .

In general, if  $f(\cdot)$  is  $(m + 1)$ -times continuously differentiable for some  $m \geq 1$ , then the processes  $\Phi^{(n)} = \{\Phi_t^{(n)}, \mathcal{F}_t; t \geq 0\}$ ,  $1 \leq n \leq m$  follow the dynamics

$$\begin{aligned} \Phi_0^{(n)} &= \frac{p}{1-p} f^{(n)}(0), \\ d\Phi_t^{(n)} &= [\lambda(f^{(n)}(0) + \Phi_t^{(n)}) + \Phi_t^{(n+1)}]dt + \Phi_t^{(n)} \int_E \left( \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(z) - 1 \right) \\ &\quad \times [N(dt, dz) - \lambda_0 dt v_0(dz)], \quad t \geq 0, \quad 1 \leq n \leq m, \end{aligned} \tag{2.2}$$

the derivation of which is a straight-forward application of Itô rule. For every  $1 \leq n \leq m$ , the drift of  $\Phi^{(n)}$  depends on  $\Phi^{(n+1)}$ , and the process  $\{(\Phi_t^{(1)}, \dots, \Phi_t^{(m)}), \mathcal{F}_t; t \geq 0\}$  is in general not a Markov process. If  $f(\cdot)$  is infinitely continuously differentiable, then under suitable conditions  $\{(\Phi_t^{(1)}, \Phi_t^{(2)}, \dots), \mathcal{F}_t; t \geq 0\}$ , will be an infinite-dimensional Markov process. The finite system of stochastic differential equations in (2.2) is ‘‘closable,’’ for example, if  $\Phi^{(m+1)}$  can be expressed in terms of  $\Phi^{(1)}, \dots, \Phi^{(m)}$ , in which case the  $m$ -dimensional process  $\{(\Phi_t^{(1)}, \dots, \Phi_t^{(m)}), \mathcal{F}_t; t \geq 0\}$  is a Markov sufficient statistic for the sequential change detection problem.

In attacking general sequential statistical problems for processes with stationary and independent increments, the importance and the explicit form of Markov sufficient statistics were noticed first in discrete time by Shiryaev (1964) and then in continuous time by Shiryaev (2008, Theorem 1 and Remark 4), who recently derived them for the generalized Bayesian nonlinear quickest detection of an abrupt change in the drift of a Brownian motion and in the arrival rate of a simple Poisson process.

**Example 2.1.** In each of the following examples, the system in (2.2) is closable, and the  $m$ -dimensional process  $\{(\Phi_t^{(1)}, \dots, \Phi_t^{(m)}), \mathcal{F}_t; t \geq 0\}$  is a piecewise deterministic strong Markov process.

(1) Suppose that  $f(t) = a_0 + a_1(t - b_1) + a_2(t - b_2)^2 + \dots + a_m(t - b_m)^m$  for every  $t \geq 0$  for some constants  $a_0, b_0, \dots, a_m, b_m$ . Then  $f^{(m+1)}(\cdot) \equiv 0$  and  $\mathbb{P}_\infty$ -a.s.  $\Phi_t^{(m+1)} = 0$  for every  $t \geq 0$ . The simple Poisson disorder problem (i.e.,  $\lambda_0 \neq \lambda_1$  and  $v_1 \equiv v_0$ ) with linear detection delay penalty function  $f(t) = t$  was formulated and partially solved by Galchuk and Rozovskii (1971). The complete solution was later given by Peskir and Shiryaev (2002) by using method of variational inequalities. Later, Dayanik and Sezer (2006) described the solution of compound Poisson disorder problem with linear detection delay penalty by first reducing the original problem to a discrete-time optimal stopping problem, which is then solved with successive approximations.

(2) Suppose that  $f(t) = ae^{bt} + c$ ,  $t \geq 0$  for some constants  $a, b \neq 0$ , and  $c$ . Then  $f^{(1)}(t) = abe^{bt}$  and  $f^{(2)}(t) = ab^2e^{bt} = bf^{(1)}(t)$ . Therefore,  $\Phi_t^{(2)} = b\Phi_t^{(1)}$  for every  $t \geq 0$ , and  $m = 1$  because

$$\Phi_0^{(1)} = \frac{abp}{1-p},$$

$$d\Phi_t^{(1)} = [\lambda(\Phi_t^{(1)} + ab) + b\Phi_t^{(1)}]dt + \Phi_t^{(1)} \int_E \left( \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(z) - 1 \right) \times [N(dt, dz) - \lambda_0 dt dz], \quad t \geq 0$$

is autonomous, and the sufficient statistic  $\{\Phi_t^{(1)}, \mathcal{F}_t; t \geq 0\}$  for the sequential change detection problem is a one-dimensional piecewise deterministic strong Markov process. The simple Poisson disorder problem with exponential detection delay penalty and  $a = -c = 1$  was solved by Bayraktar et al. (2005) ? by the method of variational inequalities. The compound Poisson disorder problem with the same exponential detection delay penalty function was later solved by Dayanik and Sezer (2006) with successive approximations applied to an equivalent essentially discrete-time optimal stopping problem.

(3) Suppose that the detection delay penalty function  $f$  is  $(m + 1)$ -times continuously differentiable, and that  $f^{(1)}$  solves  $m$ th order constant coefficient homogeneous ordinary differential equation

$$0 = c_1 f^{(1)}(t) + c_2 f^{(2)}(t) + \dots + c_m f^{(m)}(t) + f^{(m+1)}(t) \quad \text{for every } t \geq 0.$$

Then we have  $\mathbb{P}_\infty$ -a.s.  $\Phi_t^{(m+1)} = -\sum_{n=1}^m c_n \Phi_t^{(n)}$  for all  $t \geq 0$ , and the system of  $m$  stochastic differential equations in (2.2) is autonomous. Hence, the  $m$ -dimensional process  $\{(\Phi_t^{(1)}, \dots, \Phi_t^{(m)}), \mathcal{F}_t; t \geq 0\}$  is a strong Markov sufficient statistic for the sequential change detection problem. For suitable constants  $\rho_n, a_n, \alpha_n, b_n, \beta_n, \gamma_n$  for  $1 \leq n \leq m$ , the general solution of the  $m$ th order constant coefficient homogeneous ordinary differential equation is

$$f(t) = \sum_{n=1}^m (a_n \cos \alpha_n t + b_n \sin \beta_n t) t^{\gamma_n} e^{\rho_n t}, \quad t \geq 0.$$

In the remainder, we will specialize to the detection delay penalty function  $f(t) = t^m, t \geq 0$  for an arbitrary but fixed  $m \geq 1$  and describe in detail the solution of compound Poisson disorder problem. The method easily extends other cases with finite-dimensional Markov sufficient statistics. For every  $a > 0$ ,

$$\lim_{m \rightarrow \infty} \left( \frac{t}{a} \right)^m = \begin{cases} 0, & \text{if } 0 \leq t < a, \\ 1, & \text{if } t = a, \\ \infty, & \text{if } t > a, \end{cases}$$

and for large  $m, t \mapsto (t/a)^m$  is a reasonable penalty function for the sequential change detection problems, where detection delay less than  $a$  is tolerable, but detection delay more than  $a$  is completely unacceptable. For convenience, we take  $a = 1$ . Proposition 2.2 now follows from Proposition 2.1 and (2.2).

**Proposition 2.2.** *Suppose that the detection delay penalty function is  $f(t) = t^m$  for every  $t \geq 0$  for some  $m \geq 1$ . Then the minimum Bayes risk equals*

$$\inf_{\tau \in \mathcal{S}} R_\tau(p) = 1 - p + (1 - p)V\left(0, \dots, 0, \frac{pm!}{1 - p}\right), \quad 0 \leq p < 1$$



in terms of the value function of the discounted optimal stopping problem

$$V(\phi) = \inf_{\tau \in \mathcal{F}} \mathbb{E}_\infty^\phi \left[ \int_0^\tau e^{-\lambda t} g(\Phi_t) dt \right], \quad \phi \in \mathbb{R}_+^m = \overbrace{\mathbb{R}_+ \times \cdots \times \mathbb{R}_+}^{m\text{-times}} \quad (2.3)$$

with running cost function  $g: \mathbb{R}_+^m \mapsto \mathbb{R}$  defined by  $g(\phi) = e_1^T \phi - \lambda \equiv \phi_1 - \lambda$  for the  $m$ -dimensional piecewise deterministic Markov process  $\Phi = \{\Phi_t = (\Phi_t^{(1)}, \dots, \Phi_t^{(m)}), \mathcal{F}_t; t \geq 0\}$ , whose dynamics are

$$d\Phi_t^{(n)} = [\lambda \Phi_t^{(n)} + \Phi_t^{(n+1)}] dt + \Phi_{t-}^{(n)} \int_E \left( \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(z) - 1 \right) \times [N(dt, dz) - \lambda_0 dt v_0(dz)], \quad t \geq 0,$$

$$\Phi_0^{(n)} = \frac{p}{1-p} f^{(n)}(0) = 0 \quad \text{for every } n = 1, \dots, m-1, \text{ and}$$

$$d\Phi_t^{(m)} = \lambda(\Phi_t^{(m)} + m!) dt + \Phi_{t-}^{(m)} \int_E \left( \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(z) - 1 \right) [N(dt, dz) - \lambda_0 dt v_0(dz)], \quad t \geq 0,$$

$$\Phi_0^{(m)} = \frac{p}{1-p} f^{(m)}(0) = \frac{pm!}{1-p},$$

where for every  $\phi \in \mathbb{R}_+^m$  expectation  $\mathbb{E}_\infty^\phi$  is taken under  $\mathbb{P}_\infty$  such that  $\mathbb{P}_\infty\{\Phi_0 = \phi\} = 1$ .

The jumps and deterministic evolution between jumps of process  $\Phi$  can be separated, and its dynamics can be written compactly as

$$\Phi_0 = \frac{pm!}{1-p} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}_{m \times 1}, \quad d\Phi_t = (A\Phi_t + b)dt + \Phi_{t-} \int_E \left( \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(z) - 1 \right) N(dt, dz), \quad t \geq 0,$$

$$\Phi_t = \begin{bmatrix} \Phi_t^{(1)} \\ \vdots \\ \Phi_t^{(m)} \end{bmatrix}, \quad A = - \begin{bmatrix} \bar{\lambda} & 1 & 0 & 0 & \cdots & 0 \\ 0 & \bar{\lambda} & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \cdots & \vdots \\ 0 & \cdots & \bar{\lambda} & 1 & 0 & \\ 0 & \cdots & 0 & \bar{\lambda} & 1 & \\ 0 & \cdots & 0 & 0 & \bar{\lambda} & \end{bmatrix}_{m \times m},$$

$$b = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \lambda m! \end{bmatrix}_{m \times 1}, \quad \bar{\lambda} = \lambda_1 - \lambda_0 - \lambda.$$

**Proposition 2.3.** *The process  $\Phi = \{\Phi_t = (\Phi_t^{(1)}, \dots, \Phi_t^{(m)}), \mathcal{F}_t, t \geq 0\}$  is an  $m$ -dimensional piecewise deterministic strong Markov process under  $\mathbb{P}_\infty$ , and  $\mathbb{P}_\infty$ -a.s. for every  $t \geq 0$*

$$\Phi_t = \begin{cases} \varphi(t - T_n, \Phi_{T_n}), & \text{if } t \in [T_n, T_{n+1}) \text{ for some } n \geq 0, \\ \varphi(T_{n+1} - T_n, \Phi_{T_n}), \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(Z_{n+1}) & \text{if } t = T_{n+1} \text{ for some } n \geq 0, \end{cases}$$

where  $T_0 \equiv 0$ , and mapping  $(t, \phi) \mapsto \varphi(t, \phi) = (\varphi_1(t, \phi), \dots, \varphi_m(t, \phi)) : \mathbb{R}_+ \times \mathbb{R}_+^m \mapsto \mathbb{R}_+^m$  defined by

$$\varphi(t, \phi) = \begin{cases} e^{At}(\phi + A^{-1}b) - A^{-1}b, & \text{if } \bar{\lambda} \neq 0 \\ e^{At}\phi + \left(\int_0^t e^{As} ds\right)b, & \text{if } \bar{\lambda} = 0 \end{cases} \text{ for every } t \geq 0 \text{ and } \phi \in \mathbb{R}_+^m,$$

is the solution of the system of  $m$  linear ordinary differential equations  $dx(t) = Ax(t) + b, t \geq 0$  with initial condition  $x(0) = \phi$ . If  $\bar{\lambda} > 0$ , then  $\lim_{t \rightarrow \infty} \varphi(t, \phi) = -A^{-1}b$  for every  $\phi \in \mathbb{R}_+^m$ . If  $\bar{\lambda} \leq 0$ , then  $t \mapsto \varphi_n(t, \phi)$  is strictly increasing and  $\lim_{t \rightarrow \infty} \varphi_n(t, \phi) = +\infty$  for every  $1 \leq n \leq m$  and  $\phi \in \mathbb{R}_+^m$ .

### 3. A DYNAMIC PROGRAMMING OPERATOR AND THE SOLUTION

Let us define for every bounded function  $w : \mathbb{R}_+^m \mapsto \mathbb{R}$

$$(Kw)(\phi) = \int_E w\left(\phi \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(z)\right) v_0(dz), \quad \phi \in \mathbb{R}_+^m, \tag{3.1}$$

$$(Jw)(\phi, r) = \int_0^r e^{-(\lambda+\lambda_0)t} [g + \lambda_0(Kw)](\varphi(t, \phi)) dt, \quad r \geq 0, \quad \phi \in \mathbb{R}_+^m, \tag{3.2}$$

$$(J_t w)(\phi) = \inf_{r \geq t} (Jw)(\phi, r), \quad t \geq 0, \quad \phi \in \mathbb{R}_+^m. \tag{3.3}$$

Operators  $J$  and  $J_t$  naturally appear in the optimality equation satisfied by the value function of the optimal stopping problem in (2.3). This important connection is the result of the special characterization of stopping times  $\tau \in \mathcal{S}$  as described below by Proposition 3.1, the proof of which essentially follows from the arguments of Liptser and Shiryaev (2001, pp. 270–271).

**Proposition 3.1.** *For all  $\tau \in \mathcal{S}$  and  $n \geq 0$ , there is an  $\mathcal{F}_{T_n}$ -measurable nonnegative r.v.  $R_n$  such that  $\mathbb{P}_\infty$ -a.s.*

$$1_{\{\tau \geq T_n\}}[\tau \wedge T_{n+1}] = 1_{\{\tau \geq T_n\}}[(T_n + R_n) \wedge T_{n+1}], \tag{3.4}$$

$$\{\tau \geq T_n\} = \{R_0 \geq T_1, T_1 + R_1 \geq T_2, \dots, T_{n-1} + R_{n-1} \geq T_n\}, \tag{3.5}$$

$$\{T_n \leq \tau < T_{n+1}\} = \{R_0 \geq T_1, T_1 + R_1 \geq T_2, \dots, T_{n-1} + R_{n-1} \geq T_n, T_n + R_n < T_{n+1}\}. \tag{3.6}$$

Toward a solution of the optimal stopping problem in (2.3) with detection delay penalty function  $f(t) = t^m, t \geq 0$  for arbitrary but fixed  $m \geq 1$ , let us consider the following policy: suppose that we agreed to stop at some fixed stopping time  $\tau \in \mathcal{S}$  if  $\tau < T_1$ , namely, if no mark (and therefore no new information) has arrived before the alarm time set by the stopping rule, and otherwise take optimal action at time  $T_1$  based on the value  $\Phi_{T_1}$  of sufficient statistic, which will then incorporate new information contained in the mark just arrived at time  $T_1$ . The strong Markov property of process  $\Phi$  at  $(\mathcal{F}_t)_{t \geq 0}$ -stopping time  $\tau \wedge T_1$  suggests that the expected value of this policy should equal

$$\mathbb{E}_\infty \left[ \int_0^{\tau \wedge T_1} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_1\}} e^{-\lambda T_1} V(\Phi_{T_1}) \right].$$

Let  $R_0 \equiv R_0(\Phi_0)$  be  $\mathcal{F}_0$ -measurable random variable such that  $\mathbb{P}_\infty$ -a.s.  $\tau \wedge T_1 = R_0 \wedge T_1$  and  $\{\tau \geq T_1\} = \{R_0 \geq T_1\}$  as in the characterization of  $\tau$  by Proposition 3.1. Because by Proposition 2.3  $\Phi_t = \varphi(t, \Phi_0)$  for  $t \in [0, T_1)$  and  $\Phi_{T_1} = \varphi(T_1, \Phi_0) \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(Z_1)$  and since  $(T_n, Z_n)_{n \geq 1}$  is a compound Poisson process with arrival rate  $\lambda_0$  and mark distribution  $v_0$  on  $(E, \varepsilon)$ , and since  $\mathcal{F}_0$  and  $(T_1, Z_1)$  are independent due to independent increments of  $(T_n, Z_n)_{n \geq 1}$  under  $\mathbb{P}_\infty$ , we can rewrite the expected value of the policy as

$$\begin{aligned} \mathbb{E}_\infty & \left[ \int_0^{R_0} 1_{\{T_1 \geq t\}} e^{-\lambda t} g(\varphi(t, \Phi_0)) dt + 1_{\{R_0 \geq T_1\}} e^{-\lambda T_1} V \left( \varphi(T_1, \Phi_0) \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(Z_1) \right) \right] \\ & = \int_0^{R_0} e^{-(\lambda + \lambda_0)t} g(\varphi(t, \Phi_0)) dt + \int_0^{R_0} \lambda_0 e^{-(\lambda + \lambda_0)t} \int_E V \left( \varphi(t, \Phi_0) \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(z) \right) v_0(dz) dt \\ & = \int_0^{R_0} e^{-(\lambda + \lambda_0)t} [g + \lambda_0(KV)](\varphi(t, \Phi_0)) dt \equiv (JV)(\Phi_0, R_0). \end{aligned}$$

Therefore, the minimum expected total discounted cost should be given by

$$\inf_{\tau \in \mathcal{S}} \mathbb{E}_\infty \left[ \int_0^{\tau \wedge T_1} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_1\}} e^{-\lambda T_1} V(\Phi_{T_1}) \right] = \inf_{r \geq 0} (JV)(\Phi_0, r) \equiv (J_0 V)(\Phi_0).$$

Because  $V(\Phi_0)$  is by definition the minimum expected total discounted cost, the optimality principle of dynamic programming suggests that  $V(\Phi_0) = (J_0 V)(\Phi_0)$  and that  $J_0$  can be seen as a dynamic programming operator. We later show that  $V(\cdot)$  is indeed a solution of the optimality equation. In fact,  $V(\cdot)$  is the unique bounded fixed point of operator  $J_0$  and can be approximated successively by the elements of the sequence

$$v_0(\phi) = 0, \quad \phi \in \mathbb{R}_+^m \quad \text{and} \quad v_n(\phi) = (J_0 v_{n-1})(\phi), \quad \phi \in \mathbb{R}_+^m, \quad n \geq 1. \quad (3.7)$$

Let us first introduce the finite-horizon problems

$$V_n(\phi) = \inf_{\tau \in \mathcal{S}} \mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_n} e^{-\lambda t} g(\Phi_t) dt \right] \quad \text{for every } \phi \in \mathbb{R}_+^m \text{ and } n \geq 0. \quad (3.8)$$

obtained from the original problem in (2.3) by requiring a decision at or before the arrival time  $T_n$  of the  $n$ th mark. The next lemma shows that  $V(\phi)$  can be approximated successively by the elements of sequence  $(V_n(\phi))_{n \geq 0}$  as  $n \rightarrow \infty$ , uniformly in  $\phi \in \mathbb{R}_+^m$ .

**Lemma 3.1.** *The sequence  $(V_n(\phi))_{n \geq 0}$  decreases to  $V(\phi)$  as  $n \rightarrow \infty$  uniformly in  $\phi \in \mathbb{R}_+^m$ . More precisely,*

$$0 \leq V_n(\phi) - V(\phi) \leq \left( \frac{\lambda}{\lambda + \lambda_0} \right)^n \quad \text{for every } \phi \in \mathbb{R}_+^m \text{ and } n \geq 1.$$

*Proof.* The first inequality is obvious. On the other hand, because  $g(\phi) \geq -\lambda$  for every stopping time  $\tau \in \mathcal{S}$ , and under  $\mathbb{P}_\infty$  the random variable  $T_n$  has Erlang distribution with parameters  $n$  and  $\lambda_0$ , we have

$$\mathbb{E}_\infty^\phi \left[ \int_0^\tau e^{-\lambda t} g(\Phi_t) dt \right] = \mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_n} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_n\}} \int_{T_n}^\tau e^{-\lambda t} g(\Phi_t) dt \right]$$

$$\begin{aligned} &\geq \mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_n} e^{-\lambda t} g(\Phi_t) dt \right] - \mathbb{E}_\infty^\phi \left[ \int_{T_n}^\infty \lambda e^{-\lambda t} dt \right] \\ &\geq V_n(\phi) - \mathbb{E}_\infty^\phi [e^{-\lambda T_n}] = V_n(\phi) - \left( \frac{\lambda}{\lambda + \lambda_0} \right)^n. \quad \square \end{aligned}$$

Propositions 3.2 and 3.3 show that  $V_n(\cdot) = v_n(\cdot)$  for every  $n \geq 0$ . Namely,  $(V_n(\cdot))_{n \geq 0}$  can be calculated iteratively by successive applications of the dynamic programming operator  $J_0$  to function  $v_0 \equiv 0$ . Since each  $v_n(\cdot)$  is obtained as the solution of a straightforward deterministic optimization problem, Lemma 3.1 and Propositions 3.2 and 3.3 suggest for the problem in (2.3) an effective numerical solution method, which also turns out to be very useful to identifying the structural properties of the solution.

**Proposition 3.2.** *For every  $n \geq 0$  and  $\phi \in \mathbb{R}_+^m$ , we have  $V_n(\phi) > v_n(\phi)$ .*

**Proposition 3.3.** *For every  $\varepsilon > 0$ ,  $\phi \in \mathbb{R}_+^m$ , and  $n \geq 1$ , let  $r_{n,\varepsilon}(\phi)$  be a nonnegative number such that  $(Jv_{n-1})(\phi, r_{n,\varepsilon}(\phi)) \leq \varepsilon + (J_0 v_{n-1})(\phi) \equiv \varepsilon + v_n(\phi)$ , and define*

$$\tau_{0,\varepsilon} \equiv 0 \quad \text{and} \quad \tau_{n,\varepsilon} = \begin{cases} r_{n,\varepsilon/2}(\Phi_0), & \text{if } r_{n,\varepsilon/2}(\Phi_0) < T_1, \\ T_1 + T_{n-1,\varepsilon/2} \circ \theta_{T_1}, & \text{if } r_{n,\varepsilon/2} \circ \phi_{T_1} \geq T_1. \end{cases}$$

*Then for every  $\varepsilon > 0$  and  $n \geq 1$ , we have  $\tau_{n,\varepsilon} \in \mathcal{S}$ , and*

$$\mathbb{E}_\infty^\phi \left[ \int_0^{\tau_{n,\varepsilon} \wedge T_n} e^{-\lambda t} g(\Phi_t) dt \right] \leq v_n(\phi) + \varepsilon.$$

**Corollary 3.1.** *For every  $n \geq 0$  and  $\phi \in \mathbb{R}_+^m$ , we have  $V_n(\phi) = v_n(\phi)$ , and stopping time  $\tau_{n,\varepsilon}$  is  $\varepsilon$ -optimal for the problem in (3.8):  $\mathbb{E}_\infty^\phi \left[ \int_0^{\tau_{n,\varepsilon} \wedge T_n} e^{-\lambda t} g(\Phi_t) dt \right] \leq V_n(\phi) + \varepsilon$  for every  $\phi \in \mathbb{R}_+^m$ ,  $\varepsilon > 0$ , and  $n \geq 1$ .*

*Proof.* The last displayed equation of Proposition 3.3 implies  $V_n(\cdot) \leq v_n(\cdot) + \varepsilon$  for every  $\varepsilon > 0$ , and since  $\varepsilon > 0$  is arbitrary, we conclude that  $V_n(\cdot) \leq v(\cdot)$ . Since the opposite inequality is also true by Proposition 3.2, the equality  $V_n(\phi) \equiv v_n(\phi)$  holds for every  $\phi \in \mathbb{R}_+^m$ . Replacing  $v_n(\phi)$  with  $V_n(\phi)$  in the last displayed equation of Proposition 3.3 now shows that  $\tau_{n,\varepsilon}$  is  $\varepsilon$ -optimal for the problem in (3.8).  $\square$

Lemma 3.2 identifies important properties of the dynamic programming operator  $J_0$ . Particularly,  $J_0$  preserves boundedness, concavity, and monotonicity. It may have at most one fixed point in the space of bounded functions defined on  $\mathbb{R}_+^m$ . Corollary 3.2 below shows that  $J_0$  has one and only one bounded fixed point, which is the value function  $V(\cdot)$  of problem in (2.3).

**Lemma 3.2.** *If  $w : \mathbb{R}_+^m \mapsto \mathbb{R}$  is concave and bounded between  $-1$  and  $0$ , then so is  $(J_0 w)$ . If  $w_1(\cdot) \leq w_2(\cdot)$ , then  $(J_0 w_1)(\cdot) \leq (J_0 w_2)(\cdot)$ . Moreover,  $J_0$  is a contraction mapping on the collection of bounded functions defined on  $\mathbb{R}_+^m$ , and for every bounded  $w_1(\cdot)$  and  $w_2(\cdot)$ , we have  $\|J_0 w_1 - J_0 w_2\| \leq \frac{\lambda_0}{\lambda + \lambda_0} \|w_1 - w_2\|$ .*

**Corollary 3.2.** *The functions  $V_n(\cdot)$ ,  $n \geq 0$  and  $V(\cdot)$  are bounded between  $-1$  and  $0$ , concave, and continuous on  $\mathbb{R}_+^m$ . Moreover,  $V(\cdot)$  is the unique bounded fixed point of operator  $J_0$ .*

*Proof.* Because  $V_0 \equiv 0$  is bounded between  $-1$  and  $0$  and concave on  $\mathbb{R}_+^m$ , an induction on  $n \geq 1$ , Corollary 3.1, and Lemma 3.2 show that  $V_n = J_0 V_{n-1}$ ,  $n \geq 1$  are bounded between  $-1$  and  $0$  and concave. Since they are concave on  $\mathbb{R}_+^m$ , they are also continuous on  $\mathbb{R}_{++}^m$ . Because  $V(\cdot)$  is the uniform pointwise limit of  $(V_n(\cdot))_{n \geq 0}$  by Lemma 3.1,  $V(\cdot)$  is bounded between  $-1$  and  $0$ , concave on  $\mathbb{R}_+^m$ , and continuous on  $\mathbb{R}_{++}^m$ . Finally, Lemmas 3.2 and 3.1 imply that

$$\|J_0 V_n - J_0 V\| \leq \frac{\lambda_0}{\lambda + \lambda_0} \|V_n - V\| \leq \frac{\lambda_0}{\lambda + \lambda_0} \left(\frac{\lambda}{\lambda + \lambda_0}\right)^n \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Hence,  $(J_0 V)(\phi) = \lim_{n \rightarrow \infty} (J_0 V_n)(\phi) = \lim_{n \rightarrow \infty} V_{n+1}(\phi) = V(\phi)$  for every  $\phi \in \mathbb{R}_+^m$  by Corollary 3.1. If  $\tilde{V}(\cdot)$  is another bounded fixed point of  $J_0$ , then  $\|V - \tilde{V}\| = \|J_0 V - J_0 \tilde{V}\| \leq [\lambda_0/(\lambda + \lambda_0)] \|V - \tilde{V}\|$  implies that  $\|V - \tilde{V}\| = 0$ ; i.e.,  $V(\cdot)$  is the unique bounded fixed point of operator  $J_0$ .  $\square$

The next major result is Theorem 3.1, which states that for every  $\varepsilon \geq 0$  the  $(\mathcal{F}_t)_{t \geq 0}$ -stopping time  $\tau_\varepsilon := \inf\{t \geq 0; V(\Phi_t) \geq -\varepsilon\}$  is  $\varepsilon$ -optimal for the problem in (2.3). For its proof, we will need the next few lemmas and their corollaries.

**Lemma 3.3.** *For every bounded  $w : \mathbb{R}_+^m \mapsto \mathbb{R}$ , we have  $(Jw)(\phi, s) + e^{-(\lambda + \lambda_0)s} (J_0 w)(\varphi(s, \phi)) = (J_s w)(\phi)$  for every  $\phi \in \mathbb{R}_+^m$  and  $s \geq 0$ . If  $(J_0 w)(\varphi(s, \phi)) < 0$  for every  $0 \leq s < t$ , then  $(J_s w)(\phi) = (J_t w)(\phi)$  for every  $0 \leq s \leq t$ .*

The second part of Corollary 3.3 implies that, as long as the value function  $V(\cdot)$  of the optimal stopping problem in (2.3) remains strictly negative along the path  $t \mapsto \varphi(t, \phi)$ , postponing the stopping decision does not cause any regrets. This is the crucial result needed for the proof of the optimality of the stopping time  $\tau_0 := \inf\{t \geq 0; V(\Phi_t) = 0\}$ .

**Corollary 3.3.** *If we take  $w = V$  in Lemma 3.3, then we have  $(JV)(\phi, s) + e^{-(\lambda + \lambda_0)s} V(\varphi(s, \phi)) = (J_s V)(\phi)$  for every  $s \geq 0$  and  $\phi \in \mathbb{R}_+^m$ , because  $V = J_0 V$ . If  $V(\varphi(s, \phi)) < 0$  for every  $0 \leq s < t$ , then  $V(\phi) = (J_s V)(\phi)$  and  $(JV)(\phi, s) + e^{-(\lambda + \lambda_0)s} V(\varphi(s, \phi)) = V(\phi)$  for every  $0 \leq s \leq t$ .*

**Lemma 3.4.** *Let us define  $(\mathcal{F}_t)_{t \geq 0}$ -stopping time  $\tau_\varepsilon = \inf\{t \geq 0; V(\Phi_t) \geq -\varepsilon\}$  for every  $\varepsilon \geq 0$ . Then*

$$\tau_\varepsilon = \begin{cases} r_\varepsilon(\Phi_0), & \text{if } r_\varepsilon(\Phi_0) < T_1 \\ T_1 + \tau_\varepsilon \circ \theta_{T_1}, & \text{if } r_\varepsilon(\Phi_0) \geq T_1 \end{cases} \quad \text{and} \quad \tau_\varepsilon 1_{\{T_n \leq \tau_\varepsilon < T_{n+1}\}} = [T_n + r_\varepsilon(\Phi_{T_n})] 1_{\{T_n \leq \tau_\varepsilon < T_{n+1}\}}$$

for every  $\varepsilon \geq 0$  and  $n \geq 0$ ,

where  $r_\varepsilon(\phi) := \inf\{t \geq 0; V(\varphi(t, \phi)) \geq -\varepsilon\}$  for every  $\phi \in \mathbb{R}_+^m$  and  $\varepsilon \geq 0$ .

Proposition 3.4 states that postponing the stopping decision until time  $\tau_\varepsilon$  does not cause any regrets, and this observation almost immediately leads to the  $\varepsilon$ -optimality of  $\tau_\varepsilon$  for the problem in (2.3), which is established by Theorem 3.1.

**Proposition 3.4.** *Let us define  $M_t := \int_0^t e^{-\lambda u} g(\Phi_u) du + e^{-\lambda t} V(\Phi_t)$  for every  $t \geq 0$ . For every  $n \geq 0$ ,  $(\mathcal{F}_t)_{t \geq 0}$ -stopping time  $\tau$ , and  $\phi \in \mathbb{R}_+^m$ , we have  $\mathbb{E}_\infty^\phi[M_{\tau \wedge \tau_\varepsilon \wedge T_n}] = \mathbb{E}_\infty^\phi[M_0] = V(\phi)$ .*

**Theorem 3.1.** For every  $\varepsilon \geq 0$ , the  $(\mathcal{F}_t)_{t \geq 0}$ -stopping time  $\tau_\varepsilon$  of Lemma 3.4 is  $\varepsilon$ -optimal for problem in (2.3). Particularly,  $\tau_0$  is an optimal  $(\mathcal{F}_t)_{t \geq 0}$ -stopping time for problem in (2.3).

*Proof.* By Proposition 3.4 for every  $\varepsilon \geq 0$ ,  $n \geq 0$ , and  $\tau \equiv T_n$ , we have

$$\begin{aligned} V(\phi) &= \mathbb{E}_\infty^\phi[M_{\tau_\varepsilon \wedge T_n}] \\ &= \mathbb{E}_\infty^\phi \left[ \int_0^{\tau_\varepsilon \wedge T_n} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau_\varepsilon < T_n\}} e^{-\lambda \tau_\varepsilon} V(\Phi_{\tau_\varepsilon}) + 1_{\{\tau_\varepsilon \geq T_n\}} e^{-\lambda T_n} V(\Phi_{T_n}) \right] \\ &\geq \mathbb{E}_\infty^\phi \left[ \int_0^{\tau_\varepsilon \wedge T_n} e^{-\lambda t} g(\Phi_t) dt - \varepsilon 1_{\{\tau_\varepsilon < T_n\}} e^{-\lambda \tau_\varepsilon} - 1_{\{\tau_\varepsilon \geq T_n\}} e^{-\lambda T_n} \right] \\ &\geq \mathbb{E}_\infty^\phi \left[ \int_0^{\tau_\varepsilon \wedge T_n} e^{-\lambda t} g(\Phi_t) dt \right] - \varepsilon - \mathbb{E}_\infty^\phi[e^{-\lambda T_n}]. \end{aligned}$$

Since  $\int_0^{\tau_\varepsilon \wedge T_n} e^{-\lambda t} g(\Phi_t) dt$ ,  $n \geq 0$  is bounded from below, and  $\lim_{n \rightarrow \infty} \mathbb{E}_\infty e^{-\lambda T_n} = 0$ , Fatou's lemma gives  $V(\phi) \geq \underline{\lim}_{k \rightarrow \infty} \mathbb{E}_\infty^\phi \left[ \int_0^{\tau_\varepsilon \wedge T_n} e^{-\lambda t} g(\Phi_t) dt \right] - \varepsilon \geq \mathbb{E}_\infty^\phi \left[ \int_0^{\tau_\varepsilon} e^{-\lambda t} g(\Phi_t) dt \right] - \varepsilon$ .  $\square$

Since  $V(\cdot) = \lim_{n \rightarrow \infty} V_n(\cdot)$  can be calculated in the limit, optimal stopping rule  $\tau_0$  may not be implementable. In practice,  $V(\cdot)$  is approximated by  $V_n(\cdot) \equiv v_n(\cdot)$  for some sufficiently large  $n \geq 0$ , and optimal performance of  $\tau_0$  can be approximated arbitrarily closely by stopping times  $\sigma_{n,\varepsilon}$ ,  $\varepsilon > 0$  of Theorem 3.2.

**Theorem 3.2.** Define  $\sigma_{n,\varepsilon} = \inf\{t \geq 0; V_n(\Phi_t) \geq -\varepsilon\}$  for every  $\varepsilon \geq 0$  and  $n \geq 0$ . If  $N(\varepsilon) := \min\{n \geq 0; [\lambda/(\lambda + \lambda_0)]^n \leq \varepsilon\} = \lfloor \log \varepsilon / \log \frac{\lambda}{\lambda + \lambda_0} \rfloor$  for every  $\varepsilon > 0$ , then  $(\mathcal{F}_t)_{t \geq 0}$ -stopping times  $\sigma_{N(\varepsilon/2), \varepsilon/2}$  and  $\sigma_{N(\varepsilon), 0}$  are  $\varepsilon$ -optimal for every  $\varepsilon > 0$  for the problem in (2.3).

*Proof.* Because  $V_n(\cdot) \geq V(\cdot)$ , we have  $\mathbb{P}_\infty$ -a.s.  $\sigma_{n,\varepsilon} \leq \tau_\varepsilon$ , and Proposition 3.4 with  $\tau \equiv \sigma_{n,\varepsilon}$  implies  $\mathbb{E}_\infty^\phi[M_{\sigma_{n,\varepsilon} \wedge T_k}] = \mathbb{E}_\infty^\phi[M_{\sigma_{n,\varepsilon} \wedge \tau_\varepsilon \wedge T_k}] = \mathbb{E}_\infty^\phi[M_0] - V(\phi)$  for every  $k \geq 0$ . Therefore,

$$\begin{aligned} V(\phi) &= \mathbb{E}_\infty^\phi[M_{\sigma_{n,\varepsilon} \wedge T_k}] = \mathbb{E}_\infty^\phi \left[ \int_0^{\sigma_{n,\varepsilon} \wedge T_k} e^{-\lambda t} g(\Phi_t) dt + e^{-\lambda(\sigma_{n,\varepsilon} \wedge T_k)} V(\Phi_{\sigma_{n,\varepsilon} \wedge T_k}) \right] \\ &= \mathbb{E}_\infty^\phi \left[ \int_0^{\sigma_{n,\varepsilon} \wedge T_k} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\sigma_{n,\varepsilon} < T_k\}} e^{-\lambda \sigma_{n,\varepsilon}} V(\Phi_{\sigma_{n,\varepsilon}}) + 1_{\{\sigma_{n,\varepsilon} \geq T_k\}} e^{-\lambda T_k} V(\Phi_{T_k}) \right]. \end{aligned}$$

Lemma 3.1 gives  $0 \leq V_n(\phi) - V(\phi) \leq [\lambda/(\lambda + \lambda_0)]^n$  for every  $\phi \in \mathbb{R}_+^m$  and  $n \geq 0$ , and  $V_n(\cdot) \geq -1$  by Corollary 3.2. Therefore,  $V(\phi)$  is greater than or equal to

$$\mathbb{E}_\infty^\phi \left[ \int_0^{\sigma_{n,\varepsilon} \wedge T_k} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\sigma_{n,\varepsilon} < T_k\}} e^{-\lambda \sigma_{n,\varepsilon}} \left( V_n(\Phi_{\sigma_{n,\varepsilon}}) - \left( \frac{\lambda}{\lambda + \lambda_0} \right)^n \right) - 1_{\{\sigma_{n,\varepsilon} \geq T_k\}} e^{-\lambda T_k} \right].$$

The rest of the proof is similar to that of Theorem 3.1 and is completed by the definition of  $N(\varepsilon)$ .  $\square$

**Theorem 3.3.** Let us define stopping regions

$$\Gamma := \{\phi \in \mathbb{R}_+^m; V(\phi) = 0\} \quad \text{and} \quad \Gamma_n := \{\phi \in \mathbb{R}_+^m; V_n(\phi) = 0\}, \quad n \geq 1.$$

Then the sets  $\mathbb{R}_+^m = \Gamma_0 \supseteq \Gamma_1 \supseteq \dots \supseteq \Gamma_n \supseteq \dots \supseteq \Gamma$ , which are closed and convex, and  $\bigcap_{k \geq 0} \Gamma_k = \Gamma$ . We have that  $\tau_0 = \inf\{t \geq 0; \Phi_t \in \Gamma\}$  and  $\sigma_{n,0} = \inf\{t \geq 0; \Phi_t \in \Gamma_n\}$  for every  $n \geq 0$ . Moreover, the sequence  $(\sigma_{n,0})_{n \geq 0}$  increases  $\mathbb{P}_\infty$ -a.s. to  $\tau_0$  as  $n \rightarrow \infty$ .

*Proof.* Since  $V_n(\cdot)$  decreases to  $V(\cdot)$  as  $n \rightarrow \infty$ , we have  $\Gamma_0 \supseteq \Gamma_1 \supseteq \dots \supseteq \Gamma_n \supseteq \dots \supseteq \Gamma$ , which are closed and concave because  $V_n(\cdot)$ ,  $n \geq 0$  and  $V(\cdot)$  are continuous, concave, and nonpositive by Corollary 3.2. The stopping times  $\tau_0$  and  $\sigma_{n,0}$ ,  $n \geq 0$  of Theorems 3.1 and 3.2 are by definition the first hitting times of process  $\Phi$  to stopping sets  $\Gamma$  and  $\Gamma_n$ ,  $n \geq 0$ . Because the sets  $\Gamma_n$ ,  $n \geq 0$  are decreasing, the hitting times  $\sigma_{n,0}$ ,  $n \geq 0$  are increasing with  $\sigma_0 := \lim_{n \rightarrow \infty} \sigma_{n,0} \leq \tau_0$ .

Because  $\{\Phi_t, t \geq 0\}$  has left-limits, the limit  $\lim_{n \rightarrow \infty} \Phi_{\sigma_{n,0}}$  exists. Since jump times of process  $\Phi$  are totally unpredictable,  $\mathbb{P}_\infty\{\sigma_0 = T_n \text{ for some } n \geq 1\} = 0$ . Therefore,  $\mathbb{P}_\infty$ -a.s.  $\lim_{n \rightarrow \infty} \Phi_{\sigma_{n,0}} = \Phi_{\sigma_0}$ .

On  $\{\sigma_0 = \infty\}$ , we obviously have  $\mathbb{P}_\infty$ -a.s.  $\sigma_0 = \tau_0$ . On  $\{\sigma_0 < \infty\}$ , we have  $\sigma_{n,0} < \infty$  for every  $n \geq 0$  and  $\Phi_{\sigma_{n,0}} \in \Gamma_n \subseteq \Gamma_k$  for every  $n \geq k$  because  $\Gamma_n$ ,  $n \geq 0$  are closed and  $t \mapsto \Phi_t$  is right-continuous. Therefore,  $\Phi_{\sigma_0} = \lim_{n \rightarrow \infty} \Phi_{\sigma_{n,0}} \in \Gamma_k$  for every  $k \geq 0$ , equivalently  $\Phi_{\sigma_0} \in \bigcap_{k \geq 0} \Gamma_k$  on  $\{\sigma_0 < \infty\}$ . Thus, we will have proved that  $\sigma_0 \geq \tau_0$  on  $\{\sigma_0 < \infty\}$  as well, if we show that  $\bigcap_{k \geq 0} \Gamma_k = \Gamma$ .

We already know that  $\bigcap_{k \geq 0} \Gamma_k \supseteq \Gamma$ . To prove the opposite inclusion, take any  $\phi \in \bigcap_{k \geq 0} \Gamma_k$ . Then  $0 = V_k(\phi)$  for every  $k \geq 0$ . Therefore,  $V(\phi) = \lim_{n \rightarrow \infty} V_k(\phi) = 0$  and  $\phi \in \Gamma$ . Hence,  $\bigcap_{k \geq 0} \Gamma_k \subseteq \Gamma$ . □

Finally, Proposition 2.1 or the first part of Proposition 2.3 guarantee that  $\tau_0$  is a Bayes-optimal alarm time, and for every  $\varepsilon > 0$ , stopping time  $\sigma_{N(\varepsilon),0}$  is an  $\varepsilon$ -optimal alarm time for the original compound Poisson disorder problem.

#### 4. AN ILLUSTRATION

Consider the compound Poisson disorder problem with  $\lambda_0 \neq \lambda_1$ ,  $v_0 \equiv v_1$ , and  $m = 2$ ; namely, the detection delay penalty cost function is  $f(t) = t^2$ . We shall use the results of Section 3 to identify as explicitly as possible the structure of the optimal solution of the auxiliary optimal stopping problem in (2.3).

The sufficient statistic is the two-dimensional piecewise deterministic strong Markov process  $\Phi = \{\Phi_t = (\Phi_t^{(1)}, \Phi_t^{(2)}); t \geq 0\}$  which follows the dynamics

$$\Phi_t = \left\{ \begin{array}{ll} \varphi(t - T_n, \Phi_{T_n}), & \text{if } t \in [T_n, T_{n+1}) \\ \varphi(T_{n+1} - T_n, \Phi_{T_n}), \frac{\lambda_1}{\lambda_0}, & \text{if } t = T_{n+1} \end{array} \right\} \quad \text{and}$$

$$\begin{aligned} \varphi(t, \phi) &= e^{At} \phi + \left( \int_0^t e^{A(t-s)} ds \right) b \\ &\text{for every } t \geq 0, \phi \in \mathbb{R}_+^2, \text{ and } n \geq 0, \end{aligned}$$

where  $\bar{\lambda} = \lambda_1 - \lambda_0 - \lambda$ , and

$$\begin{aligned} A &= \begin{bmatrix} -\bar{\lambda} & 1 \\ 0 & -\bar{\lambda} \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 2\lambda \end{bmatrix}, \quad e^{At} = e^{-\bar{\lambda}t} \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \\ A^{-1}b &= - \begin{bmatrix} \bar{\lambda}^{-1} & \bar{\lambda}^{-2} \\ 0 & \bar{\lambda}^{-1} \end{bmatrix} \begin{bmatrix} 0 \\ 2\lambda \end{bmatrix} = - \begin{bmatrix} 2\lambda\bar{\lambda}^{-2} \\ 2\lambda\bar{\lambda}^{-1} \end{bmatrix} \quad \text{if } \bar{\lambda} \neq 0, \end{aligned}$$

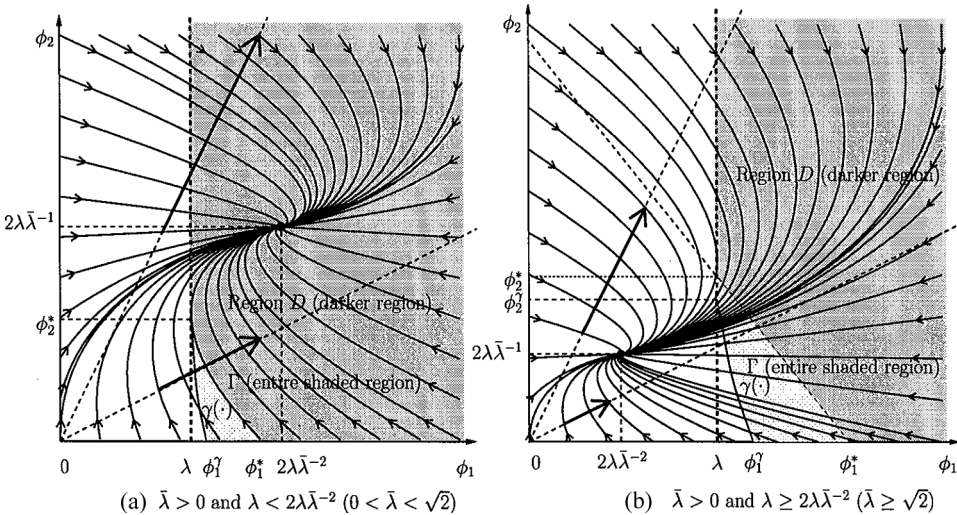
$$\varphi(t, \phi) = \begin{cases} \left[ \begin{array}{l} e^{-\bar{\lambda}t}(\phi_1 - 2\lambda\bar{\lambda}^{-2} + t(\phi_2 - 2\lambda\bar{\lambda}^{-1})) + 2\lambda\bar{\lambda}^{-2} \\ e^{-\bar{\lambda}t}(\phi_2 - 2\lambda\bar{\lambda}^{-1}) + 2\lambda\bar{\lambda}^{-1} \end{array} \right], & \text{if } \bar{\lambda} \neq 0 \\ \left[ \begin{array}{l} \phi_1 + \phi_2 t + \lambda t^2 \\ \phi_2 + 2\lambda t \end{array} \right], & \text{if } \bar{\lambda} = 0 \end{cases}$$

for every  $t \geq 0$  and  $\phi = (\phi_1, \phi_2) \in \mathbb{R}_+^2$ .

Depending on the relationships between the parameters of the problem, the sample paths of process  $\Phi$  can take one of two major forms, and each can further be divided into two subcases. We will describe qualitatively the form of the optimal solution of the problem in (2.3) for each case. Note, however, that under all circumstances it is never optimal to stop before the process  $\Phi$  leaves the strip  $C_0 := \{\phi = (\phi_1, \phi_2) \in \mathbb{R}_+^2; \phi_1 < \lambda\}$ , because the integrand in (2.3) remains negative until the first exit time  $\tau_0 := \inf\{t \geq 0; \Phi_t \notin C_0\} = \inf\{t \geq 0; \Phi_t^{(1)} \geq \lambda\}$  of process  $\Phi$  from  $C_0$ .

**4.1. Case I:  $\bar{\lambda} > 0$**

The solution  $x(t) = \varphi(t, \phi)$  of the system of linear ordinary differential equations  $dx(t)/dt = Ax(t) + b$  with initial condition  $x(0) = \phi \in \mathbb{R}_+^2$  has unique equilibrium point at  $-A^{-1}b = [2\lambda\bar{\lambda}^{-2} \ 2\lambda\bar{\lambda}^{-1}]^T$ . Since  $\bar{\lambda} > 0$ , we have  $\lambda_1/\lambda_0 > 1$ , in which case at each mark arrival time, the process  $\Phi$  jumps away from the origin along the ray emanating from the origin and passing through the position of  $\Phi$  before jump; see Figure 1. The structure of the optimal solution depends on the position of the root  $\lambda$  of the running cost function  $g(\phi) = \phi_1 - \lambda$  in (2.3) relative to the first coordinate  $2\lambda\bar{\lambda}^{-2}$  of the equilibrium point  $-A^{-1}b$ .



**Figure 1.** Case I:  $\bar{\lambda} > 0$ . The sufficient statistic  $\Phi$  follows the integral curves of a system of two linear ordinary differential equations, which have unique equilibrium point. Moreover, since  $\lambda_1 > \lambda_0$ , at every arrival time of a mark,  $\Phi$  jumps away from the origin.



Case I(a).  $\bar{\lambda} > 0$  and  $\lambda < 2\lambda\bar{\lambda}^{-2}$  (equivalently,  $0 < \bar{\lambda} < \sqrt{2}$ ). Let  $\phi_1 = \phi_1^* \geq 0$  be the unique number such that  $\lambda$  equals the unique minimum of the mapping  $t \mapsto \varphi_1(t, (\phi_1; 0))$ , which it is attained at  $t = t^*(\phi_1)$ . Let us denote by  $\phi_2^* = \varphi_2(t^*(\phi_1^*), (\phi_1^*, 0))$  the second coordinate at time  $t^*(\phi_1^*)$ , when the minimum value of the first coordinate is attained and equals  $\lambda$ , starting initially at  $(\phi_1^*, 0)$ ; see Figure 1(a). For every fixed  $\phi_1 \geq 0$ , taking the derivative of  $\varphi_1(t, (\phi_1, 0)) = e^{-\lambda t}(\phi_1 - 2\lambda\bar{\lambda}^{-2} - 2\lambda\bar{\lambda}^{-1}t) + 2\lambda\bar{\lambda}^{-2}$  with respect to  $t$  gives  $\partial\varphi_1(t, (\phi_1, 0))/\partial t = e^{-\lambda t}(-\bar{\lambda}\phi_1 + \lambda t)$ , and equating it to zero and solving for  $t$  leads to  $t^*(\phi_1) = \bar{\lambda}\lambda^{-1}\phi_1$  for every  $\phi_1 \geq 0$ . Therefore,  $\lambda = \varphi_1(t^*(\phi_1^*), (\phi_1^*, 0)) = 2\lambda\bar{\lambda}^{-2}[1 - e^{-(\bar{\lambda}^2/\lambda)\phi_1^*}]$  gives

$$\phi_1^* = -\lambda\bar{\lambda}^{-2} \ln\left(1 - \frac{\bar{\lambda}^2}{2}\right) \quad \text{and} \quad \phi_2^* = \lambda\bar{\lambda}.$$

Since  $\bar{\lambda}^2/2 \in (0, 1)$ , we have  $0 < \phi_1^* < \infty$ . Moreover, (i)  $\phi_1^* \leq 2\lambda\bar{\lambda}^{-2}$  if and only if  $\bar{\lambda} \leq \sqrt{2(1 - e^{-2})}$ , (ii)  $\phi_2^* < 2\lambda\bar{\lambda}^{-2}$ , which is the second coordinate of the equilibrium point  $-A^{-1}b$ . Let us define

$$D := \{(\phi_1, \phi_2) \in \mathbb{R}_+^2; \phi_1 = \varphi_1(t, (\phi_1^*, 0)), \phi_2 \geq \varphi_2(t, (\phi_1^*, 0)), 0 \leq t \leq t^*(\phi_1^*)\} \\ \cup ((\phi_1^*, \infty) \times \mathbb{R}_+),$$

which is the dark shaded region in Figure 1(a). Because  $\lambda_1/\lambda_0 > 1$  and the equilibrium point  $-A^{-1}b$  belongs to  $D$ , we have  $(\lambda_1/\lambda_0)D \subseteq D$ , and  $\mathbb{P}_\infty^\phi\{\Phi_t \in D \text{ for every } t \geq 0\} = 1$  for every  $\phi \in D$ . Since  $g(\phi) - \lambda \geq 0$  for every  $\phi \in D \subset \mathbb{R}_+^2 \setminus C_0$ , we have  $\mathbb{E}_\infty^\phi[\int_0^\tau e^{-\lambda t} g(\Phi_t) dt] \geq 0$  for every  $\tau \in \mathcal{S}$  and  $\phi \in D$ . Therefore,  $V(\phi) = 0$  for every  $\phi \in D$ , and  $D$  is a subset of the optimal stopping region  $\Gamma = \{\phi \in \mathbb{R}_+^2; V(\phi) = 0\}$ . Because  $D \subset \Gamma \subseteq \mathbb{R}_+^2 \setminus C_0$  and  $\Gamma$  is closed and convex, the optimal stopping boundary  $\partial\Gamma$  coincides with the infinite line segment  $\{(\lambda, \phi_2); \phi_2 \geq \phi_2^*\}$  and with some nondecreasing convex continuous curve  $\gamma : [\lambda, \phi_1^*] \mapsto \mathbb{R}$  such that  $\gamma(\lambda) = \phi_2^*$ . There is also some  $\lambda < \phi_1^\gamma \leq \phi_1^*$  such that  $\gamma(\cdot)$  is strictly decreasing on  $[\lambda, \phi_1^\gamma]$  and equals zero on  $[\phi_1^\gamma, \phi_1^*]$ ; see Figure 1(a). All of those conclusions are direct consequences of the convexity of the optimal stopping region  $\Gamma$ .

In this subcase, starting initially at any  $\Phi_0 = \phi$  on the vertical axis (namely,  $\phi = (0, \phi_2)$  for any  $\phi_2 \geq 0$ ), the process  $\Phi$  never returns to  $C_0$  once it leaves that region. Therefore, the first exit time  $\tau_0$  of  $\Phi$  from  $C_0$  is optimal for the problem in (2.3) if  $\Phi_0 = (0, \phi_2)$  for some  $\phi_2 \geq 0$ . Since by Proposition 2.2 we have

$$\inf_{\tau \in \mathcal{S}} R_\tau(p) = 1 - p + (1 - p)V\left(0, \frac{2p}{1 - p}\right), \quad 0 \leq p < 1$$

depends on  $V(\phi)$  evaluated on  $\{\phi = (0, \phi_2); \phi_2 \geq 0\}$ , the  $(\mathcal{F}_t)_{t \geq 0}$  stopping time  $\tau_0$  is an optimal change-detection alarm time if  $0 < \bar{\lambda} \leq \sqrt{2}$ .

Case I(b).  $\bar{\lambda} > 0$  and  $\lambda \geq 2\lambda\bar{\lambda}^{-2}$  (equivalently,  $\bar{\lambda} \geq \sqrt{2}$ ). We shall first state and prove a comparison lemma for the sample paths of the process  $\Phi$ .

**Lemma 4.1.** *For every  $i = 1, 2$ , it we have  $\mathbb{P}_\infty$ -a.s.  $\Phi_t^{(i)} \geq \varphi_i(t, \Phi_0)$  for every  $t \geq 0$ .*

*Proof.* Clearly,  $\Phi_t^{(i)} \geq \varphi_i(t, \Phi_0)$  for every  $0 \leq t < T_1$  and  $i = 1, 2$ . Suppose that for some  $n \geq 1$ , we have  $\mathbb{P}_\infty$ -a.s.  $\Phi_t^{(i)} \geq \varphi_i(t, \Phi_0)$  for every  $0 \leq t < T_n$  and  $i = 1, 2$ . Let

us prove that the same inequality also holds  $\mathbb{P}_\infty$ -a.s. for  $T_n \leq t < T_{n+1}$ , and hence for  $0 \leq t < T_{n+1}$ , which will then complete the proof of the lemma since  $\mathbb{P}_\infty$ -a.s.  $T_n \uparrow \infty$  as  $n \rightarrow \infty$ .

It is clear from the explicit form of  $\varphi(\cdot, \cdot)$  in (2.2) that, if  $x = (x_1, x_2)$  and  $y = (y_1, y_2)$  are in  $\mathbb{R}_+^2$  such that  $x_1 \leq y_1$  and  $x_2 \leq y_2$ , then  $\varphi_i(t, x) \leq \varphi_i(t, y)$  for every  $t \geq 0$  and  $i = 1, 2$ . Because  $\lambda_1/\lambda_0 > 1$ , we have  $\Phi_{T_n}^{(i)} = (\lambda_1/\lambda_0)\Phi_{T_n-}^{(i)} \geq \Phi_{T_n-}^{(i)}$  for  $i = 1, 2$ , and

$$\begin{aligned} \Phi_t^{(i)} &= \varphi_i(t - T_n, \Phi_{T_n}) = \varphi_i\left(t - T_n, \frac{\lambda_1}{\lambda_0}\Phi_{T_n-}\right) \\ &\geq \varphi_i(t - T_n, \Phi_{T_n-}), \quad T_n \leq t < T_{n+1}, \quad i = 1, 2. \end{aligned}$$

Since  $\Phi_{T_n-}^{(i)} \geq \varphi_i(T_n, \Phi_0)$  for  $i = 1, 2$  by the induction hypothesis, we have  $\Phi_t^{(i)} \geq \varphi_i(t - T_n, \Phi_{T_n-}) \geq \varphi_i(t - T_n, \varphi_i(T_n, \Phi_0)) = \varphi_i(t - T_n + T_n, \Phi_0) = \phi_i(t, \Phi)$  for every  $T_n \leq t < T_{n+1}$  and  $i = 1, 2$ .  $\square$

Lemma 4.1 implies that

$$\begin{aligned} V(\phi) &= \inf_{\tau \in \mathcal{S}} \mathbb{E}_\infty^\phi \left[ \int_0^\tau e^{-\lambda t} g(\Phi_t) dt \right] = \inf_{\tau \in \mathcal{S}} \mathbb{E}_\infty^\phi \left[ \int_0^\tau e^{-\lambda t} (\Phi_t^{(1)} - \lambda) dt \right] \\ &\geq \inf_{\tau \in \mathcal{S}} \mathbb{E}_\infty^\phi \left[ \int_0^\tau e^{-\lambda t} (\varphi_1(t, \Phi_0) - \lambda) dt \right] \\ &= \inf_{r \geq 0} \int_0^r e^{-\lambda t} (\varphi_1(t, \phi) - \lambda) dt =: h(\phi), \quad \phi \in \mathbb{R}_+^2. \end{aligned}$$

Therefore,  $\{\phi \in \mathbb{R}_+^2; h(\phi) = 0\} \subseteq \{\phi \in \mathbb{R}_+^2; V(\phi) = 0\} \equiv \Gamma$ . On the other hand, for every  $\phi = (\phi_1, \phi_2) \in \mathbb{R}_+^2$  such that  $\phi_1 \geq \lambda$ , we have  $h(\phi) = \min\{0, \int_0^\infty e^{-\lambda t} (\varphi_1(t, \phi) - \lambda) dt\}$ , and  $h(\phi) = 0$  if

$$\begin{aligned} 0 &\leq \int_0^\infty e^{-\lambda t} (\varphi_1(t, \phi) - \lambda) dt \\ &= \int_0^\infty e^{-\lambda t} [e^{-\bar{\lambda} t} (\phi_1 - 2\lambda\bar{\lambda}^{-2} + t(\phi_2 - 2\lambda\bar{\lambda}^{-1})) + 2\lambda\bar{\lambda}^{-2} - \lambda] dt \\ &= \frac{\phi_1}{\lambda_1 - \lambda_0} + \frac{\phi_2}{(\lambda_1 - \lambda_0)^2} - \frac{2\lambda\bar{\lambda}^{-2}}{\lambda_1 - \lambda_0} - \frac{2\lambda\bar{\lambda}^{-1}}{(\lambda_1 - \lambda_0)^2} + 2\bar{\lambda}^{-2} - 1, \end{aligned}$$

and after multiplying both sides by  $(\lambda_1 - \lambda_0)^2$  and rearranging the terms we obtain

$$\phi_2 \geq \bar{\gamma}(\phi_1) := -(\lambda_1 - \lambda_0)\phi_1 + 2\lambda\bar{\lambda}^{-2}(\lambda_1 - \lambda_0) + 2\lambda\bar{\lambda}^{-1} + (1 - 2\bar{\lambda}^{-2})(\lambda_1 - \lambda_0)^2.$$

Hence, we have

$$\begin{aligned} D &:= \{(\phi_1, \phi_2) \in \mathbb{R}_+^2; \phi_1 \geq \lambda, \phi_2 \geq \bar{\gamma}(\phi_1)\} \\ &\subseteq \{\phi \in \mathbb{R}_+^2; h(\phi) = 0\} \subseteq \{\phi \in \mathbb{R}_+^2; V(\phi) = 0\} = \Gamma. \end{aligned}$$

**Lemma 4.2.** *Let  $\phi_1 = \phi_1^*$  be the root  $\bar{\gamma}(\phi) = 0$  and define  $\phi_2^* := \bar{\gamma}(\lambda)$ . Then*

$$\phi_1^* = \lambda + \frac{\phi_2^*}{\lambda_1 - \lambda_0} > \lambda \quad \text{and} \quad \phi_2^* = 2\lambda\bar{\lambda}^{-1} + (1 - 2\bar{\lambda}^{-2})(\lambda_1 - \lambda_0)\bar{\lambda} > 2\lambda\bar{\lambda}^{-1}.$$

*Proof.* Direct calculation gives  $\phi_2^* = \bar{\gamma}(\lambda) = 2\lambda\bar{\lambda}^{-1} + (1 - 2\bar{\lambda}^{-2})(\lambda_1 - \lambda_0)\bar{\lambda} > 2\lambda\bar{\lambda}^{-1}$ , because  $\bar{\lambda} \geq \sqrt{2}$  implies that  $1 - 2\bar{\lambda}^{-2} > 0$ ,  $\lambda_1 - \lambda_0 > \lambda > 0$ ,  $\bar{\lambda} > 0$ . Because  $\phi_1 \rightarrow \bar{\gamma}(\phi_1)$  is a straight line with slope  $-(\lambda_1 - \lambda_0)$ , we have  $(0 - \phi_2^*)/(\phi_1^* - \lambda) = -(\lambda_1 - \lambda_0)$ .

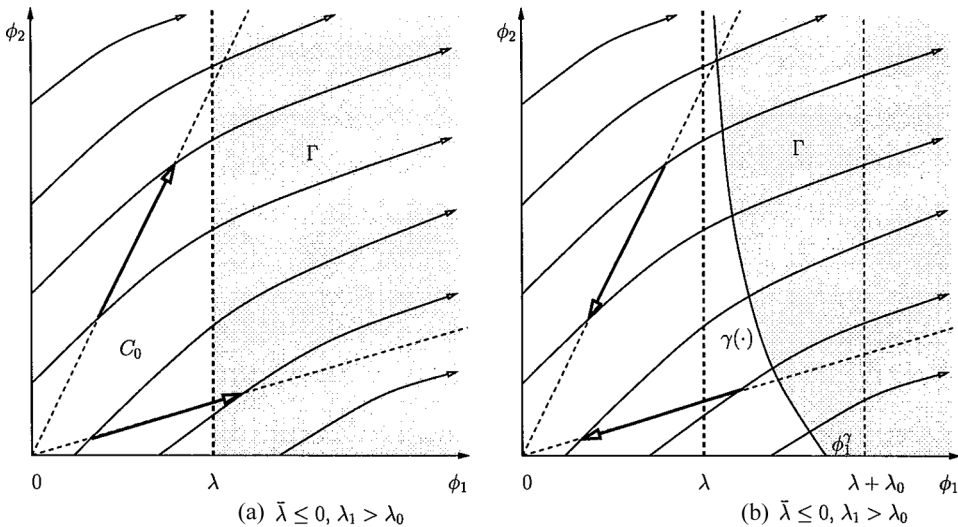
Because optimal stopping region  $\Gamma$  is closed and convex, and  $D \subseteq \Gamma \subseteq \mathbb{R}_+^2 \setminus C_0$ , there exist some  $2\lambda\bar{\lambda}^{-1} < \phi_2^\gamma < \phi_2^*$ ,  $\lambda < \phi_1^\gamma \leq \phi_1^*$ , and some nondecreasing convex continuous curve  $\gamma : [\lambda, \phi_1^*] \mapsto \mathbb{R}$  such that optimal stopping boundary  $\partial\Gamma$  coincides with the infinite line segment  $\{(\lambda, \phi_2); \phi_2 \geq \phi_2^\gamma\}$  and with  $\{(\phi_1, \gamma(\phi_1)); \lambda \leq \phi_1 \leq \phi_1^*\}$ . Moreover,  $\phi_1 \mapsto \gamma(\phi_1)$  is strictly decreasing on  $\phi_1 \in [\lambda, \phi_1^\gamma]$  and equals zero on  $\phi_1 \in [\phi_1^\gamma, \phi_1^*]$ ; see Figure 1(b).

**4.2. Case II:  $\bar{\lambda} < 0$**

The components of  $t \mapsto \varphi(t, \phi)$  are strictly increasing for all  $\phi \in \mathbb{R}_+^2$ . Both  $\lambda_1 > \lambda_0$  and  $\lambda_1 < \lambda_0$  are possible.

*Case II(a).*  $\bar{\lambda} < 0$  and  $\lambda_1 > \lambda_0$ . The process  $\Phi$  runs away from the origin both at and between jump times. It never returns to region  $C_0$  once it leaves that region. Therefore, optimal stopping region  $\Gamma$  coincides with  $\mathbb{R}_+^2 \setminus C_0$ , optimal stopping boundary  $\partial\Gamma$  is the straight line  $\phi_1 = \lambda$ , and the first exit time  $\tau_0$  of process  $\Phi$  from region  $C_0$  is optimal for the problem in (2.3) and an optimal alarm time for the compound Poisson disorder problem; see Figure 2(a).

*Case II(b).*  $\bar{\lambda} < 0$  and  $\lambda_1 < \lambda_0$ . The process  $\Phi$  runs away from the origin between jump times, but is pulled back toward the origin at every jump. Therefore,



**Figure 2.** Case II:  $\bar{\lambda} \leq 0$ . The process  $\Phi$  follows integral curves, both coordinates of which are strictly increasing. It jumps away from the origin if  $\lambda_1 > \lambda_0$  and toward the origin if  $\lambda_1 < \lambda_0$ .

$\Phi$  may return to region  $C_0$  after a jump with positive probability; see Figure 2(b). Since  $V(\cdot) \geq -1$ , for every  $\phi \in [\lambda + \lambda_0, \infty) \times \mathbb{R}_+$  we have  $V(\phi)$  equals

$$\begin{aligned} (J_0 V)(\phi) &= \inf_{r \geq 0} \int_0^r e^{-(\lambda + \lambda_0)t} [g + \lambda(KV)](\varphi(t, \phi)) dt \\ &\geq \inf_{r \geq 0} \int_0^r e^{-(\lambda + \lambda_0)t} [\varphi_1(t, \phi) - \lambda - \lambda_0] dt = 0, \end{aligned}$$

which implies that  $[\lambda + \lambda_0, \infty) \times \mathbb{R}_+ \subseteq \Gamma$ . Because the optimal stopping region  $\Gamma \subseteq \mathbb{R}_+ \setminus C_0$  is closed and convex, there exist some  $\lambda < \phi_1^\gamma \leq \lambda + \lambda_0$  and some nonincreasing convex continuous curve  $\gamma : [\lambda, \lambda + \lambda_0] \mapsto \mathbb{R}$  such that the optimal stopping boundary  $\partial\Gamma$  coincides with  $\gamma(\cdot)$ , which is strictly decreasing on  $[\lambda, \phi_1^\gamma]$  and vanishes on  $[\phi_1^\gamma, \lambda + \lambda_0]$ .

## APPENDIX: PROOFS OF SELECTED RESULTS

*Proof of Proposition 3.2.* From the definitions we immediately have  $V_0(\phi) = v_0(\phi) = 0$  for every  $\phi \in \mathbb{R}_+^m$ . For every  $n \geq 1$  and  $\tau \in \mathcal{S}$  we shall prove that

$$\begin{aligned} &\mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_k} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_k\}} e^{-\lambda T_k} v_{n-k}(\Phi_{T_k}) \right] \\ &\geq \mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_{k-1}} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_{k-1}\}} e^{-\lambda T_{k-1}} v_{n-k+1}(\Phi_{T_{k-1}}) \right] \quad \text{for every } 1 \leq k \leq n, \end{aligned} \tag{A.1}$$

which will then imply that

$$\begin{aligned} \mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_n} e^{-\lambda t} g(\Phi_t) dt \right] &= \mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_n} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_n\}} e^{-\lambda T_n} v_0(\Phi_{T_n}) \right] \\ &\geq \mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_0} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_0\}} e^{-\lambda T_0} v_n(\Phi_{T_0}) \right] = v_n(\phi), \end{aligned}$$

and taking the infimum of both sides over all  $\tau \in \mathcal{S}$  gives  $V_n(\phi) = \inf_{\tau \in \mathcal{S}} \mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_n} e^{-\lambda t} g(\Phi_t) dt \right] \geq v_n(\phi)$  for every  $\phi \in \mathbb{R}_+^m$ , which is the conclusion of the proposition. Let us fix any  $n \geq 1$ ,  $\tau \in \mathcal{S}$  and prove (A.1). For every  $1 \leq k \leq n$ , by Proposition 3.1, there is a nonnegative  $\mathcal{F}_{T_{k-1}}$ -measurable random variable  $R_{k-1}$  such that  $1_{\{\tau \geq T_{k-1}\}}(\tau \wedge T_k) = 1_{\{\tau \geq T_{k-1}\}}[(T_{k-1} + R_{k-1}) \wedge T_k]$  and we have

$$\begin{aligned} &\mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_k} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_k\}} e^{-\lambda T_k} v_{n-k}(\Phi_{T_k}) \right] \\ &= \mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_{k-1}} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_{k-1}\}} \int_{T_{k-1}}^{\tau \wedge T_k} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_k\}} e^{-\lambda T_k} v_{n-k}(\Phi_{T_k}) \right] \\ &= \mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_{k-1}} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_{k-1}\}} e^{-\lambda T_{k-1}} \left\{ \int_{T_{k-1}}^{(T_{k-1} + R_{k-1}) \wedge T_k} e^{-\lambda(t - T_{k-1})} \right. \right. \\ &\quad \times g(\varphi(t - T_{k-1}, \Phi_{T_{k-1}})) dt + 1_{\{T_{k-1} + R_{k-1} \geq T_k\}} e^{-\lambda(T_k - T_{k-1})} v_{n-k} \\ &\quad \left. \left. \times \left( \varphi(T_k - T_{k-1}, \Phi_{T_{k-1}}) \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(Z_k) \right) \right\} \right] \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_{k-1}} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_{k-1}\}} e^{-\lambda T_{k-1}} \mathbb{E}_\infty^\phi \left\{ \int_0^{R_{k-1} \wedge (T_k - T_{k-1})} e^{-\lambda t} \right. \right. \\
&\quad \times g(\varphi(t, \Phi_{T_{k-1}})) dt + 1_{\{R_{k-1} \geq T_k - T_{k-1}\}} e^{-\lambda(T_k - T_{k-1})} v_{n-k} \\
&\quad \left. \left. \times \left( \varphi(T_k - T_{k-1}, \Phi_{T_{k-1}}) \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(Z_k) \right) \Big| \mathcal{F}_{T_{k-1}} \right\} \right].
\end{aligned}$$

Because  $R_{k-1}$  and  $\Phi_{T_{k-1}}$  are  $\mathcal{F}_{T_{k-1}}$ -measurable, and  $T_k - T_{k-1}$  and  $Z_k$  are independent of  $\mathcal{F}_{T_{k-1}}$  and have the same distributions as  $T_1$  and  $Z_1$ , respectively, under  $\mathbb{P}_\infty$ , the conditional expectation becomes

$$\begin{aligned}
&\mathbb{E}_\infty^\phi \left\{ \int_0^{R_{k-1} \wedge (T_k - T_{k-1})} e^{-\lambda t} g(\varphi(t, \Phi_{T_{k-1}})) dt \right. \\
&\quad \left. + 1_{\{R_{k-1} \geq T_k - T_{k-1}\}} e^{-\lambda(T_k - T_{k-1})} v_{n-k} \left( \varphi(T_k - T_{k-1}, \Phi_{T_{k-1}}) \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(Z_k) \right) \Big| \mathcal{F}_{T_{k-1}} \right\} \\
&= \mathbb{E}_\infty^\phi \left\{ \int_0^{r \wedge T_1} e^{-\lambda t} g(\varphi(t, \phi)) dt + 1_{\{r \geq T_1\}} e^{-\lambda T_1} v_{n-k} \left( \varphi(T_1, \phi) \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(Z_1) \right) \right\} \Big|_{\substack{r=R_{k-1} \\ \phi=\Phi_{T_{k-1}}}} \\
&= \left\{ \int_0^r e^{-(\lambda+\lambda_0)t} [g + \lambda_0(Kv_{n-k})](\varphi(t, \phi)) dt \right\} \Big|_{\substack{r=R_{k-1} \\ \phi=\Phi_{T_{k-1}}}} = (Jv_{n-k})(\Phi_{T_{k-1}}, R_{k-1}),
\end{aligned}$$

and substituting into previous displayed equation gives

$$\begin{aligned}
&\mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_k} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_k\}} e^{-\lambda T_k} v_{n-k}(\Phi_{T_k}) \right] \\
&= \mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_{k-1}} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_{k-1}\}} e^{-\lambda T_{k-1}} (Jv_{n-k})(\Phi_{T_{k-1}}, R_{k-1}) \right] \\
&\geq \mathbb{E}_\infty^\phi \left[ \int_0^{\tau \wedge T_{k-1}} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \geq T_{k-1}\}} e^{-\lambda T_{k-1}} v_{n-k+1}(\Phi_{T_{k-1}}) \right],
\end{aligned}$$

since  $(Jv_{n-k})(\phi, r) \geq \inf_{t \geq 0} (Jv_{n-k})(\phi, t) = (J_0 v_{n-k})(\phi) = v_{n-k+1}(\phi)$  for every  $r \geq 0$  and  $\phi \in \mathbb{R}_+^m$ , and this completes the proof of (A.1) and Proposition 3.2.  $\square$

*Proof of Proposition 3.3 (By Induction on  $n$ ).* For  $n = 0$ , the last displayed equation of Proposition 3.3 becomes  $0 \leq 0 + \varepsilon$ , which is obviously true for every  $\varepsilon > 0$ . Suppose now that the last inequality of Proposition 3.3 holds for every  $\varepsilon > 0$  for some  $n \geq 0$ . Note that  $\mathbb{E}_\infty^\phi \left[ \int_0^{\tau_{n+1, \varepsilon} \wedge T_{n+1}} e^{-\lambda t} g(\Phi_t) dt \right]$  equals

$$\begin{aligned}
&\mathbb{E}_\infty^\phi \left[ \int_0^{\tau_{n+1, \varepsilon} \wedge T_1} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau_{n+1, \varepsilon} \geq T_1\}} \int_{T_1}^{\tau_{n+1, \varepsilon} \wedge T_{n+1}} e^{-\lambda t} g(\Phi_t) dt \right] \\
&= \mathbb{E}_\infty^\phi \left[ \int_0^{r_{n+1, \varepsilon/2}(\Phi_0) \wedge T_1} e^{-\lambda t} g(\varphi(t, \Phi_0)) dt \right. \\
&\quad \left. + 1_{\{r_{n+1, \varepsilon/2}(\Phi_0) \geq T_1\}} \int_{T_1}^{[T_1 + \tau_{n, \varepsilon/2} \circ \theta_{T_1}] \wedge T_{n+1}} e^{-\lambda t} g(\Phi_t) dt \right].
\end{aligned}$$

By the strong Markov property of process  $\Phi$  at the first jump time  $T_1$ , the conditional expectation of the last integral with respect to  $\mathcal{F}_{T_1}$  becomes

$$\begin{aligned} & \mathbb{E}_\infty^\phi \left[ \int_{T_1}^{[T_1 + \tau_{n,\varepsilon/2} \circ \theta_{T_1}] \wedge T_{n+1}} e^{-\lambda t} g(\Phi_t) dt \middle| \mathcal{F}_{T_1} \right] \\ &= e^{-\lambda T_1} \mathbb{E}_\infty^\phi \left[ \left( \int_0^{\tau_{n,\varepsilon/2} \wedge T_n} e^{-\lambda t} g(\Phi_t) dt \right) \circ \theta_{T_1} \middle| \mathcal{F}_{T_1} \right] \\ &= e^{-\lambda T_1} \mathbb{E}_\infty^{\Phi_{T_1}} \left[ \int_0^{\tau_{n,\varepsilon/2} \wedge T_n} e^{-\lambda t} g(\Phi_t) dt \right] \leq e^{-\lambda T_1} \left( v_n(\Phi_{T_1}) + \frac{\varepsilon}{2} \right) \end{aligned}$$

by the induction hypothesis. Therefore,  $\mathbb{E}_\infty^\phi \left[ \int_0^{\tau_{n+1,\varepsilon} \wedge T_{n+1}} e^{-\lambda t} g(\Phi_t) dt \right]$  equals

$$\begin{aligned} & \mathbb{E}_\infty^\phi \left[ \int_0^{r_{n+1,\varepsilon/2}(\Phi_0) \wedge T_1} e^{-\lambda t} g(\varphi(t, \Phi_0)) dt + 1_{\{r_{n+1,\varepsilon/2}(\Phi_0) \geq T_1\}} e^{-\lambda T_1} \left( v_n(\Phi_{T_1}) + \frac{\varepsilon}{2} \right) \right] \\ & \leq \mathbb{E}_\infty^\phi \left[ \int_0^{r_{n+1,\varepsilon/2}(\Phi_0) \wedge T_1} e^{-\lambda t} g(\varphi(t, \Phi_0)) dt + 1_{\{r_{n+1,\varepsilon/2}(\Phi_0) \geq T_1\}} e^{-\lambda T_1} v_n(\Phi_{T_1}) \right] + \frac{\varepsilon}{2} \\ & = (Jv_n)(\phi, r_{n+1,\varepsilon/2}(\phi)) + \frac{\varepsilon}{2} \leq v_{n+1}(\phi) + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = v_{n+1}(\phi) + \varepsilon. \end{aligned} \quad \square$$

*Proof of Lemma 3.2.* Suppose that  $w : \mathbb{R}_+^m \mapsto \mathbb{R}$  is concave and bounded between  $-1$  and  $0$ . Then  $(Jw)(\phi, r) = \int_0^r e^{-(\lambda+\lambda_0)t} [g + \lambda_0(Kw)](\varphi(t, \phi)) dt \geq \int_0^r e^{-(\lambda+\lambda_0)t} (-\lambda - \lambda_0) dt \geq -1$ . Taking the infimum over  $r \geq 0$  gives  $-1 \leq \inf_{r \geq 0} (Jw)(\phi, r) = (J_0w)(\phi) \leq (Jw)(\phi, 0) = 0$ . Moreover,  $g(\cdot)$  is affine and therefore concave. Because  $\phi \mapsto \varphi(t, \phi)$  is affine for every fixed  $t \geq 0$ , and  $w(\cdot)$  is concave, the mapping  $\phi \mapsto (Kw)(\varphi(t, \phi)) = \int_E w(\varphi(t, \phi) \frac{z_1}{\lambda_0} \frac{dv_1}{dv_0}(z)) v_0(dz)$  is concave. Therefore,  $\phi \mapsto (Jw)(\phi, r)$  is concave for every fixed  $r \geq 0$ . Because the pointwise infimum of every family of concave functions is also concave, the mapping  $\phi \mapsto (J_0w)(\phi) = \inf_{r \geq 0} (Jw)(\phi, r)$  is concave. If  $w_1(\cdot) \leq w_2(\cdot)$ , then  $(Kw_1)(\cdot) \leq (Kw_2)(\cdot)$ ,  $(Jw_1)(\cdot, \cdot) \leq (Jw_2)(\cdot, \cdot)$ , and  $(J_0w_1)(\cdot) \leq (J_0w_2)(\cdot)$ .

Let  $w_1(\cdot)$  and  $w_2(\cdot)$  be two bounded functions on  $\mathbb{R}_+^m$ . Fix any  $\varepsilon > 0$  and  $\phi \in \mathbb{R}_+^m$ . Then there are constants  $r_\varepsilon^{(i)}(\phi)$ ,  $i = 1, 2$  such that  $(Jw_i)(\phi, r_\varepsilon^{(i)}(\phi)) \leq (J_0w_i)(\phi) + \varepsilon$  for every  $i = 1, 2$ . Then

$$\begin{aligned} (J_0w_1)(\phi) - (J_0w_2)(\phi) & \leq (Jw_1)(\phi, r_\varepsilon^{(2)}(\phi)) - (Jw_2)(\phi, r_\varepsilon^{(2)}(\phi)) + \varepsilon \\ & = \int_0^{r_\varepsilon^{(2)}(\phi)} \lambda_0 e^{-(\lambda+\lambda_0)t} (K(w_1 - w_2))(\varphi(t, \phi)) dt + \varepsilon \\ & \leq \|w_1 - w_2\| \int_0^\infty \lambda_0 e^{-(\lambda+\lambda_0)t} dt + \varepsilon \\ & = \|w_1 - w_2\| \frac{\lambda_0}{\lambda + \lambda_0} + \varepsilon. \end{aligned}$$

Similarly,  $(J_0w_1)(\phi) - (J_0w_2)(\phi) \geq -(\|w_1 - w_2\| \frac{\lambda_0}{\lambda + \lambda_0} + \varepsilon)$ , and we have  $|(J_0w_1)(\phi) - (J_0w_2)(\phi)| \leq \frac{\lambda_0}{\lambda + \lambda_0} \|w_1 - w_2\| + \varepsilon$  for every  $\varepsilon > 0$  and  $\phi \in \mathbb{R}_+^m$ . Letting first  $\varepsilon \downarrow 0$  and then taking the supremum over  $\phi \in \mathbb{R}_+^m$  gives the desired inequality.

*Proof of Lemma 3.3.* Note that, since  $\varphi(u, \varphi(s, \phi)) = \varphi(s + u, \phi)$  for every  $s, u \geq 0$  and  $\phi \in \mathbb{R}_+^m$ , we can write  $(J_0 w)(\varphi(s, \phi))$  as

$$\begin{aligned} & \inf_{r \geq 0} \int_0^r e^{-(\lambda+\lambda_0)u} [g + \lambda_0(Kw)](\varphi(s+u, \phi)) du \\ & = e^{(\lambda+\lambda_0)s} \inf_{r \geq s} \int_s^r e^{-(\lambda+\lambda_0)u} [g + \lambda_0(Kw)](\varphi(u, \phi)) du, \end{aligned}$$

and  $(Jw)(\phi, s) + e^{-(\lambda+\lambda_0)s}(J_0 w)(\varphi(s, \phi)) = \inf_{r \geq s} \int_0^r e^{-(\lambda+\lambda_0)u} [g + \lambda_0(Kw)](\varphi(u, \phi)) du = (J_s w)(\phi)$ . If  $(J_0 w)(\varphi(s, \phi)) < 0$  for every  $0 \leq s < t$ , then  $(Jw)(\phi, s) > (Jw)(\phi, s) + e^{-(\lambda+\lambda_0)s}(J_0 w)(\varphi(s, \phi)) = (J_s w)(\phi) = \inf_{r \geq s} (Jw)(\phi, r)$  for every  $0 \leq s < t$ . Therefore,  $(J_s w)(\phi) = \inf_{r \geq s} (Jw)(\phi, r) = \inf_{r \geq t} (Jw)(\phi, r) = (J_t w)(\phi)$  for every  $0 \leq s \leq t$ .

*Proof of Proposition 3.4.* The result holds for  $n = 0$ . Suppose for some  $n \geq 0$  we have  $\mathbb{E}_\infty^\phi[M_{\tau \wedge \tau_\varepsilon \wedge T_n}] = \mathbb{E}_\infty^\phi[M_0]$  for every  $\phi \in \mathbb{R}_+^m$ ,  $\varepsilon \geq 0$ , and  $(\mathcal{F}_t)_{t \geq 0}$ -stopping time  $\tau$ . Fix  $\phi \in \mathbb{R}_+^m$ ,  $\varepsilon \geq 0$ , and  $(\mathcal{F}_t)_{t \geq 0}$ -stopping time  $\tau$ . Then  $\mathbb{E}_\infty^\phi[M_{\tau \wedge \tau_\varepsilon \wedge T_{n+1}}] = \mathbb{E}_\infty^\phi[M_{\tau \wedge \tau_\varepsilon \wedge T_n} + 1_{\{\tau \wedge \tau_\varepsilon \wedge T_n\}}(M_{\tau \wedge \tau_\varepsilon \wedge T_{n+1}} - M_{T_n})] = \mathbb{E}_\infty^\phi[M_0] + \mathbb{E}_\infty^\phi[1_{\{\tau \wedge \tau_\varepsilon \geq T_n\}}(M_{\tau \wedge \tau_\varepsilon \wedge T_{n+1}} - M_{T_n})]$  by induction hypothesis. We shall prove that second term equals zero. Since by Proposition 3.1 there is an  $(\mathcal{F}_t)_{t \geq 0}$ -measurable  $R_n$  such that  $(\tau \wedge \tau_\varepsilon \wedge T_{n+1})1_{\{\tau \wedge \tau_\varepsilon \geq T_n\}} = [(T_n + R_n) \wedge T_{n+1}]1_{\{\tau \wedge \tau_\varepsilon \geq T_n\}}$  we can write  $\mathbb{E}_\infty^\phi[1_{\{\tau \wedge \tau_\varepsilon \geq T_n\}}(M_{\tau \wedge \tau_\varepsilon \wedge T_{n+1}} - M_{T_n})]$  as

$$\begin{aligned} & \mathbb{E}_\infty^\phi \left[ 1_{\{\tau \wedge \tau_\varepsilon \geq T_n\}} \left( \int_{T_n}^{\tau \wedge \tau_\varepsilon \wedge T_{n+1}} e^{-\lambda t} g(\Phi_t) dt + 1_{\{\tau \wedge \tau_\varepsilon \geq T_{n+1}\}} e^{-\lambda T_{n+1}} V(\Phi_{T_{n+1}}) \right. \right. \\ & \quad \left. \left. + 1_{\{\tau \wedge \tau_\varepsilon < T_{n+1}\}} e^{-\lambda(\tau \wedge \tau_\varepsilon)} V(\Phi_{\tau \wedge \tau_\varepsilon}) - e^{-\lambda T_n} V(\Phi_{T_n}) \right) \right] \\ & = \mathbb{E}_\infty^\phi \left[ 1_{\{\tau \wedge \tau_\varepsilon \geq T_n\}} e^{-\lambda T_n} \left( \int_{T_n}^{(T_n + R_n) \wedge T_{n+1}} e^{-\lambda(t - T_n)} g(\varphi(t - T_n, \Phi_{T_n})) dt \right. \right. \\ & \quad \left. \left. + 1_{\{R_n \geq T_{n+1} - T_n\}} e^{-\lambda(T_{n+1} - T_n)} V(\varphi(T_{n+1} - T_n, \Phi_{T_n})) \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(Z_{n+1}) \right. \right. \\ & \quad \left. \left. + 1_{\{R_n < T_{n+1} - T_n\}} e^{-\lambda R_n} V(\varphi(R_n, \Phi_{T_n})) - V(\Phi_{T_n}) \right) \right]. \end{aligned}$$

Because the random variables  $1_{\{\tau \wedge \tau_\varepsilon \geq T_n\}}$ ,  $T_n$ ,  $R_n$ ,  $\Phi_{T_n}$  are  $\mathcal{F}_{T_n}$ -measurable, and since  $T_{n+1} - T_n$  and  $Z_{n+1}$  are independent of  $\mathcal{F}_{T_n}$  with exponential distribution with rate  $\lambda_0$  and with distribution  $v_0$  under  $\mathbb{P}_\infty$ , respectively, taking the conditional expectation with respect to  $\mathcal{F}_{T_n}$  inside the above expectation gives

$$\begin{aligned} & \mathbb{E}_\infty^\phi \left[ 1_{\{\tau \wedge \tau_\varepsilon \geq T_n\}} e^{-\lambda T_n} \left( \int_0^{R_n} e^{-(\lambda+\lambda_0)t} g(\varphi(t, \Phi_{T_n})) dt \right. \right. \\ & \quad \left. \left. + \int_0^{R_n} \lambda_0 e^{-(\lambda+\lambda_0)t} \underbrace{V\left(\varphi(t, \Phi_{T_n}) \frac{\lambda_1}{\lambda_0} \frac{dv_1}{dv_0}(z)\right)}_{(KV)(\varphi(t, \Phi_{T_n}))} v_0(dz) dt \right. \right. \\ & \quad \left. \left. + e^{-(\lambda+\lambda_0)R_n} V(\varphi(R_n, \Phi_{T_n})) - V(\Phi_{T_n}) \right) \right] \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E}^\phi \left[ 1_{\{\tau \wedge \tau_\varepsilon \geq T_n\}} e^{-\lambda T_n} \left( \int_0^{R_n} e^{-(\lambda+\lambda_0)t} [g + \lambda_0(KV)](\varphi(t, \Phi_{T_n})) dt \right. \right. \\
&\quad \left. \left. + e^{-(\lambda+\lambda_0)R_n} V(\varphi(R_n, \Phi_{T_n})) - V(\Phi_{T_n}) \right) \right] \\
&= \mathbb{E}^\phi \left[ 1_{\{\tau \wedge \tau_\varepsilon \geq T_n\}} e^{-\lambda T_n} \left( \underbrace{(JV)(\Phi_{T_n}, R_n) + e^{-(\lambda+\lambda_0)R_n} V(\varphi(R_n, \Phi_{T_n})) - V(\Phi_{T_n})}_{(J_{R_n} V)(\Phi_{T_n}) \text{ by Corollary 3.3}} \right) \right] \\
&= \mathbb{E}^\phi \left[ 1_{\{\tau \wedge \tau_\varepsilon \geq T_n\}} e^{-\lambda T_n} ((J_{R_n} V)(\Phi_{T_n}) - V(\Phi_{T_n})) \right].
\end{aligned}$$

Because  $(\tau \wedge \tau_\varepsilon) 1_{\{T_n \leq \tau \wedge \tau_\varepsilon < T_{n+1}\}} = (T_n + R_n) 1_{\{T_n \leq \tau \wedge \tau_\varepsilon < T_{n+1}\}}$  and  $\tau_\varepsilon 1_{\{T_n \leq \tau \wedge \tau_\varepsilon < T_{n+1}\}} = [T_n + r_\varepsilon(\Phi_{T_n})] 1_{\{T_n \leq \tau_\varepsilon < T_{n+1}\}}$ , where  $r_\varepsilon(\cdot)$  is defined as in Lemma 3.4, and since  $\tau \wedge \tau_\varepsilon \leq \tau_\varepsilon$ , we have  $R_n \leq r_\varepsilon(\Phi_{T_n})$  on  $\{T_n \leq \tau \wedge \tau_\varepsilon < T_{n+1}\}$ . However, since  $R_n$  and  $r_\varepsilon(\Phi_{T_n})$  are  $\mathcal{F}_{T_n}$ -measurable, we must also have  $R_n \leq r_\varepsilon(\Phi_{T_n})$  on  $\{\tau \wedge \tau_\varepsilon \geq T_n\}$ . Because  $V(\varphi(s, \Phi_{T_n})) < -\varepsilon \leq 0$  for every  $0 \leq s < r_\varepsilon(\Phi_{T_n})$ , we also have  $V(\varphi(s, \Phi_{T_n})) < -\varepsilon \leq 0$  for every  $0 \leq s < R_n$  on  $\{\tau \wedge \tau_\varepsilon \geq T_n\}$ . Then Corollary 3.3 guarantees that  $(J_{R_n} V)(\Phi_{T_n}) = V(\Phi_{T_n})$  on  $\{\tau \wedge \tau_\varepsilon \geq T_n\}$  and we finally have  $\mathbb{E}_\infty^\phi [1_{\{\tau \wedge \tau_\varepsilon \geq T_n\}} (M_{\tau \wedge \tau_\varepsilon \wedge T_{n+1}} - M_{T_n})] = \mathbb{E}_\infty^\phi [1_{\{\tau \wedge \tau_\varepsilon \geq T_n\}} e^{-\lambda T_n} ((J_{R_n} V)(\Phi_{T_n}) - V(\Phi_{T_n}))] = 0$ .

## ACKNOWLEDGMENTS

The author thanks the anonymous referee and the editor Alexander Tartakovsky for thoughtful remarks and valuable suggestions, which improved the exposition of the paper.

## REFERENCES

- Baron, M. and Tartakovsky, A. G. (2006). Asymptotic Optimality of Change-Point Detection Schemes in General Continuous-Time Models, *Sequential Analysis* 25: 257–296.
- Basseville, M. and Nikiforov, I. V. (1993). *Detection of Abrupt Changes: Theory and Application*, Prentice Hall Information and System Sciences Series, Englewood Cliffs, NJ: Prentice Hall Inc.
- Bayraktar, E. and Dayanik, S. (2006). Poisson Disorder Problem with Exponential Penalty for Delay, *Mathematics of Operations Research* 31: 217–233.
- Bayraktar, E., Dayanik, S., and Karatzas, I. (2005). The Standard Poisson Disorder Problem Revisited, *Stochastic Processes and Their Applications* 115: 1437–1450.
- Beibel, M. (2000). A Note on Sequential Detection with Exponential Penalty for the Delay, *Annals of Statistics* 28: 1696–1701.
- Davis, M. H. A. (1993). *Markov Models and Optimization*, vol. 49 of Monographs on Statistics and Applied Probability, London: Chapman & Hall.
- Dayanik, S., Poor, H. V., and Sezer, S. O. (2008). Multisource Bayesian Sequential Change Detection, *Annals of Applied Probability* 18: 552–590.
- Dayanik, S. and Sezer, S. O. (2006). Compound Poisson Disorder Problem, *Mathematics of Operations Research* 31: 649–672.
- Galchuk, L. I. and Rozovskii, B. L. (1971). The Problem of “Disorder” for a Poisson Process, *Teoriya Veroyatnostei i ee Primeneniya* 16: 729–734.



- Gapeev, P. V. (2005). The Disorder Problem for Compound Poisson Processes with Exponential Jumps, *Annals of Applied Probability* 15: 487–499.
- Gapeev, P. V. (2006). Nonadditive Disorder Problems for Some Diffusion Processes, *WIAS Preprint No. 1181*, London School of Economics.
- Gugerli, U. S. (1986). Optimal Stopping of a Piecewise-Deterministic Markov Process, *Stochastics* 19: 221–236.
- Hadjiliadis, O. (2005). Optimality of the 2-CUSUM Drift Equalizer Rules for Detecting Two-Sided Alternatives in the Brownian Motion Model, *Journal of Applied Probability* 42: 1183–1193.
- Hadjiliadis, O. and Moustakides, V. (2005). Optimal and Asymptotically Optimal CUSUM Rules for Change Point Detection in the Brownian Motion Model with Multiple Alternatives, *Teoriya Veroyatnostei ee Primeneniya* 50: 131–144.
- Liptser, R. S. and Shiryaev, A. N. (2001). *Statistics of Random Processes II*, vol. 6 of Applications of Mathematics (New York), expanded edition, Berlin: Springer-Verlag.
- Lorden, G. (1971). Procedures for Reacting to a Change in Distribution, *Annals of Mathematical Statistics* 42: 1897–1908.
- Moustakides, G. V. (1986). Optimal Stopping Times for Detecting Changes in Distributions, *Annals of Statistics* 14: 1379–1387.
- Moustakides, G. V. (2004). Optimality of the CUSUM Procedure in Continuous Time, *Annals of Statistics* 32: 302–315.
- Moustakides, G. V. (2008). Sequential Change Detection Revisited, *Annals of Statistics* 36: 787–807.
- Peskir, G. and Shiryaev, A. (2006). *Optimal Stopping and Free-Boundary Problems*, Lectures in Mathematics ETH Zurich, Basel: Birkhäuser Verlag.
- Peskir, G. and Shiryaev, A. N. (2002). Solving the Poisson Disorder Problem, in *Advances in Finance and Stochastics*, Berlin: Springer, pp. 295–312.
- Pollak, M. (1985). Optimal Detection of a Change in Distribution, *Annals of Statistics* 13: 206–227.
- Poor, H. V. (1998). Quickest Detection with Exponential Penalty for Delay, *Annals of Statistics* 26: 2179–2205.
- Poor, H. V. and Hadjiliadis, O. (2008). *Quickest Detection*, Cambridge University Press, New York.
- Shiryaev, A. N. (1963). On Optimum Methods in Quickest Detection Problems, *Theory of Probability and Its Applications* 8: 22–26.
- Shiryaev, A. N. (1964). On Markov Sufficient Statistics in Non-Additive Bayes Problems of Sequential Analysis, *Theory of Probability and Its Applications* 9: 604–618.
- Shiryaev, A. N. (1996). Minimax Optimality of the Method of Cumulative Sums (CUSUM) in the Continuous Time Case, *Uspekhi Matematicheskikh Nauk* 51: 173–174.
- Shiryaev, A. N. (2008). Generalized Bayesian Nonlinear Quickest Detection Problems: On Markov Family of Sufficient Statistics, in *Mathematical Control Theory and Finance*, Sarychev, A., Shiryaev, A., Guerra, M., Grossinho, M. d. R. (eds.), Berlin: Springer, pp. 377–386.
- Shiryaev, A. N. and Zryumov, P. Y. (2009). Volume Dedicated to the 60th Birthday of Yu. Kabanov, *On the Linear and Nonlinear Generalized Bayesian Disorder Problem (Discrete Time Case)*, Berlin: Springer.
- Tartakovsky, A. G. (2008). Multidecision Quickest Change-Point Detection: Previous Achievements and Open Problems, *Sequential Analysis* 27: 201–231.
- Tartakovsky, A. G. and Veeravalli, V. V. (2004). General Asymptotic Bayesian Theory of Quickest Change Detection, *Teoriya Veroyatnostei i ee Primeneniya* 49: 538–582.