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
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 which depends only on the observations of the point process and which 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 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, 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 [1], Hadjiliadis [14], Hadjiliadis and Moustakides [15], Moustakides

Address correspondence to S. Dayanik, Department of Operations Research and Financial Engineering, Princeton University, Sherrerd Hall, Princeton NJ 08544, USA; Fax: 609-258-3791; E-mail: sdayanik@princeton.edu

[19, 20], Shiryaev [28] in continuous time, and by Lorden [17], Moustakides [18], Pollak [23], Tartakovsky [30], Tartakovsky and Veeravalli [31] in discrete time. Shiryaev [26] 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 [11] formulated simple Poisson disorder problem and provided partial solution, which has been completed by Peskir and Shiryaev [22]. Gapeev [12] solved compound Poisson disorder problem with exponentially distributed jumps. The solution for the general case was provided by Dayanik and Sezer [10]. Bayesian sequential detection of a change in the local characteristics of a finite-activity Lévy process has been formulated and solved by Dayanik, Poor, and Sezer [9]. Basseville and Nikiforov [2], Peskir and Shiryaev [21], and Poor and Hadjiliadis [25] 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 [1] and Tartakovsky and Veeravalli [31] 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 [24] in discrete time, by Beibel [5] in detecting a change in the drift of a Brownian motion, by Bayraktar and Dayanik [3] and Bayraktar, Dayanik, and Karatzas [4] in simple Poisson disorder problem, and by Dayanik and Sezer [10] in compound Poisson disorder problem. Shiryaev [27, 29] 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.

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 [29] 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 which 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 which 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 [13] and Davis [8]. 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 the entire state space uniformly and exponentially fast, and the 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 λ and mark distribution ν on some measurable space $(E, \mathcal{E})$ changes from (λ_0, ν_0) to (λ_1, ν_1) at some unobservable disorder time Θ, which has zero-modified exponential distribution

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

for some known constants $\lambda_0 > 0$, $\lambda_1 > 0$, $\lambda > 0$, $0 \leq p < 1$, and known probability distributions ν_0 and ν_1 on $(E, \mathcal{E})$. We want to detect the disorder time Θ by means of a stopping time τ of the observation filtration

$$\mathcal{F}_t = \sigma\{(T_n, Z_n);\ n \geq 1 \text{ such that } T_n \leq t\}, \quad t \geq 0$$

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

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

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

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

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 λ_0 and mark distribution ν_0 on $(E, \mathcal{E})$, and Θ is an independent random variable with zero-modified exponential distribution. Suppose that λ_1 is a positive constant and ν_1 is a probability distribution on $(E, \mathcal{E})$ absolutely continuous with respect to ν_0, and either $\lambda_0 \neq \lambda_1$ or $\nu_0 \not\equiv \nu_1$. Let

$$\mathcal{G}_t := \sigma(\Theta) \vee \mathcal{F}_t, \qquad t \geq 0$$

be the filtration obtained by augmenting $(\mathcal{F}_t)_{t\geq 0}$ with the information about Θ and define the probability measure $\mathbb{P}$ locally on $(\Omega, \mathcal{G}_\infty)$ through the Radon-Nikodym derivatives

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

where

$$L_t := e^{-(\lambda_1-\lambda_0)t} \prod_{n:\,T_n \leq t}\left(\frac{\lambda_1}{\lambda_0}\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(Z_n)\right) = \exp\left\{-(\lambda_1-\lambda_0)t + \int_0^t \int_E \left(\log\frac{\lambda_1}{\lambda_0}\frac{\nu_1}{\nu_0}(z)\right) N(\mathrm{d}s, \mathrm{d}z)\right\},$$

and $N(\mathrm{d}s, \mathrm{d}z)$ is the Poisson random measure on $(\mathbb{R}_+ \times E, \mathscr{B}(\mathbb{R}_+) \times \mathcal{E})$ with mean measure $\lambda_0 \mathrm{d}s\, \nu_0(\mathrm{d}z)$ under $\mathbb{P}_\infty$. Girsanov's change-of-measure theorem guarantees that $(T_n, Z_n)_{n\geq 1}$ and Θ 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

$$L_0 = 1 \qquad \text{and} \qquad \mathrm{d}L_t = L_{t-} \int_E \left(\frac{\lambda_1}{\lambda_0} \frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(z) - 1 \right) \left[N(\mathrm{d}t, \mathrm{d}z) - \lambda_0 \mathrm{d}t\, \nu_0(\mathrm{d}z) \right], \qquad t \geq 0. \tag{2.1}$$

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

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

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

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

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

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

in terms of the first element of the sequence of processes

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

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 $\tau \in \mathcal{S}$ can be written as

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

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*

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

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. time.

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 not in general 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

$$\begin{aligned} \Phi_0^{(n)} &= \frac{p}{1-p} f^{(n)}(0), \qquad\qquad 1 \leq n \leq m, \\ d\Phi_t^{(n)} &= \left[\lambda(f^{(n)}(0) + \Phi_t^{(n)}) + \Phi_t^{(n+1)}\right] dt \\ &\quad + \Phi_{t-}^{(n)} \int_E \left(\frac{\lambda_1}{\lambda_0}\frac{d\nu_1}{d\nu_0}(z) - 1\right)\left[N(dt, dz) - \lambda_0 dt\, \nu_0(dz)\right], \quad t \geq 0,\ 1 \leq n \leq m, \end{aligned} \tag{2.2}$$

the derivation of which is included in the appendix. For every $1 \leq n \leq m$, the drift of $\Phi^{(n)}$ depends on $\Phi^{(n+1)}$, and the process $\{(\Phi_t^{(1)}, \ldots, \Phi_t^{(m)}), \mathcal{F}_t;\ t \geq 0\}$ is in general not a Markov process. If $f(\cdot)$ is infinitely continuously differentiable, then under suitable conditions $\{(\Phi_t^{(1)}, \Phi_t^{(2)}, \ldots), \mathcal{F}_t;\ t \geq 0\}$ 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 $\Phi^{(1)}, \ldots, \Phi^{(m)}$, in which case the m-dimensional process $\{(\Phi_t^{(1)}, \ldots, \Phi_t^{(m)}), \mathcal{F}_t;\ t \geq 0\}$ is a Markov sufficient statistic for the sequential change detection problem.

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

(i) Suppose that $f(t) = a_0 + a_1(t-b_1) + a_2(t-b_2)^2 + \ldots + a_m(t-b_m)^m$ for every $t \geq 0$ for some constants $a_0, b_0, \ldots, a_m, b_m$. 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 $\nu_1 \equiv \nu_0$) with linear detection delay penalty function $f(t) = t$ was formulated and partially solved by Galchuk, Rozovskii [11]. The complete solution was later given by Peskir and Shiryaev [22] by using method of variational inequalities. Later, Dayanik and Sezer [10] 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.

(ii) 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

$$\Phi_0^{(1)} = \frac{abp}{1-p}, \quad d\Phi_t^{(1)} = \left[\lambda(\Phi_t^{(1)} + ab) + b\Phi_t^{(1)}\right] dt \\ + \Phi_t^{(1)} \int_E \left(\frac{\lambda_1}{\lambda_0}\frac{d\nu_1}{d\nu_0}(z) - 1\right)\left[N(dt, dz) - \lambda_0 dt\, dz\right],\ t \geq 0$$

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. [4] 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 [10] with successive approximations applied to an equivalent essentially discrete-time optimal stopping problem.

(iii) 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

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

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 $\{(\Phi_t^{(1)}, \ldots, \Phi_t^{(m)}), \mathcal{F}_t;\ t \geq 0\}$ is a strong Markov sufficient statistic for the sequential change detection problem. The general solution of the homogeneous constant coefficient ordinary differential equation is in the form of

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

for suitable constants $\rho_n, a_n, \alpha_n, b_n, \beta_n, \gamma_n$ for $1 \leq n \leq m$.

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 $a > 0$,

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

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 that 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*

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

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

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

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

$$\begin{aligned} \mathrm{d}\Phi_t^{(n)} &= \left[\lambda \Phi_t^{(n)} + \Phi_t^{(n+1)}\right] \mathrm{d}t + \Phi_{t-}^{(n)} \int_E \left(\frac{\lambda_1}{\lambda_0} \frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(z) - 1\right) [N(\mathrm{d}t, \mathrm{d}z) - \lambda_0 \mathrm{d}t\, \nu_0(\mathrm{d}z)], \quad t \geq 0, \\ \Phi_0^{(n)} &= \frac{p}{1-p} f^{(n)}(0) = 0 \quad \textit{for every } n = 1, \ldots, m-1, \end{aligned}$$

and

$$\mathrm{d}\Phi_t^{(m)} = \lambda \left(\Phi_t^{(m)} + m!\right) \mathrm{d}t + \Phi_{t-}^{(m)} \int_E \left(\frac{\lambda_1}{\lambda_0} \frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(z) - 1\right) [N(\mathrm{d}t, \mathrm{d}z) - \lambda_0 \mathrm{d}t\, \nu_0(\mathrm{d}z)], \quad t \geq 0,$$

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

where for every $\phi \in \mathbb{R}_+^m$ the 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 Φ can be separated, and its dynamics can be written compactly as

$$\Phi_0 = \frac{p\, m!}{1-p} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}_{m\times 1}, \quad \mathrm{d}\Phi_t = (A\Phi_t + b)\mathrm{d}t + \Phi_{t-} \int_E \left(\frac{\lambda_1}{\lambda_0} \frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(z) - 1 \right) N(\mathrm{d}t, \mathrm{d}z), \qquad t \geq 0,$$

where

$$\Phi_t = \begin{bmatrix} \Phi_t^{(1)} \\ \vdots \\ \Phi_t^{(m)} \end{bmatrix}, \quad A = \begin{bmatrix} -\bar{\lambda} & 1 & 0 & 0 & \cdots & 0 \\ 0 & -\bar{\lambda} & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \cdots & \vdots \\ 0 & \cdots & & -\bar{\lambda} & 1 & 0 \\ 0 & \cdots & & 0 & -\bar{\lambda} & 1 \\ 0 & \cdots & & 0 & 0 & -\bar{\lambda} \end{bmatrix}_{m\times m}, \quad b = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \lambda\, m! \end{bmatrix}_{m\times 1}, \quad \bar{\lambda} = \lambda_1 - \lambda_0 - \lambda,$$

$$A = -\bar{\Lambda} + N, \qquad \bar{\Lambda} = \begin{bmatrix} \bar{\lambda} & & & \\ & \bar{\lambda} & & \\ & & \ddots & \\ & & & \bar{\lambda} \end{bmatrix}_{m\times m}, \qquad N = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \cdots & \vdots \\ 0 & \cdots & & 0 & 1 & 0 \\ 0 & \cdots & & 0 & 0 & 1 \\ 0 & \cdots & & 0 & 0 & 0 \end{bmatrix}_{m\times m}$$

The matrix A is the sum of diagonal matrix $-\bar{\Lambda}$, whose diagonal elements equal $-\bar{\lambda}$, and matrix N, which is nilpotent with index m. Therefore, for every $t \geq 0$

$$e^{Nt} = \sum_{k=0}^{\infty} N^k \frac{t^k}{k!} = \sum_{k=0}^{m-1} N^k \frac{t^k}{k!} = \begin{bmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{m-2}}{(m-2)!} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & & \ddots & t \\ 0 & \cdots & & \cdots & 1 \end{bmatrix} \quad \text{and} \quad e^{At} = e^{-\bar{\Lambda}t} e^{Nt} = e^{-\bar{\lambda}t} e^{Nt}. \tag{2.4}$$

Between jump times T_n and T_{n+1}, the process Φ follows the integral curves of the system of m linear ordinary differential equations $\mathrm{d}\Phi_t = (A\Phi_t + b)\mathrm{d}t$, $t \in [T_n, T_{n+1})$, which admits the solution (see Coddington [7, Theorem 3.4])

$$\begin{aligned} \Phi_t &= e^{A(t-T_n)} \Phi_{T_n} + \int_{T_n}^t e^{A(t-s)} b \mathrm{d}s = e^{A(t-T_n)} \Phi_{T_n} + \left(\int_0^{t-T_n} e^{As} \mathrm{d}s \right) b \\ &= \varphi(t - T_n, \Phi_{T_n}) \qquad \text{for every } t \in [T_n, T_{n+1}), \end{aligned}$$

where we define

$$\varphi(t, \phi) = e^{At}\phi + \left(\int_0^t e^{As} \mathrm{d}s \right) b \quad \text{for every } t \geq 0 \text{ and } \phi \in \mathbb{R}_+^m.$$

If $\bar\lambda \neq 0$, then A is invertible, and

$$A^{-1} = -\begin{bmatrix} \bar\lambda^{-1} & \bar\lambda^{-2} & \bar\lambda^{-3} & \cdots & \bar\lambda^{-m} \\ 0 & \bar\lambda^{-1} & \bar\lambda^{-2} & \cdots & \bar\lambda^{-m+1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \ddots & \ddots & \bar\lambda^{-2} \\ 0 & \cdots & \cdots & & \bar\lambda^{-1} \end{bmatrix}, \qquad \begin{array}{c} \displaystyle\int_0^t e^{As}\mathrm{d}s = (e^{At} - I)A^{-1}, \quad t \geq 0, \\ A^{-1}b = -\lambda m! \begin{bmatrix} \bar\lambda^{-m} \\ \vdots \\ \bar\lambda^{-3} \\ \bar\lambda^{-2} \\ \bar\lambda^{-1} \end{bmatrix}, \quad \text{and} \end{array} \tag{2.5}$$

$$\varphi(t,\phi) = e^{At}\phi + (e^{At} - I)A^{-1}b = e^{At}(\phi + A^{-1}b) - A^{-1}b, \quad \forall\, t \geq 0,\ \forall\, \phi \in \mathbb{R}_+^m.$$

If $\bar\lambda > 0$, then $\lim_{t\to\infty} e^{At} = \lim_{t\to\infty} e^{-\bar\lambda t}e^{Nt} = 0$ and $\lim_{t\to\infty}\varphi(t,\phi) = -A^{-1}b$ for every $\phi \in \mathbb{R}_+^m$. If $\bar\lambda < 0$, then for every $\phi \in \mathbb{R}_+^m$ we have $\phi + A^{-1}b \neq 0$, and because $\mathrm{d}\varphi(t,\phi)/\mathrm{d}t = A\varphi(t,\phi) + b > 0$, the nth component $\varphi_n(t,\phi)$ of m-dimensional $\varphi(t,\phi)$ is strictly increasing in $t \in \mathbb{R}_+$ and increases to $+\infty$ as $t \to \infty$ for every $n = 1,\ldots,m$. Finally, if $\bar\lambda = 0$, then $\bar\Lambda = 0$ and $A = -\bar\Lambda + N = N$ is not invertible, but we can still directly calculate that

$$\int_0^t e^{Ns}\mathrm{d}s = \begin{bmatrix} t & \frac{t^2}{2} & \frac{t^3}{3!} & \cdots & \frac{t^m}{m!} \\ 0 & t & \frac{t^2}{2} & \cdots & \frac{t^{m-1}}{(m-1)!} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & & \ddots & \frac{t^2}{2} \\ 0 & \cdots & & \cdots & t \end{bmatrix}, \quad \text{and} \quad \begin{array}{c} \varphi(t,\phi) = e^{Nt}\phi + \left(\displaystyle\int_0^t e^{Ns}\mathrm{d}s\right)b, \\ \forall\, t \geq 0,\ \forall\, \phi \in \mathbb{R}_+^m, \end{array} \tag{2.6}$$

and obviously $t \mapsto \varphi(t,\phi)$ is strictly increasing with $\lim_{t\to\infty}\varphi(t,\phi) = +\infty$ for every $n = 1,\ldots,m$ and $\phi \in \mathbb{R}_+^m$. Proposition 2.3 summarizes the sample-path properties of process Φ described so far.

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

$$\Phi_t = \begin{cases} \varphi(t - T_n, \Phi_{T_n}), & \text{if} \quad t \in [T_n, T_{n+1}) \quad \text{for some } n \geq 0, \\ \varphi(T_{n+1} - T_n, \Phi_{T_n})\dfrac{\lambda_1}{\lambda_0}\dfrac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(Z_{n+1}), & \text{if} \quad t = T_{n+1} \quad \text{for some } n \geq 0, \end{cases}$$

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

$$\varphi(t,\phi) = \left\{\begin{matrix} e^{At}(\phi + A^{-1}b) - A^{-1}b, & \text{if} \quad \bar\lambda \neq 0 \\ e^{Nt}\phi + \left(\displaystyle\int_0^t e^{Ns}\mathrm{d}s\right)b, & \text{if} \quad \bar\lambda = 0 \end{matrix}\right\} \qquad \text{for every } t \geq 0 \text{ and } \phi \in \mathbb{R}_+^m,$$

and e^{At}, $A^{-1}b$, e^{Nt}, and $\int_0^t e^{Ns}\mathrm{d}s$ can be calculated explicitly by (2.4), (2.5), (2.4), and (2.6), respectively. If $\bar\lambda > 0$, then $\lim_{t\to\infty}\varphi(t,\phi) = -A^{-1}b$ for every $\phi \in \mathbb{R}_+^m$. If $\bar\lambda \leq 0$, then $t \mapsto \varphi_n(t,\phi)$ is strictly increasing and $\lim_{t\to\infty}\varphi_n(t,\phi) = +\infty$ for every $1 \leq n \leq m$ and $\phi \in \mathbb{R}_+^m$.

3. A DYNAMIC PROGRAMMING OPERATOR AND SOLUTION

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

$$(Kw)(\phi) = \int_E w\left(\phi\frac{\lambda_1}{\lambda_0}\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(z)\right)\nu_0(\mathrm{d}z), \qquad \phi \in \mathbb{R}_+^m, \tag{3.1}$$

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

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

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 by Proposition 3.1, the proof of which is deferred to the appendix.

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

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

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

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

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 Φ_{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 Φ at $(\mathcal{F}_t)_{t\geq 0}$-stopping time $\tau\wedge T_1$ suggests that the expected value of this policy should equal

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

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 τ by Proposition 3.1. Because by Proposition 2.3

$$\Phi_t = \varphi(t,\Phi_0) \quad \text{for } t\in[0,T_1) \qquad\text{and}\qquad \Phi_{T_1} = \varphi(T_1,\Phi_0)\frac{\lambda_1}{\lambda_0}\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(Z_1),$$

and since $(T_n,Z_n)_{n\geq 1}$ is a compound Poisson process with arrival rate λ_0 and mark distribution ν_0 on $(E,\mathcal{E})$, 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

$$\begin{aligned}
\mathbb{E}_\infty&\left[\int_0^{R_0} 1_{\{T_1\geq t\}}e^{-\lambda t}g(\varphi(t,\Phi_0))\mathrm{d}t + 1_{\{R_0\geq T_1\}}e^{-\lambda T_1}V\left(\varphi(T_1,\Phi_0)\frac{\lambda_1}{\lambda_0}\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(Z_1)\right)\right]\\
&= \int_0^{R_0} e^{-(\lambda+\lambda_0)t}g(\varphi(t,\Phi_0))\mathrm{d}t + \int_0^{R_0}\lambda_0 e^{-(\lambda+\lambda_0)t}\int_E V\left(\varphi(t,\Phi_0)\frac{\lambda_1}{\lambda_0}\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(z)\right)\nu_0(\mathrm{d}z)\mathrm{d}t\\
&= \int_0^{R_0} e^{-(\lambda+\lambda_0)t}[g+\lambda_0(KV)](\varphi(t,\Phi_0))\mathrm{d}t \equiv (JV)(\Phi_0,R_0).
\end{aligned}$$

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

$$\inf_{\tau\in\mathcal{S}}\,\mathbb{E}_\infty\left[\int_0^{\tau\wedge T_1} e^{-\lambda t}g(\Phi_t)\mathrm{d}t + 1_{\{\tau\geq T_1\}}e^{-\lambda T_1}V(\Phi_{T_1})\right] = \inf_{r\geq 0}\,(JV)(\Phi_0,r)\equiv(J_0V)(\Phi_0).$$

Because $V(\Phi_0)$ is by definition the minimum expected total discounted cost, the optimality principle of dynamic programming suggests that $V(\Phi_0)=(J_0V)(\Phi_0)$ 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

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

Let us first introduce the finite-horizon problems

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

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 $n \to \infty$, uniformly in $\phi \in \mathbb{R}_+^m$.

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

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

Proof. Because $\tau \wedge T_n \in \mathcal{S}$ for every $\tau \in \mathcal{S}$, we have $V(\phi) \leq \mathbb{E}_\infty^\phi \left[\int_0^{\tau \wedge T_n} e^{-\lambda t} g(\Phi_t) \mathrm{d}t \right]$ for every $\tau \in \mathcal{S}$, and taking the infimum of both sides over $\tau \in \mathcal{S}$ gives the first inequality $0 \leq V_n(\phi) - V(\phi)$. 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 λ_0, we have

$$\begin{aligned}
&\mathbb{E}_\infty^\phi \left[\int_0^\tau e^{-\lambda t} g(\Phi_t) \mathrm{d}t \right] = \mathbb{E}_\infty^\phi \left[\int_0^{\tau \wedge T_n} e^{-\lambda t} g(\Phi_t) \mathrm{d}t + 1_{\{\tau \geq T_n\}} \int_{T_n}^\tau e^{-\lambda t} g(\Phi_t) \mathrm{d}t \right] \\
&\geq \mathbb{E}_\infty^\phi \left[\int_0^{\tau \wedge T_n} e^{-\lambda t} g(\Phi_t) \mathrm{d}t \right] - \mathbb{E}_\infty^\phi \left[\int_{T_n}^\infty \lambda e^{-\lambda t} \mathrm{d}t \right] \geq V_n(\phi) - \mathbb{E}_\infty^\phi \left[e^{-\lambda T_n} \right] = V_n(\phi) - \left(\frac{\lambda}{\lambda + \lambda_0} \right)^n,
\end{aligned}$$

which proves the second inequality and completes the proof of the lemma. □

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 straight-forward 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 $V_n(\phi) \geq v_n(\phi)$.*

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*

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

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

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

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 ε-optimal for the problem in (3.8): $\mathbb{E}_\infty^\phi \left[\int_0^{\tau_{n,\varepsilon}\wedge T_n} e^{-\lambda t} g(\Phi_t)\mathrm{d}t\right] \leq V_n(\phi) + \varepsilon$ 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 $V_n(\phi) = v_n(\phi)$ 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 ε-optimal for the problem in (3.8). □

Lemma 3.2 describes an explicit decomposition for the stopping times $\tau_{n,\varepsilon}$, which is consistent with the general characterization of stopping times of $(\mathcal{F}_t)_{t\geq 0}$ described by Proposition 3.1.

Lemma 3.2. *For every $n \geq 1$ and $\varepsilon > 0$, let $\tau_{n,\varepsilon}$ be the stopping time in Proposition 3.3. If $0 \leq k \leq n$, then*

$$\{\tau_{n,\varepsilon} \geq T_k\} = \big\{r_{n,\varepsilon/2}(\Phi_{T_0}) \geq T_1 - T_0, r_{n-1,\varepsilon/4}(\Phi_{T_1}) \geq T_2 - T_1, \ldots, \\ r_{n-k+1,\varepsilon/2^k}(\Phi_{T_{k-1}}) \geq T_k - T_{k-1}\big\} = \bigcap_{\ell=0}^{k-1} \left\{r_{n-\ell,\varepsilon/2^{\ell+1}}(\Phi_{T_\ell}) \geq T_{\ell+1} - T_\ell\right\}.$$

If $0 \leq k \leq n-1$, then

$$\{T_k \leq \tau_{n,\varepsilon} < T_{k+1}\} = \left(\bigcap_{\ell=0}^{k-1} \left\{r_{n-\ell,\varepsilon/2^{\ell+1}}(\Phi_{T_\ell}) \geq T_{\ell+1} - T_\ell\right\}\right) \bigcap \left\{r_{n-k,\varepsilon/2^{k+1}}(\Phi_{T_k}) < T_{k+1} - T_k\right\},$$

$$\tau_{n,\varepsilon} 1_{\{T_k \leq \tau_{n,\varepsilon} < T_{k+1}\}} = \left[T_k + r_{n-k,\varepsilon/2^{k+1}}(\Phi_{T_k})\right] 1_{\{T_k \leq \tau_{n,\varepsilon} < T_{k+1}\}},$$

$$(\tau_{n,\varepsilon} \wedge T_{k+1}) 1_{\{\tau_{n,\varepsilon} \geq T_k\}} = \left[(T_k + r_{n-k,\varepsilon/2^{k+1}}(\Phi_{T_k})) \wedge T_{k+1}\right] 1_{\{\tau_{n,\varepsilon} \geq T_k\}}.$$

Lemma 3.3 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 3.3. *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 3.3 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 3.3 and 3.1 imply that

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

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

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 ε-optimal for the problem in (2.3). For its proof, we will need the next few lemmas and their corollaries.

Lemma 3.4. *For every bounded $w : \mathbb{R}_+^m \mapsto \mathbb{R}$, we have*

$$(Jw)(\phi, s) + e^{-(\lambda+\lambda_0)s}(J_0 w)(\varphi(s,\phi)) = (J_s w)(\phi) \quad \textit{for every } \phi \in \mathbb{R}_+^m \textit{ and } s \geq 0.$$

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 3.4, then we have*

$$(JV)(\phi, s) + e^{-(\lambda+\lambda_0)s}V(\varphi(s,\phi)) = (J_s V)(\phi) \quad \textit{for every } s \geq 0 \textit{ and } \phi \in \mathbb{R}_+^m,$$

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

$$(JV)(\phi, s) + e^{-(\lambda+\lambda_0)s}V(\varphi(s,\phi)) = V(\phi) \quad \textit{for every } 0 \leq s \leq t.$$

Lemma 3.5 gives the explicit decomposition of stopping times τ_ε, $\varepsilon \geq 0$, announced earlier by Proposition 3.1 for all stopping times of $(\mathcal{F}_t)_{t\geq 0}$.

Lemma 3.5. *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*

$$\tau_\varepsilon = \begin{cases} r_\varepsilon(\Phi_0), & \textit{if} \quad r_\varepsilon(\Phi_0) < T_1 \\ T_1 + \tau_\varepsilon \circ \theta_{T_1}, & \textit{if} \quad r_\varepsilon(\Phi_0) \geq T_1 \end{cases} \quad \textit{and} \quad \begin{array}{l} \tau_\varepsilon 1_{\{T_n \leq \tau_\varepsilon < T_{n+1}\}} = [T_n + r_\varepsilon(\Phi_{T_n})] 1_{\{T_n \leq \tau_\varepsilon < T_{n+1}\}} \\ \textit{for every } \varepsilon \geq 0 \textit{ and } n \geq 0, \end{array}$$

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 and Corollary 3.4 state that postponing the stopping decision until time τ_ε does not cause any regrets, and this observation almost immediately leads to the ε-optimality of τ_ε for the problem in (2.3), which is established by Theorem 3.1.

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

Corollary 3.4. *The stopped process $\{M_{t\wedge\tau_\varepsilon\wedge T_n}, \mathcal{F}_t;\ t \geq 0\}$ is a uniformly integrable martingale under $\mathbb{P}_\infty$ for every $n \geq 0$.*

Theorem 3.1. *For every $\varepsilon \geq 0$, the $(\mathcal{F}_t)_{t\geq 0}$-stopping time τ_ε of Lemma 3.5 is ε-optimal for problem in (2.3). Particularly, τ_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

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

On the one hand, on the event $\{\tau_\varepsilon < \infty\} \supseteq \{\tau_\varepsilon < T_n\}$, we have $V(\Phi_{\tau_\varepsilon}) > -\varepsilon$, because $V(\cdot)$ is continuous by Corollary 3.2, and $t \mapsto V(\Phi_t)$ is right-continuous. On the other hand, we always have $V(\cdot) \geq -1$ by Corollary 3.2 as well. Therefore,

$$\begin{aligned} V(\phi) &\geq \mathbb{E}^\phi_\infty \left[\int_0^{\tau_\varepsilon \wedge T_n} e^{-\lambda t} g(\Phi_t) \mathrm{d}t - \varepsilon 1_{\{\tau_\varepsilon < T_n\}} e^{-\lambda \tau_\varepsilon} - 1_{\{\tau_\varepsilon \geq T_n\}} e^{-\lambda T_n} \right] \\ &\geq \mathbb{E}^\phi_\infty \left[\int_0^{\tau_\varepsilon \wedge T_n} e^{-\lambda t} g(\Phi_t) \mathrm{d}t \right] - \varepsilon - \mathbb{E}^\phi_\infty \left[e^{-\lambda T_n} \right] \end{aligned}$$

The sequence of random variables $\int_0^{\tau_\varepsilon \wedge T_n} e^{-\lambda t} g(\Phi_t) \mathrm{d}t$, $n \geq 0$ is bounded from below, since

$$\int_0^{\tau_\varepsilon \wedge T_n} e^{-\lambda t} g(\Phi_t) \mathrm{d}t \geq -\int_0^{\tau_\varepsilon \wedge T_n} \lambda e^{-\lambda t} \mathrm{d}t \geq -\int_0^{\infty} \lambda e^{-\lambda t} \mathrm{d}t = -1,$$

and $\lim_{n\to\infty} \mathbb{E}_\infty e^{-\lambda T_n} = 0$ by the bounded convergence theorem, because $T_n \uparrow +\infty$ as $n \to \infty$. Then taking the limit infimum and using the Fatou's lemma give

$$V(\phi) \geq \varliminf_{k\to\infty} \mathbb{E}^\phi_\infty \left[\int_0^{\tau_\varepsilon \wedge T_n} e^{-\lambda t} g(\Phi_t) \mathrm{d}t \right] - \varepsilon \geq \mathbb{E}^\phi_\infty \left[\int_0^{\tau_\varepsilon} e^{-\lambda t} g(\Phi_t) \mathrm{d}t \right] - \varepsilon,$$

which proves that τ_ε is ε-optimal for problem in (2.3). □

Because $V(\cdot) = \lim_{n\to\infty} V_n(\cdot)$ can be calculated only in the limit, optimal stopping rule τ_0 may not be implementable. In practice, $V(\cdot)$ will be approximated by $V_n(\cdot) \equiv v_n(\cdot)$ for some sufficiently large $n \geq 0$, and the optimal performance of τ_0 can be approximated arbitrarily closely by the 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*

$$N(\varepsilon) := \min\left\{ n \geq 0;\ \left(\frac{\lambda}{\lambda+\lambda_0}\right)^n \leq \varepsilon \right\} = \left\lfloor \frac{\log \varepsilon}{\log \frac{\lambda}{\lambda+\lambda_0}} \right\rfloor \qquad \text{for every } \varepsilon > 0,$$

then $(\mathcal{F}_t)_{t\geq 0}$-stopping times $\sigma_{N(\varepsilon/2),\varepsilon/2}$ and $\sigma_{N(\varepsilon),0}$ are ε-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

$$\mathbb{E}^\phi_\infty[M_{\sigma_{n,\varepsilon} \wedge T_k}] = \mathbb{E}^\phi_\infty\left[M_{\sigma_{n,\varepsilon} \wedge \tau_\varepsilon \wedge T_k}\right] = \mathbb{E}^\phi_\infty[M_0] = V(\phi) \quad k \geq 0.$$

Then as in the proof of Theorem 3.1, we can write

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

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

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

On $\{\sigma_{n,\varepsilon} < \infty\} \supseteq \{\sigma_{n,\varepsilon} < T_k\}$ we have $V_n(\Phi_{\sigma_{n,\varepsilon}}) \geq -\varepsilon$ because $V_n(\cdot)$ is continuous by Corollary 3.2, and $t \mapsto V_n(\Phi_t)$ is right-continuous. Therefore,

$$\begin{aligned} V(\phi) &\geq \mathbb{E}^\phi_\infty \left[\int_0^{\sigma_{n,\varepsilon} \wedge T_k} e^{-\lambda t} g(\Phi_t) \mathrm{d}t - 1_{\{\sigma_{n,\varepsilon} < T_k\}} e^{-\lambda \sigma_{n,\varepsilon}} \left(\varepsilon + \left(\frac{\lambda}{\lambda + \lambda_0} \right)^n \right) - 1_{\{\sigma_{n,\varepsilon} \geq T_k\}} e^{-\lambda T_k} \right] \\ &\geq \mathbb{E}^\phi_\infty \left[\int_0^{\sigma_{n,\varepsilon} \wedge T_k} e^{-\lambda t} g(\Phi_t) \mathrm{d}t \right] - \left(\varepsilon + \left(\frac{\lambda}{\lambda + \lambda_0} \right)^n \right) - \mathbb{E}^\phi_\infty \left[e^{-\lambda T_k} \right]. \end{aligned}$$

For every fixed $n \geq 0$ and $\varepsilon \geq 0$, the sequence $\int_0^{\sigma_{n,\varepsilon} \wedge T_k} e^{-\lambda t} g(\Phi_t) \mathrm{d}t$, $k \geq 0$ is bounded from below since

$$\int_0^{\sigma_{n,\varepsilon} \wedge T_k} e^{-\lambda t} g(\Phi_t) \mathrm{d}t \geq - \int_0^{\sigma_{n,\varepsilon} \wedge T_k} \lambda e^{-\lambda t} \mathrm{d}t \geq - \int_0^\infty \lambda e^{-\lambda t} \mathrm{d}t = -1,$$

and $\lim_{k\to\infty} \mathbb{E}_\infty e^{-\lambda T_k} = 0$ by the bounded convergence theorem, because $T_k \uparrow +\infty$ as $k \to \infty$. Then taking the limit infimum and using the Fatou's lemma give

$$\begin{aligned} V(\phi) &\geq \varliminf_{k\to\infty} \mathbb{E}^\phi_\infty \left[\int_0^{\sigma_{n,\varepsilon} \wedge T_k} e^{-\lambda t} g(\Phi_t) \mathrm{d}t \right] - \varepsilon - \left(\frac{\lambda}{\lambda + \lambda_0} \right)^n \\ &\geq \mathbb{E}^\phi_\infty \left[\int_0^{\sigma_{n,\varepsilon}} e^{-\lambda t} g(\Phi_t) \mathrm{d}t \right] - \varepsilon - \left(\frac{\lambda}{\lambda + \lambda_0} \right)^n \qquad \text{for every } n \geq 0 \text{ and } \varepsilon \geq 0. \end{aligned}$$

The conclusion now immediately follows from the definition of $N(\varepsilon)$. □

We have seen that $V(\cdot)$ is the limit of $V_n(\cdot)$ as $n \to \infty$. The final result of this section shows that optimal stopping time τ_0 is similarly $\mathbb{P}_\infty$-a.s. limit of the sequence of increasingly accurate stopping rules $(\sigma_{n,0})_{n\geq 0}$. This result is established by first observing that τ_0 and $\sigma_{n,0}$, $n \geq 0$ are the first hitting times of the process Φ to the nested stopping regions Γ and Γ_n, $n \geq 0$, respectively, which are defined as the subsets of the state space $\mathbb{R}^m_+$ where $V(\cdot)$ and $V_n(\cdot)$, $n \geq 0$ vanish.

Theorem 3.3. *Let us define stopping regions*

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

Then the sets $\mathbb{R}^m_+ = \Gamma_0 \supseteq \Gamma_1 \supseteq \ldots \supseteq \Gamma_n \supseteq \ldots \supseteq \Gamma$ *are closed and convex, and* $\cap_{k\geq 0}\Gamma_k = \Gamma$. *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* τ_0 *as* $n \to \infty$.

Proof. Since $V_n(\cdot)$ decreases to $V(\cdot)$ as $n \to \infty$, we have $\Gamma_0 \supseteq \Gamma_1 \supseteq \ldots \supseteq \Gamma_n \supseteq \ldots \supseteq \Gamma$, 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 τ_0 and $\sigma_{n,0}$, $n \geq 0$ of Theorems 3.1 and 3.2 are by definition the first hitting times of process Φ to stopping sets Γ and Γ_n, $n \geq 0$. Because the sets Γ_n, $n \geq 0$ are decreasing, the hitting times $\sigma_{n,0}$, $n \geq 0$ are increasing with $\sigma_0 := \lim_{n\to\infty} \sigma_{n,0} \leq \tau_0$.

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

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 Γ_n, $n \geq 0$ are closed and $t \mapsto \Phi_t$ is right-continuous. Therefore, $\Phi_{\sigma_0} = \lim_{n\to\infty} \Phi_{\sigma_{n,0}} \in \Gamma_k$ for every $k \geq 0$, equivalently $\Phi_{\sigma_0} \in \cap_{k\geq 0}\Gamma_k$ 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 $\cap_{k\geq 0}\Gamma_k = \Gamma$.

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

Finally, Proposition 2.1 or the first part of Proposition 2.3 guarantee that τ_0 is a Bayes-optimal alarm time, and for every $\varepsilon > 0$, stopping time $\sigma_{N(\varepsilon),0}$ is an ε-optimal alarm time for the original compound Poisson disorder problem.

4. ILLUSTRATION

Consider the compound Poisson disorder problem with $\lambda_0 \neq \lambda_1$, $\nu_0 \equiv \nu_1$, 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

$$\Phi_t = \begin{cases} \varphi(t - T_n, \Phi_{T_n}), & \text{if} \quad t \in [T_n, T_{n+1}) \\ \varphi(T_{n+1} - T_n, \Phi_{T_n})\dfrac{\lambda_1}{\lambda_0}, & \text{if} \quad t = T_{n+1} \end{cases} \quad \text{and} \quad \begin{array}{l} \varphi(t,\phi) = e^{At}\phi + \left(\int_0^t e^{A(t-s)} \mathrm{d}s\right) b \\ \text{for every } t \geq 0,\ \phi \in \mathbb{R}_+^2, \text{ and } n \geq 0, \end{array}$$

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

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

$$\varphi(t,\phi) = \begin{cases} \begin{bmatrix} e^{-\bar{\lambda}t}(\phi_1 - 2\lambda\bar{\lambda}^{-2} + t(\phi_2 - 2\lambda\bar{\lambda}^{-1})) + 2\lambda\bar{\lambda}^{-2} \\ e^{-\bar{\lambda}t}(\phi_2 - 2\lambda\bar{\lambda}^{-1}) + 2\lambda\bar{\lambda}^{-1} \end{bmatrix}, & \text{if} \quad \bar{\lambda} \neq 0 \\ \begin{bmatrix} \phi_1 + \phi_2 t + \lambda t^2 \\ \phi_2 + 2\lambda t \end{bmatrix}, & \text{if} \quad \bar{\lambda} = 0 \end{cases} \quad \begin{array}{l} \text{for every } t \geq 0 \\ \text{and } \phi = (\phi_1, \phi_2) \in \mathbb{R}_+^2. \end{array}$$

Depending on the relationships between the parameters of the problem, the sample paths of process Φ 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 Φ 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 Φ from C_0.

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

The solution $x(t) = \varphi(t,\phi)$ of the system of linear ordinary differential equations $\mathrm{d}x(t)/\mathrm{d}t = Ax(t) + b$ with initial condition $x(0) = \phi \in \mathbb{R}_+^2$ has unique equilibrium point at $-A^{-1}b = [2\lambda\bar{\lambda}^{-2}\ 2\lambda\bar{\lambda}^{-1}]^{\mathrm{T}}$. Because $\bar{\lambda} > 0$, we also have $\lambda_1 - \lambda_0 > \lambda > 0$ and $\lambda_1/\lambda_0 > 1$, in which case at each mark arrival time, the process Φ jumps away from the origin along the ray emanating from the origin and passing through the position of Φ before jump; see Figures 1 and 2. The structure of the optimal solution depends on the position of the root λ 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 $-A^{-1}b$.

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 λ 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 also 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 λ, starting initially at $(\phi_1^*, 0)$ on the ϕ_1-axis; see Figure 1. For every fixed $\phi_1 \geq 0$, taking the derivative of $\varphi_1(t, (\phi_1, 0)) = e^{-\bar{\lambda}t}(\phi_1 - 2\lambda\bar{\lambda}^{-2} - 2\lambda\bar{\lambda}^{-1}t) + 2\lambda\bar{\lambda}^{-2}$ with

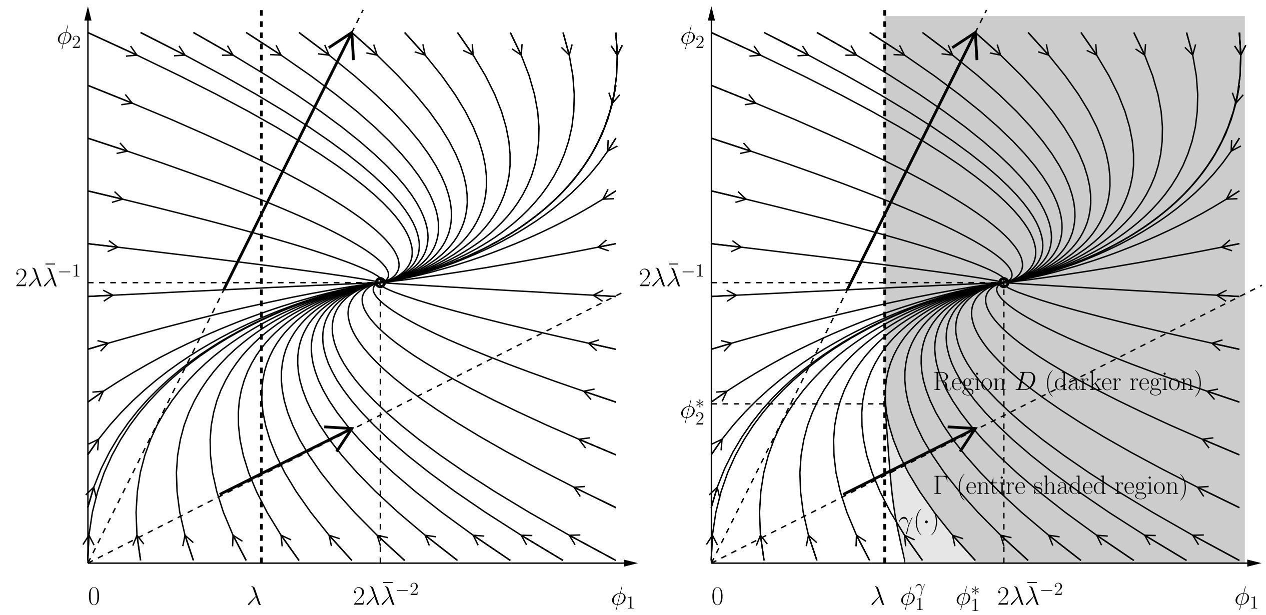


Figure 1. **Case I (a): $\bar{\lambda} > 0$ and $\lambda < 2\lambda\bar{\lambda}^{-2}$ (equivalently, $0 < \bar{\lambda} < \sqrt{2}$).** The sufficient statistic Φ follows the integral curves of a system of two linear ordinary differential equations, which have unique equilibrium point. Moreover, since $\lambda_1 > \lambda_0$, at every arrival time of a mark, Φ jumps away from the origin. On the lefthand side, integral curves and direction of jumps are drawn. On the righthand side, region D and optimal stopping region Γ are displayed.

respect to t gives $\partial\varphi_1(t,(\phi_1,0))/\partial t = e^{-\bar{\lambda}t}(-\bar{\lambda}\phi_1 + \lambda t)$, 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, $\lambda = \varphi_1(t^*(\phi_1^*),(\phi_1^*,0)) = 2\lambda\bar{\lambda}^{-2}\left[1 - e^{-(\bar{\lambda}^2/\lambda)\phi_1^*}\right]$ gives

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

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

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

which is the dark shaded region in Figure 1. 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

$$\mathbb{P}_\infty^\phi\{\Phi_t \in D \text{ for every } t \geq 0\} = 1 \qquad \text{for every } \phi \in D.$$

Since $g(\phi) - \lambda \geq 0$ for every $\phi \in D \subset \mathbb{R}_+^2 \setminus C_0$, we have $\mathbb{E}_\infty^\phi\left[\int_0^\tau e^{-\lambda t} g(\Phi_t)\mathrm{d}t\right] \geq 0$ 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 $\Gamma = \{\phi \in R_+^2;\ V(\phi) = 0\}$. Because $D \subset \Gamma \subseteq \mathbb{R}_+^2 \setminus C_0$ and Γ 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. All of those conclusions are direct consequences of the convexity of the optimal stopping region region Γ.

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 Φ never returns to C_0 once it leaves that region. Therefore, the first exit time τ_0 of Φ 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 the minimum Bayes risk

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

depends on $V(\phi)$ evaluated on $\{\phi = (0, \phi_2);\ \phi_2 \geq 0\}$, the $(\mathcal{F}_t)_{t \geq 0}$ stopping time τ_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 Φ.

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

Proof. Clearly, $\Phi_t^{(i)} = \varphi_i(t, \Phi_0)$ 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 $n \to \infty$.

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

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

Since $\Phi_{T_n-}^{(i)} \geq \varphi_i(T_n, \Phi_0)$ for every $i = 1, 2$ by the induction hypothesis, we can now write

$$\Phi_t^{(i)} \geq \varphi_i\left(t - T_n, \Phi_{T_n-}\right) \geq \varphi_i\left(t - T_n, \varphi_i(T_n, \Phi_0)\right) = \varphi_i\left(t - T_n + T_n, \Phi_0\right) = \phi_i(t, \Phi)$$

for every $T_n \leq t < T_{n+1}$ and $i = 1, 2$, which completes the proof of the induction step. □

Lemma 4.1 implies that

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

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 $h(\phi) = \min\{0, \int_0^\infty e^{-\lambda t}(\varphi_1(t, \phi) - \lambda)\mathrm{d}t\}$, and $h(\phi) = 0$ if

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

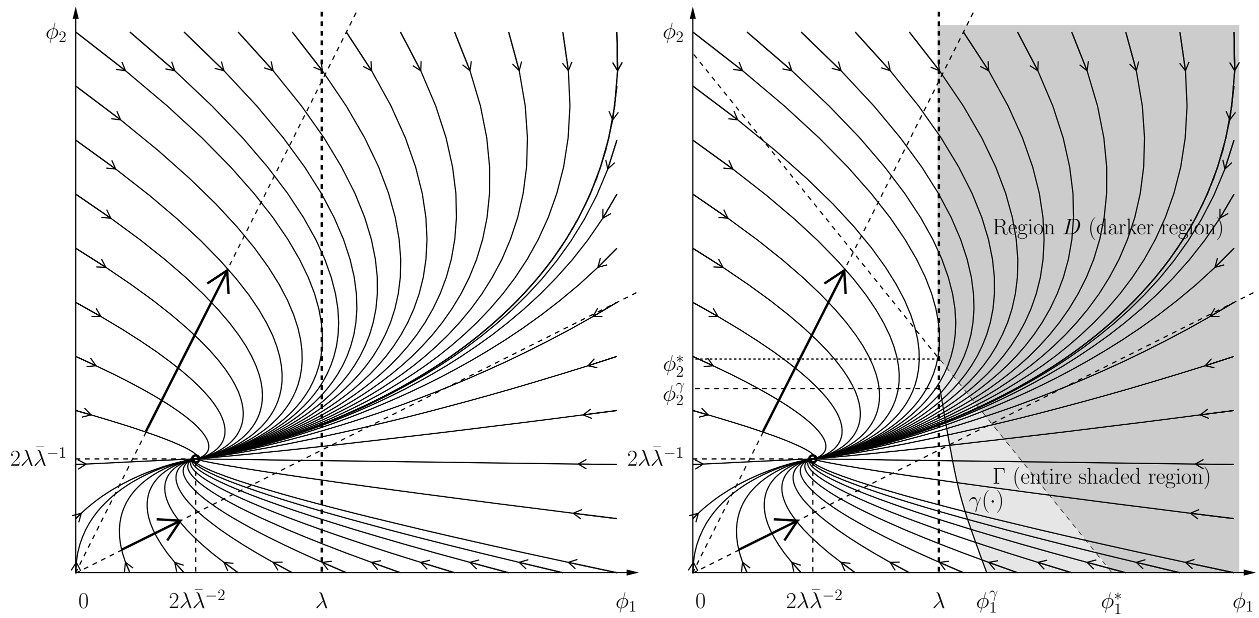


Figure 2. **Case I (b):** $\bar{\lambda} > 0$ **and** $\lambda \geq 2\lambda\bar{\lambda}^{-2}$ **(equivalently,** $\bar{\lambda} \geq \sqrt{2}$**).** As in Case I (a), the sufficient statistic Φ follows the integral curves of a system of two linear ordinary differential equations, which have unique equilibrium point, and since $\lambda_1 > \lambda_0$, at every arrival time of a mark, Φ jumps away from the origin. On the lefthand side, integral curves and direction of jumps are drawn. On the righthand side, region D and optimal stopping region Γ are displayed.

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

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

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

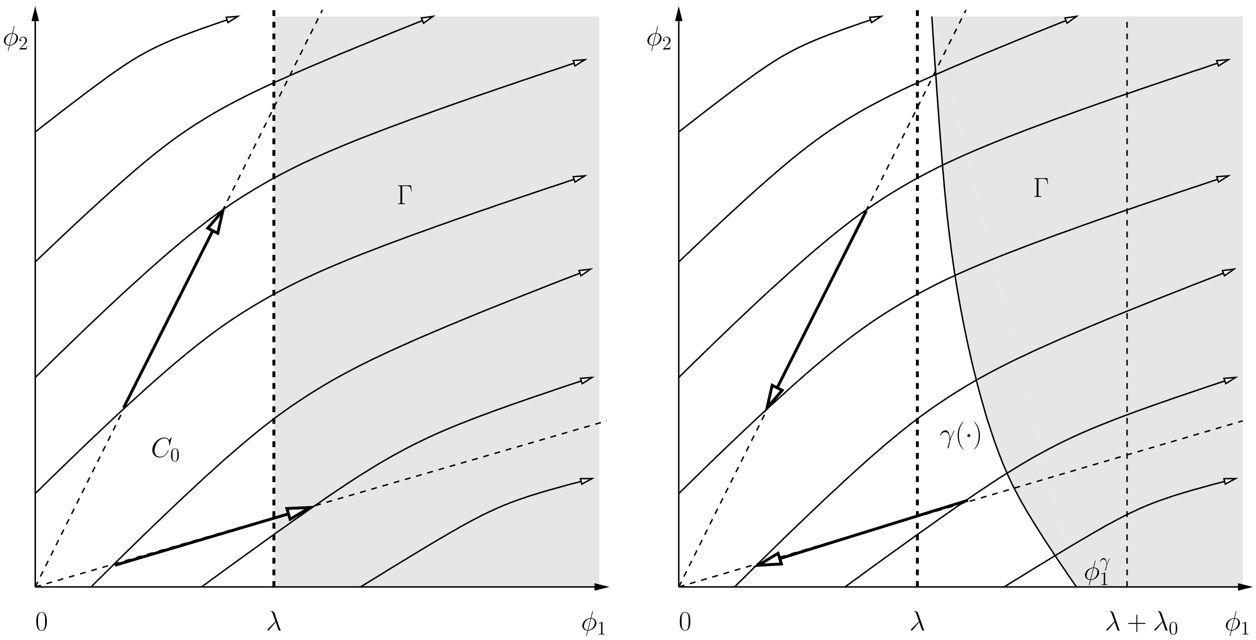
$$\phi_1^* = \lambda + \frac{\phi_2^*}{\lambda_1 - \lambda_0} > \lambda \qquad \textit{and} \qquad \phi_2^* = 2\lambda\bar{\lambda}^{-1} + (1 - 2\bar{\lambda}^{-2})(\lambda_1 - \lambda_0)\bar{\lambda} > 2\lambda\bar{\lambda}^{-1}.$$

Proof. Direct calculation gives

$$\begin{aligned}
\phi_2^* = \bar{\gamma}(\lambda) &= -\lambda(\lambda_1 - \lambda_0) + 2\lambda\bar{\lambda}^{-2}(\lambda_1 - \lambda_0) + 2\lambda\bar{\lambda}^{-1} + (1 - 2\bar{\lambda}^{-2})(\lambda_1 - \lambda_0)^2 \\
&= 2\lambda\bar{\lambda}^{-1} + (1 - 2\bar{\lambda}^{-2})(\lambda_1 - \lambda_0)^2 - (1 - 2\bar{\lambda}^{-2})\lambda(\lambda_1 - \lambda_0) \\
&= 2\lambda\bar{\lambda}^{-1} + (1 - 2\bar{\lambda}^{-2})(\lambda_1 - \lambda_0)(\lambda_1 - \lambda_0 - \lambda) = 2\lambda\bar{\lambda}^{-1} + (1 - 2\bar{\lambda}^{-2})(\lambda_1 - \lambda_0)\bar{\lambda} > 2\lambda\bar{\lambda}^{-1},
\end{aligned}$$

because $\bar{\lambda} \geq \sqrt{2}$ implies that $1 - 2\bar{\lambda}^{-2} > 0$, $\lambda_1 - \lambda_0 > \lambda > 0$, and $\bar{\lambda} > 0$. Because $\phi_1 \to \bar{\gamma}(\phi_1)$ is a straight line slope $-(\lambda_1 - \lambda_0)$, we have $(0 - \phi_2^*)/(\phi_1^* - \lambda) = -(\lambda_1 - \lambda_0)$, which completes the proof. □

Because optimal stopping region Γ 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 2.



(a) $\bar{\lambda} \leq 0$, $\lambda_1 > \lambda_0$ (b) $\bar{\lambda} \leq 0$, $\lambda_1 < \lambda_0$

***Figure 3*. Case II:** $\bar{\lambda} \leq 0$**.** The process Φ follows integral curves, both coordinates of which are strictly increasing. It jumps away from the origin in **Case II (a):** $\lambda_1 > \lambda_0$ and toward the origin in **Case II (b):** $\lambda_1 < \lambda_0$.

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

Both components of $t \mapsto \varphi(t, \phi)$ are strictly increasing for every $\phi \in \mathbb{R}_+^2$. 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 Φ 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 Γ coincides with $\mathbb{R}_+ \setminus C_0$, optimal stopping boundary $\partial\Gamma$ is the straight line $\phi_1 = \lambda$, and the first exit time τ_0 of process Φ 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 3(a).

Case II (b): $\bar{\lambda} < 0$ **and** $\lambda_1 < \lambda_0$**.** The process Φ is driven away from the origin between jump times, but is pulled back toward the origin at every jump. Therefore, Φ may return to region C_0 after a jump with positive probability; see Figure 3(b). Since $V(\cdot) \geq -1$, we have

$$\begin{aligned} V(\phi) = (J_0 V)(\phi) &= \inf_{r \geq 0} \int_0^r e^{-(\lambda+\lambda_0)t} [g + \lambda(KV)](\varphi(t, \phi)) \mathrm{d}t \\ &\geq \inf_{r \geq 0} \int_0^r e^{-(\lambda+\lambda_0)t} [\varphi_1(t, \phi) - \lambda - \lambda_0] \mathrm{d}t = 0 \qquad \text{for every } \phi \in [\lambda + \lambda_0, \infty) \times \mathbb{R}_+, \end{aligned}$$

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 $[\phi_1^\gamma, \lambda + \lambda_0]$.

A. APPENDIX: PROOFS OF SELECTED RESULTS

The Derivation of the Dynamics of Process Φ in (2.2). Because $Z_t 1_{\{\Theta \le t\}} = (L_t/L_\Theta) 1_{\{\Theta \le t\}}$, $Z_t 1_{\{\Theta > t\}} = 1_{\{\Theta > t\}}$ for every $t \ge 0$, and Θ and $\mathcal{F}_t$ are independent under $\mathbb{P}_\infty$, we have

$$\begin{aligned}
\Phi_t^{(n)} &= \frac{\mathbb{E}_\infty\left[Z_t f^{(n)}(t-\Theta) 1_{\{\Theta\le t\}} \mid \mathcal{F}_t\right]}{\mathbb{E}_\infty\left[Z_t 1_{\{\Theta>t\}} \mid \mathcal{F}_t\right]} = \frac{\mathbb{E}_\infty\left[(L_t/L_\Theta) f^{(n)}(t-\Theta) 1_{\{\Theta\le t\}} \mid \mathcal{F}_t\right]}{\mathbb{P}_\infty\{\Theta > t\}} \\
&= \frac{pL_t f^{(n)}(t) + (1-p)\int_0^t \lambda e^{-\lambda s}(L_t/L_s) f^{(n)}(t-s)\mathrm{d}s}{(1-p)e^{-\lambda t}} \\
&= L_t \left[\frac{p}{1-p} e^{\lambda t} f^{(n)}(t) + e^{\lambda t}\int_0^t \lambda e^{-\lambda s}\frac{1}{L_s} f^{(n)}(t-s)\mathrm{d}s\right] \quad \text{for every } n \ge 1.
\end{aligned}$$

Applying change-of-variable formula and using the dynamics of process L in (2.1) give

$$\begin{aligned}
\mathrm{d}\Phi_t^{(n)} = \Phi_{t-}^{(n)} &\int_E \left(\frac{\lambda_1}{\lambda_0}\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(z) - 1\right)[N(\mathrm{d}s,\mathrm{d}z) - \lambda_0 \mathrm{d}t\,\nu_0(\mathrm{d}z)] \\
&+ L_t\left[\lambda \frac{p}{1-p} e^{\lambda t} f^{(n)}(t) + \lambda e^{\lambda t}\int_0^t \lambda e^{-\lambda s}\frac{1}{L_s} f^{(n)}(t-s)\mathrm{d}s + \frac{p}{1-p} e^{\lambda t} f^{(n+1)}(t)\right. \\
&\left. + e^{\lambda t}\lambda e^{-\lambda t}\frac{1}{L_t} f^{(n)}(0) + e^{\lambda t}\int_0^t \lambda e^{-\lambda s}\frac{1}{L_s} f^{(n+1)}(t-s)\mathrm{d}s\right]\mathrm{d}t \\
= \Phi_{t-}^{(n)} &\int_E \left(\frac{\lambda_1}{\lambda_0}\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(z) - 1\right)[N(\mathrm{d}s,\mathrm{d}z) - \lambda_0 \mathrm{d}t\,\nu_0(\mathrm{d}z)] \\
&+ \left[\lambda f^{(n)}(0) + \lambda L_t\left(\frac{p}{1-p} e^{\lambda t} f^{(n)}(t) + \lambda e^{\lambda t}\int_0^t \lambda e^{-\lambda s}\frac{1}{L_s} f^{(n)}(t-s)\mathrm{d}s\right)\right. \\
&\left. + L_t\left(\frac{p}{1-p} e^{\lambda t} f^{(n+1)}(t) + e^{\lambda t}\int_0^t \lambda e^{-\lambda s}\frac{1}{L_s} f^{(n+1)}(t-s)\mathrm{d}s\right)\right]\mathrm{d}t \\
= \Phi_{t-}^{(n)} &\int_E \left(\frac{\lambda_1}{\lambda_0}\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(z) - 1\right)[N(\mathrm{d}s,\mathrm{d}z) - \lambda_0 \mathrm{d}t\,\nu_0(\mathrm{d}z)] + \left[\lambda f^{(n)}(0) + \lambda\Phi_t^{(n)} + \Phi_t^{(n+1)}\right]\mathrm{d}t,
\end{aligned}$$

which leads to (2.2) after a rearrangement of the terms on the righthand side. Finally,

$$\Phi_0^{(n)} = \left.\frac{\mathbb{E}[f^{(n)}(t-\Theta)1_{\{\Theta\le t\}} \mid \mathcal{F}_t]}{\mathbb{P}\{\Theta > t \mid \mathcal{F}_t\}}\right|_{t=0} = \frac{\mathbb{P}\{\Theta = 0\} f^{(n)}(0)}{\mathbb{P}\{\Theta > 0\}} = \frac{p}{1-p} f^{(n)}(0). \quad \square$$

Proof of Proposition 3.1. The existence of $\mathcal{F}_{T_n}$-measurable nonnegative random variables R_n satisfying (3.4) for $n \ge 0$ is proved, for example, by Brémaud [6], Davis [8], and Liptser and Shiryaev [16]. We can prove (3.5) by induction. Let us first show that $\mathbb{P}_\infty$-a.s. $\{\tau \ge T_1\} = \{R_0 \ge T_1\}$. Because $\{\tau \ge T_0\} \supseteq \{\tau \ge T_1\}$, we have

$$R_0 1_{\{\tau\ge T_1\}\cap\{R_0<T_1\}} = (R_0 \wedge T_1) 1_{\{\tau\ge T_1\}\cap\{R_0<T_1\}} = (\tau\wedge T_1) 1_{\{\tau\ge T_1\}\cap\{R_0<T_1\}} = T_1 1_{\{\tau\ge T_1\}\cap\{R_0<T_1\}},$$

which implies that $\mathbb{P}_\infty(\{\tau \ge T_1\}\cap\{R_0 < T_1\}) = 0$ and

$$\{\tau \ge T_1\} = \{\tau \ge T_1, R_0 < T_1\} \cup \{\tau \ge T_1, R_0 \ge T_1\} = \{\tau \ge T_1, R_0 \ge T_1\}, \quad \mathbb{P}_\infty\text{-a.s.} \qquad \text{(A.1)}$$

On the other hand, $\tau 1_{\{R_0\ge T_1\}\cap\{\tau<T_1\}} = (R_0\wedge T_1) 1_{\{R_0\ge T_1\}\cap\{\tau<T_1\}} = T_1 1_{\{R_0\ge T_1\}\cap\{\tau<T_1\}}$ implies that $\mathbb{P}_\infty(\{R_0 \ge T_1\}\cap\{\tau < T_1\}) = 0$, and

$$\{R_0 \ge T_1\} = \{R_0 \ge T_1, \tau < T_1\} \cup \{R_0 \ge T_1, \tau \ge T_1\} = \{R_0 \ge T_1, \tau \ge T_1\}, \quad \mathbb{P}_\infty\text{-a.s.},$$

which leads together with (A.1) to $\mathbb{P}_\infty$-a.s. $\{\tau \geq T_1\} = \{R_0 \geq T_1\}$. Now suppose that (3.5) holds for some $n \geq 1$. Let us show that it must also hold if n is replaced with $n+1$. Because $\{\tau \geq T_{n+1}\} \subseteq \{\tau \geq T_n\}$, by (3.4) we have

$$\begin{aligned}(T_n + R_n)1_{\{\tau \geq T_{n+1}, T_n + R_n < T_{n+1}\}} &= [(T_n + R_n) \wedge T_{n+1}]1_{\{\tau \geq T_{n+1}, T_n + R_n < T_{n+1}\}} \\ &= (\tau \wedge T_{n+1})1_{\{\tau \geq T_{n+1}, T_n + R_n < T_{n+1}\}} = T_{n+1}1_{\{\tau \geq T_{n+1}, T_n + R_n < T_{n+1}\}},\end{aligned}$$

which implies that $\mathbb{P}_\infty(\{\tau \geq T_{n+1}, T_n + R_n < T_{n+1}\}) = 0$, and

$$\{\tau \geq T_{n+1}\} \overset{\mathbb{P}_\infty\text{-a.s.}}{=} \{\tau \geq T_{n+1}, T_n + R_n \geq T_{n+1}\} \subseteq \{\tau \geq T_n, T_n + R_n \geq T_{n+1}\}. \tag{A.2}$$

On the other hand,

$$\begin{aligned}\tau 1_{\{T_n + R_n \geq T_{n+1}, T_n \leq \tau < T_{n+1}\}} &= (\tau \wedge T_{n+1})1_{\{T_n + R_n \geq T_{n+1}, T_n \leq \tau < T_{n+1}\}} \\ &= [(T_n + R_n) \wedge T_{n+1}]1_{\{T_n + R_n \geq T_{n+1}, T_n \leq \tau < T_{n+1}\}} = T_{n+1}1_{\{T_n + R_n \geq T_{n+1}, T_n \leq \tau < T_{n+1}\}}\end{aligned}$$

implies that $\mathbb{P}_\infty\{T_n + R_n \geq T_{n+1}, T_n \leq \tau < T_{n+1}\} = 0$ and

$$\begin{aligned}\{\tau \geq T_n, T_n + R_n \geq T_{n+1}\} &= \{T_n + R_n \geq T_{n+1}, T_n \leq \tau < T_{n+1}\} \cup \{T_n + R_n \geq T_{n+1}, \tau \geq T_{n+1}\} \\ &\overset{\mathbb{P}_\infty\text{-a.s.}}{=} \{T_n + R_n \geq T_{n+1}, \tau \geq T_{n+1}\} \subseteq \{\tau \geq T_{n+1}\},\end{aligned}$$

which in combination with (A.2) gives that

$$\begin{aligned}\{\tau \geq T_{n+1}\} &\overset{\mathbb{P}_\infty\text{-a.s.}}{=} \{\tau \geq T_n, T_n + R_n \geq T_{n+1}\} = \{\tau \geq T_n\} \cap \{T_n + R_n \geq T_{n+1}\} \\ &\overset{\mathbb{P}_\infty\text{-a.s.}}{=} \{R_0 \geq T_1, T_1 + R_1 \geq T_2, \ldots, T_{n-1} + R_{n-1} \geq T_n\} \cap \{T_n + R_n \geq T_{n+1}\} \\ &= \{R_0 \geq T_1, T_1 + R_1 \geq T_2, \ldots, T_n + R_n \geq T_{n+1}\},\end{aligned}$$

where the third equality follows from the induction hypothesis. This completes the proof of (3.5). Finally,

$$\begin{aligned}\{T_n \leq \tau < T_{n+1}\} &= \{\tau \geq T_n\} \setminus \{\tau \geq T_{n+1}\} = \{R_0 \geq T_1, T_1 + R_1 \geq T_2, \ldots, T_{n-1} + R_{n-1} \geq T_n\} \setminus \\ &\quad (\{R_0 \geq T_1, T_1 + R_1 \geq T_2, \ldots, T_{n-1} + R_{n-1} \geq T_n\} \cap \{T_n + R_n \geq T_{n+1}\}) \\ &= \{R_0 \geq T_1, T_1 + R_1 \geq T_2, \ldots, T_{n-1} + R_{n-1} \geq T_n, T_n + R_n < T_{n+1}\}\end{aligned}$$

proves (3.6) and completes the proof of Proposition 3.1. □

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

$$\begin{aligned}&\mathbb{E}^\phi_\infty\left[\int_0^{\tau \wedge T_k} e^{-\lambda t} g(\Phi_t)\mathrm{d}t + 1_{\{\tau \geq T_k\}} e^{-\lambda T_k} v_{n-k}(\Phi_{T_k})\right] \\ &\quad \geq \mathbb{E}^\phi_\infty\left[\int_0^{\tau \wedge T_{k-1}} e^{-\lambda t} g(\Phi_t)\mathrm{d}t + 1_{\{\tau \geq T_{k-1}\}} e^{-\lambda T_{k-1}} v_{n-k+1}(\Phi_{T_{k-1}})\right] \quad \text{for every } 1 \leq k \leq n,\end{aligned} \tag{A.3}$$

which will then imply that

$$\mathbb{E}^\phi_\infty\left[\int_0^{\tau \wedge T_n} e^{-\lambda t} g(\Phi_t)\mathrm{d}t\right] = \mathbb{E}^\phi_\infty\left[\int_0^{\tau \wedge T_n} e^{-\lambda t} g(\Phi_t)\mathrm{d}t + 1_{\{\tau \geq T_n\}} e^{-\lambda T_n} v_0(\Phi_{T_n})\right]$$

$$\geq \mathbb{E}_\infty^\phi \left[\int_0^{\tau\wedge T_{n-1}} e^{-\lambda t} g(\Phi_t) \mathrm{d}t + 1_{\{\tau \geq T_{n-1}\}} e^{-\lambda T_{n-1}} v_1(\Phi_{T_{n-1}}) \right] \geq \cdots$$
$$\geq \mathbb{E}_\infty^\phi \left[\int_0^{\tau\wedge T_0} e^{-\lambda t} g(\Phi_t) \mathrm{d}t + 1_{\{\tau \geq T_0\}} e^{-\lambda T_0} v_n(\Phi_{T_0}) \right] = v_n(\phi),$$

and taking the infimum of both sides over all $\tau \in \mathcal{S}$ gives $V_n(\phi) = \inf_{\tau\in\mathcal{S}} \mathbb{E}_\infty^\phi \left[\int_0^{\tau\wedge T_n} e^{-\lambda t} g(\Phi_t) \mathrm{d}t \right] \geq v_n(\phi)$ 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.3). 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

$$1_{\{\tau \geq T_{k-1}\}}(\tau \wedge T_k) = 1_{\{\tau \geq T_{k-1}\}}[(T_{k-1} + R_{k-1}) \wedge T_k],$$
$$\{\tau \geq T_k\} = \{\tau \geq T_{k-1}, T_{k-1} + R_{k-1} \geq T_k\},$$

and we have

$$\begin{aligned}
&\mathbb{E}_\infty^\phi \left[\int_0^{\tau\wedge T_k} e^{-\lambda t} g(\Phi_t) \mathrm{d}t + 1_{\{\tau \geq T_k\}} e^{-\lambda T_k} v_{n-k}(\Phi_{T_k}) \right] \\
&= \mathbb{E}_\infty^\phi \left[\int_0^{\tau\wedge T_{k-1}} e^{-\lambda t} g(\Phi_t) \mathrm{d}t + 1_{\{\tau \geq T_{k-1}\}} \int_{T_{k-1}}^{\tau\wedge T_k} e^{-\lambda t} g(\Phi_t) \mathrm{d}t + 1_{\{\tau \geq T_k\}} e^{-\lambda T_k} v_{n-k}(\Phi_{T_k}) \right] \\
&= \mathbb{E}_\infty^\phi \Bigg[\int_0^{\tau\wedge T_{k-1}} e^{-\lambda t} g(\Phi_t) \mathrm{d}t \\
&\qquad + 1_{\{\tau \geq T_{k-1}\}} e^{-\lambda T_{k-1}} \Bigg\{ \int_{T_{k-1}}^{(T_{k-1}+R_{k-1})\wedge T_k} e^{-\lambda(t-T_{k-1})} g(\varphi(t - T_{k-1}, \Phi_{T_{k-1}})) \mathrm{d}t \\
&\qquad + 1_{\{T_{k-1}+R_{k-1} \geq T_k\}} e^{-\lambda(T_k - T_{k-1})} v_{n-k} \left(\varphi\left(T_k - T_{k-1}, \Phi_{T_{k-1}}\right) \frac{\lambda_1}{\lambda_0} \frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(Z_k) \right) \Bigg\} \Bigg] \\
&= \mathbb{E}_\infty^\phi \Bigg[\int_0^{\tau\wedge T_{k-1}} e^{-\lambda t} g(\Phi_t) \mathrm{d}t + 1_{\{\tau \geq T_{k-1}\}} e^{-\lambda T_{k-1}} \mathbb{E}_\infty^\phi \Bigg\{ \int_0^{R_{k-1}\wedge(T_k - T_{k-1})} e^{-\lambda t} g(\varphi(t, \Phi_{T_{k-1}})) \mathrm{d}t \\
&\qquad + 1_{\{R_{k-1} \geq T_k - T_{k-1}\}} e^{-\lambda(T_k - T_{k-1})} v_{n-k} \left(\varphi\left(T_k - T_{k-1}, \Phi_{T_{k-1}}\right) \frac{\lambda_1}{\lambda_0} \frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(Z_k) \right) \Bigg| \mathcal{F}_{T_{k-1}} \Bigg\} \Bigg].
\end{aligned}$$

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

$$\begin{aligned}
&\mathbb{E}_\infty^\phi \Bigg\{ \int_0^{R_{k-1}\wedge(T_k - T_{k-1})} e^{-\lambda t} g(\varphi(t, \Phi_{T_{k-1}})) \mathrm{d}t \\
&\qquad + 1_{\{R_{k-1} \geq T_k - T_{k-1}\}} e^{-\lambda(T_k - T_{k-1})} v_{n-k} \left(\varphi\left(T_k - T_{k-1}, \Phi_{T_{k-1}}\right) \frac{\lambda_1}{\lambda_0} \frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(Z_k) \right) \Bigg| \mathcal{F}_{T_{k-1}} \Bigg\} \\
&= \mathbb{E}_\infty^\phi \left\{ \int_0^{r\wedge T_1} e^{-\lambda t} g(\varphi(t, \phi)) \mathrm{d}t + 1_{\{r \geq T_1\}} e^{-\lambda T_1} v_{n-k} \left(\varphi\left(T_1, \phi\right) \frac{\lambda_1}{\lambda_0} \frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(Z_1) \right) \right\} \Bigg|_{\substack{r = R_{k-1} \\ \phi = \Phi_{T_{k-1}}}} \\
&= \left\{ \int_0^r e^{-(\lambda+\lambda_0)t} [g + \lambda_0 (K v_{n-k})](\varphi(t, \phi)) \mathrm{d}t \right\} \Bigg|_{\substack{r = R_{k-1} \\ \phi = \Phi_{T_{k-1}}}} = (J v_{n-k})(\Phi_{T_{k-1}}, R_{k-1}),
\end{aligned}$$

and substituting this into previous displayed equation gives

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

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.3) and Proposition 3.2. □

Proof of Lemma 3.2. For every $n \geq 1$, $0 \leq k \leq n$, and $\varepsilon > 0$, we can write

$$\begin{aligned}
&\{\tau_{n,\varepsilon} \geq T_k\} = \{\tau_{n,\varepsilon} \geq T_1, \tau_{n,\varepsilon} \geq T_k\} = \{r_{n,\varepsilon/2}(\Phi_0) \geq T_1, \tau_{n,\varepsilon} \geq T_k\} \\
&= \left\{r_{n,\varepsilon/2}(\Phi_0) \geq T_1, T_1 + \tau_{n-1,\varepsilon/2}\circ\theta_{T_1} \geq T_k\right\} = \left\{r_{n,\varepsilon/2}(\Phi_0) \geq T_1, \tau_{n-1,\varepsilon/2}\circ\theta_{T_1} \geq T_k - T_1\right\} \\
&= \left\{r_{n,\varepsilon/2}(\Phi_0) \geq T_1, \tau_{n-1,\varepsilon/2}\circ\theta_{T_1} \geq T_{k-1}\circ\theta_{T_1}\right\} = \left\{r_{n,\varepsilon/2}(\Phi_0) \geq T_1\right\} \cap \left(\left\{\tau_{n-1,\varepsilon/2} \geq T_{k-1}\right\}\circ\theta_{T_1}\right).
\end{aligned}$$

Repeating this $k-1$ times gives

$$\begin{aligned}
\{\tau_{n,\varepsilon} \geq T_k\} &= \left\{r_{n,\varepsilon/2}(\Phi_0) \geq T_1\right\} \cap \left(\left\{r_{n-1,\varepsilon/4}(\Phi_0) \geq T_1\right\}\circ\theta_{T_1}\right) \cap \left(\left\{\tau_{n-2,\varepsilon/4} \geq T_{k-2}\right\}\circ\theta_{T_2}\right) \\
&= \left\{r_{n,\varepsilon/2}(\Phi_0) \geq T_1\right\} \cap \left(\left\{r_{n-1,\varepsilon/4}(\Phi_0) \geq T_1\right\}\circ\theta_{T_1}\right) \cap \left(\left\{r_{n-2,\varepsilon/8}(\Phi_0) \geq T_1\right\}\circ\theta_{T_2}\right) \\
&\qquad \cap \left(\left\{\tau_{n-3,\varepsilon/8} \geq T_{k-3}\right\}\circ\theta_{T_3}\right) = \ldots \\
&= \left\{r_{n,\varepsilon/2}(\Phi_0) \geq T_1\right\} \cap \left(\left\{r_{n-1,\varepsilon/4}(\Phi_0) \geq T_1\right\}\circ\theta_{T_1}\right) \cap \left(\left\{r_{n-2,\varepsilon/8}(\Phi_0) \geq T_1\right\}\circ\theta_{T_2}\right) \\
&\qquad \cap \ldots \cap \left(\left\{r_{n-k+1,\varepsilon/2^k}(\Phi_0) \geq T_1\right\}\circ\theta_{T_{k-1}}\right) \cap \left(\left\{\tau_{n-k,\varepsilon/2^k} \geq T_0\right\}\circ\theta_{T_k}\right) \\
&= \bigcap_{\ell=0}^{k-1}\left(\left\{r_{n-\ell,\varepsilon/2^{\ell+1}}(\Phi_0) \geq T_1\right\}\circ\theta_{T_\ell}\right) = \bigcap_{\ell=0}^{k-1}\left\{r_{n-\ell,\varepsilon/2^{\ell+1}}(\Phi_{T_\ell}) \geq T_{\ell+1} - T_\ell\right\},
\end{aligned}$$

because $T_0 = 0$ and $\{\tau_{n-k,\varepsilon/2^k} \geq T_0\} = \Omega$. If $0 \leq k \leq n-1$, then

$$\begin{aligned}
\{T_k \leq \tau_{n,\varepsilon} < T_{k+1}\} &= \{\tau_{n,\varepsilon} \geq T_k\} \setminus \{\tau_{n,\varepsilon} \geq T_{k+1}\} \\
&= \bigcap_{\ell=0}^{k-1}\left\{r_{n-\ell,\varepsilon/2^{\ell+1}}(\Phi_{T_\ell}) \geq T_{\ell+1} - T_\ell\right\} \setminus \bigcap_{\ell=0}^{k}\left\{r_{n-\ell,\varepsilon/2^{\ell+1}}(\Phi_{T_\ell}) \geq T_{\ell+1} - T_\ell\right\} \\
&= \left(\bigcap_{\ell=0}^{k-1}\left\{r_{n-\ell,\varepsilon/2^{\ell+1}}(\Phi_{T_\ell}) \geq T_{\ell+1} - T_\ell\right\}\right) \bigcap \left\{r_{n-k,\varepsilon/2^{k+1}}(\Phi_{T_k}) < T_{k+1} - T_k\right\},
\end{aligned}$$

which immediately implies that $\tau_{n,\varepsilon}1_{\{T_k\leq\tau_{n,\varepsilon}<T_{k+1}\}} = [T_k + r_{n-k,\varepsilon/2^{k+1}}(\Phi_{T_k})]1_{\{T_k\leq\tau_{n,\varepsilon}<T_{k+1}\}}$. Finally,

$$\begin{aligned}
&(\tau_{n,\varepsilon}\wedge T_{k+1})1_{\{\tau_{n,\varepsilon}\geq T_k\}} = \tau_{n,\varepsilon}1_{\{T_k\leq\tau_{n,\varepsilon}<T_{k+1}\}} + T_{k+1}1_{\{\tau_{n,\varepsilon}\geq T_{k+1}\}} \\
&= \left[T_k + r_{n-k,\varepsilon/2^{k+1}}(\Phi_{T_k})\right]1_{\{T_k\leq\tau_{n,\varepsilon}<T_{k+1}\}} + T_{k+1}1_{\{\tau_{n,\varepsilon}\geq T_{k+1}\}} \\
&= \left[(T_k + r_{n-k,\varepsilon/2^{k+1}}(\Phi_{T_k}))\wedge T_{k+1}\right]1_{\{T_k\leq\tau_{n,\varepsilon}<T_{k+1}\}} + \left[(T_k + r_{n-k,\varepsilon/2^{k+1}}(\Phi_{T_k}))\wedge T_{k+1}\right]1_{\{\tau_{n,\varepsilon}\geq T_{k+1}\}} \\
&= \left[(T_k + r_{n-k,\varepsilon/2^{k+1}}(\Phi_{T_k}))\wedge T_{k+1}\right]1_{\{\tau_{n,\varepsilon}\geq T_k\}},
\end{aligned}$$

because $T_k \leq \tau \equiv T_k + r_{n-k,\varepsilon/2^{k+1}}(\Phi_{T_k}) < T_{k+1}$ on $\{T_k \leq \tau_{n,\varepsilon} < T_{k+1}\}$ and by the first part of the lemma $T_k + r_{n-k,\varepsilon/2^{k+1}}(\Phi_{T_k}) \geq T_{k+1}$ on $\{\tau_{n,\varepsilon} \geq T_{n+1}\}$. □

Proof of Proposition 3.3 (by induction on n). For $n = 0$, the last displayed equation of Proposition 3.3 becomes $0 \le 0+\varepsilon$, 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 $n \ge 0$. Note that

$$\tau_{n+1,\varepsilon} = \begin{cases} r_{n+1,\varepsilon/2}(\Phi_0), & \text{if} \quad r_{n+1,\varepsilon/2}(\Phi_0) < T_1, \\ T_1 + \tau_{n,\varepsilon/2} \circ \theta_{T_1}, & \text{if} \quad r_{n+1,\varepsilon/2}(\Phi_0) \ge T_1, \end{cases}$$

and $\tau_{n+1,\varepsilon} \wedge T_1 = r_{n+1,\varepsilon/2}(\Phi_0) \wedge T_1$, $\{\tau_{n+1,\varepsilon} \ge T_1\} = \{r_{n+1,\varepsilon/2}(\Phi_0) \ge T_1\}$, and $\tau_{n+1,\varepsilon} = T_1 + \tau_{n,\varepsilon/2} \circ \theta_{T_1}$ on $\{\tau_{n+1,\varepsilon} \ge T_1\}$ by Lemma 3.2. Therefore,

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

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

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

by the induction hypothesis. Therefore,

$$\begin{aligned}
&\mathbb{E}^{\phi}_{\infty}\left[\int_0^{\tau_{n+1,\varepsilon}\wedge T_{n+1}} e^{-\lambda t} g(\Phi_t)\mathrm{d}t\right] \\
&\qquad \le \mathbb{E}^{\phi}_{\infty}\left[\int_0^{r_{n+1,\varepsilon/2}(\Phi_0)\wedge T_1} e^{-\lambda t} g(\varphi(t,\Phi_0))\mathrm{d}t + 1_{\{r_{n+1,\varepsilon/2}(\Phi_0)\ge T_1\}} e^{-\lambda T_1}\left(v_n(\Phi_{T_1}) + \frac{\varepsilon}{2}\right)\right] \\
&\qquad \le \mathbb{E}^{\phi}_{\infty}\left[\int_0^{r_{n+1,\varepsilon/2}(\Phi_0)\wedge T_1} e^{-\lambda t} g(\varphi(t,\Phi_0))\mathrm{d}t + 1_{\{r_{n+1,\varepsilon/2}(\Phi_0)\ge T_1\}} e^{-\lambda T_1} v_n(\Phi_{T_1})\right] + \frac{\varepsilon}{2} \\
&\qquad = (Jv_n)(\phi, r_{n+1,\varepsilon/2}(\phi)) + \frac{\varepsilon}{2} \le v_{n+1}(\phi) + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = v_{n+1}(\phi) + \varepsilon,
\end{aligned}$$

which completes the proof of Proposition 3.3. □

Proof of Lemma 3.3. Suppose that $w : \mathbb{R}^m_+ \mapsto \mathbb{R}$ is concave and bounded between -1 and 0. Then for every $r \ge 0$,

$$\begin{aligned}
(Jw)(\phi, r) &= \int_0^r e^{-(\lambda+\lambda_0)t}[g + \lambda_0(Kw)](\varphi(t,\phi))\mathrm{d}t \ge \int_0^r e^{-(\lambda+\lambda_0)t}(-\lambda - \lambda_0)\mathrm{d}t \\
&\ge -\int_0^\infty (\lambda+\lambda_0)e^{-(\lambda+\lambda_0)t}\mathrm{d}t = -1.
\end{aligned}$$

Taking the infimum over $r \geq 0$ gives $-1 \leq \inf_{r\geq 0}(Jw)(\phi, r) = (J_0 w)(\phi) \leq (Jw)(\phi, 0) = 0$. 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 $\phi \mapsto (Kw)(\varphi(t, \phi)) = \int_E w\left(\varphi(t, \phi)\frac{\lambda_1}{\lambda_0}\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(z)\right)\nu_0(\mathrm{d}z)$ 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 $\phi \mapsto (J_0 w)(\phi) = \inf_{r\geq 0}(Jw)(\phi, r)$ 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 $(J_0 w_1)(\cdot) \leq (J_0 w_2)(\cdot)$.

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 $r^{(i)}_\varepsilon(\phi)$, $i = 1, 2$ such that $(Jw_i)(\phi, r^{(i)}_\varepsilon(\phi)) \leq (J_0 w_i)(\phi) + \varepsilon$ for every $i = 1, 2$. Then

$$\begin{aligned}
&(J_0 w_1)(\phi) - (J_0 w_2)(\phi) \leq (Jw_1)(\phi, r^{(2)}_\varepsilon) - (Jw_2)(\phi, r^{(2)}_\varepsilon(\phi)) + \varepsilon \\
&= \int_0^{r^{(2)}_\varepsilon(\phi)} e^{-(\lambda+\lambda_0)t}[g + \lambda_0(Kw_1)](\varphi(t, \phi))\mathrm{d}t - \int_0^{r^{(2)}_\varepsilon(\phi)} e^{-(\lambda+\lambda_0)t}[g + \lambda_0(Kw_2)](\varphi(t, \phi))\mathrm{d}t + \varepsilon \\
&= \int_0^{r^{(2)}_\varepsilon(\phi)} \lambda_0 e^{-(\lambda+\lambda_0)t}(K(w_1 - w_2))(\varphi(t, \phi))\mathrm{d}t + \varepsilon \leq \|w_1 - w_2\| \int_0^\infty \lambda_0 e^{-(\lambda+\lambda_0)t}\mathrm{d}t + \varepsilon \\
&= \|w_1 - w_2\| \frac{\lambda_0}{\lambda + \lambda_0} + \varepsilon.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&(J_0 w_1)(\phi) - (J_0 w_2)(\phi) \geq (Jw_1)(\phi, r^{(1)}_\varepsilon(\phi)) - \varepsilon - (Jw_2)(\phi, r^{(1)}_\varepsilon(\phi)) \\
&= \int_0^{r^{(1)}_\varepsilon(\phi)} e^{-(\lambda+\lambda_0)t}[g + \lambda_0(Kw_1)](\varphi(t, \phi))\mathrm{d}t - \int_0^{r^{(1)}_\varepsilon(\phi)} e^{-(\lambda+\lambda_0)t}[g + \lambda_0(Kw_2)](\varphi(t, \phi))\mathrm{d}t - \varepsilon \\
&= \int_0^{r^{(1)}_\varepsilon(\phi)} \lambda_0 e^{-(\lambda+\lambda_0)t}(K(w_1 - w_2))(\varphi(t, \phi))\mathrm{d}t - \varepsilon \geq -\|w_1 - w_2\| \int_0^\infty \lambda_0 e^{-(\lambda+\lambda_0)t}\mathrm{d}t - \varepsilon \\
&= -\left(\|w_1 - w_2\| \frac{\lambda_0}{\lambda + \lambda_0} + \varepsilon\right).
\end{aligned}$$

Hence we have

$$|(J_0 w_1)(\phi) - (J_0 w_2)(\phi)| \leq \frac{\lambda_0}{\lambda + \lambda_0}\|w_1 - w_2\| + \varepsilon \qquad \text{for every } \varepsilon > 0 \text{ and } \phi \in \mathbb{R}^m_+.$$

Letting first $\varepsilon \downarrow 0$ and then taking the supremum over $\phi \in \mathbb{R}^m_+$ gives the desired inequality. Because $\lambda_0/(\lambda + \lambda_0) \in (0, 1)$, operator J_0 is a contraction on the collection of bounded functions $w : \mathbb{R}^m_+ \mapsto \mathbb{R}$. □

Proof of Lemma 3.4. 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

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

and $(Jw)(\phi, s) + e^{-(\lambda+\lambda_0)s}(J_0 w)(\varphi(s, \phi))$ equals

$$\int_0^s e^{-(\lambda+\lambda_0)u}[g + \lambda_0(Kw)](\varphi(u, \phi))\mathrm{d}u + e^{-(\lambda+\lambda_0)s}(J_0 w)(\varphi(s, \phi))$$

$$= \inf_{r\geq s} \int_0^r e^{-(\lambda+\lambda_0)u}[g+\lambda_0(Kw)](\varphi(u,\phi))\mathrm{d}u = (J_s w)(\phi).$$

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 Lemma 3.5. Let us define

$$\widetilde{\tau}_\varepsilon = \begin{Bmatrix} r_\varepsilon(\Phi_0), & \text{if} \quad r_\varepsilon(\Phi_0) < T_1 \\ T_1 + \widetilde{\tau}_\varepsilon \circ \theta_{T_1}, & \text{if} \quad r_\varepsilon(\Phi_0) \geq T_1 \end{Bmatrix}, \quad \varepsilon \geq 0.$$

Because $t \mapsto V(\varphi(t,\phi))$ is continuous for every $\phi \in \mathbb{R}^m_+$, we have $V(\varphi(r_\varepsilon(\phi),\phi)) \geq -\varepsilon$ if $r_\varepsilon(\phi) < \infty$ and $V(\Phi_{\widetilde{\tau}_\varepsilon}) \geq -\varepsilon$ on $\{\widetilde{\tau}_\varepsilon < \infty\}$, since $\widetilde{\tau}_\varepsilon 1_{[T_n,T_{n+1})}(\widetilde{\tau}_\varepsilon) = [T_n + r_\varepsilon(\Phi_{T_n})]1_{[T_n,T_{n+1})}(\widetilde{\tau}_\varepsilon)$, and

$$\begin{aligned} 1_{\{\widetilde{\tau}_\varepsilon<\infty\}}V(\Phi_{\widetilde{\tau}_\varepsilon}) &= \sum_{n=0}^\infty 1_{[T_n,T_{n+1})}(\widetilde{\tau}_\varepsilon)V(\Phi_{T_n+r_\varepsilon(\Phi_{T_n})}) = \sum_{n=0}^\infty 1_{[T_n,T_{n+1})}(\widetilde{\tau}_\varepsilon)V(\varphi(r_\varepsilon(\Phi_{T_n}),\Phi_{T_n})) \\ &\geq (-\varepsilon)1_{\{\widetilde{\tau}_\varepsilon<\infty\}}. \end{aligned}$$

Therefore, $\mathbb{P}_\infty$-a.s. $\tau_\varepsilon \leq \widetilde{\tau}_\varepsilon$. On the other hand, for every $n \geq 0$, there is a nonnegative $\mathcal{F}_{T_n}$-measurable random variable $R_{n,\varepsilon}$ such that $\tau_\varepsilon 1_{[T_n,T_{n+1})}(\tau_\varepsilon) = (T_n + R_{n,\varepsilon})1_{[T_n,T_{n+1})}(\tau_\varepsilon)$. Because $V(\cdot)$ is continuous and $t \mapsto \Phi_t$ is right-continuous and has left-limits, $t \mapsto V(\Phi_t)$ is right-continuous. Therefore,

$$\begin{aligned} (-\varepsilon)1_{[T_n,T_{n+1})}(\tau_\varepsilon) \leq V(\Phi_{\tau_\varepsilon})1_{[T_n,T_{n+1})}(\tau_\varepsilon) &= V(\Phi_{T_n+R_{n,\varepsilon}})1_{[T_n,T_{n+1})}(\tau_\varepsilon) \\ &= V(\varphi(R_{n,\varepsilon},\Phi_{T_n}))1_{[T_n,T_{n+1})}(\tau_\varepsilon) \end{aligned}$$

implies that $R_{n,\varepsilon} \geq r_\varepsilon(\Phi_{T_n})$ on $\{\tau_\varepsilon \in [T_n,T_{n+1})\}$. Thus,

$$\tau_\varepsilon 1_{\{\tau_\varepsilon<\infty\}} = \sum_{n=0}^\infty (T_n + R_{n,\varepsilon})1_{[T_n,T_{n+1})}(\tau_\varepsilon) \geq \sum_{n=0}^\infty (T_n + r_\varepsilon(\Phi_{T_n}))1_{[T_n,T_{n+1})}(\tau_\varepsilon) = \widetilde{\tau}_\varepsilon 1_{\{\tau_\varepsilon<\infty\}}.$$

Hence, $\tau_\varepsilon \geq \widetilde{\tau}_\varepsilon$ on $\{\tau_\varepsilon < \infty\}$ or simply $\tau_\varepsilon \geq \widetilde{\tau}_\varepsilon$. This proves that $\mathbb{P}_\infty$-a.s. $\tau_\varepsilon = \widetilde{\tau}_\varepsilon$, and without loss of generality we can take $R_{n,\varepsilon} = r_\varepsilon(\Phi_{T_n})$. □

Proof of Proposition 3.4. The result holds for $n = 0$. Suppose that for some $n \geq 0$ we have $\mathbb{E}^\phi_\infty[M_{\tau\wedge\tau_\varepsilon\wedge T_n}] = \mathbb{E}^\phi_\infty[M_0]$ for every $\phi \in \mathbb{R}^m_+$, $\varepsilon \geq 0$, and $(\mathcal{F}_t)_{t\geq 0}$-stopping time τ. Fix any $\phi \in \mathbb{R}^m_+$, $\varepsilon \geq 0$, and $(\mathcal{F}_t)_{t\geq 0}$-stopping time τ. Then

$$\begin{aligned} \mathbb{E}^\phi_\infty[M_{\tau\wedge\tau_\varepsilon\wedge T_{n+1}}] &= \mathbb{E}^\phi_\infty\left[M_{\tau\wedge\tau_\varepsilon\wedge T_n} + 1_{\{\tau\wedge\tau_\varepsilon\geq T_n\}}\left(M_{\tau\wedge\tau_\varepsilon\wedge T_{n+1}} - M_{T_n}\right)\right] \\ &= \mathbb{E}^\phi_\infty\left[M_{\tau\wedge\tau_\varepsilon\wedge T_n}\right] + \mathbb{E}^\phi_\infty\left[1_{\{\tau\wedge\tau_\varepsilon\geq T_n\}}\left(M_{\tau\wedge\tau_\varepsilon\wedge T_{n+1}} - M_{T_n}\right)\right] \\ &= \mathbb{E}^\phi_\infty[M_0] + \mathbb{E}^\phi_\infty\left[1_{\{\tau\wedge\tau_\varepsilon\geq T_n\}}\left(M_{\tau\wedge\tau_\varepsilon\wedge T_{n+1}} - M_{T_n}\right)\right] \end{aligned}$$

by the induction hypothesis. We shall prove that the second term on the righthand side equals zero. Since by Proposition 3.1 there exists a nonnegative $(\mathcal{F}_t)_{t\geq 0}$-measurable random variable R_n such that

$$\begin{aligned} (\tau\wedge\tau_\varepsilon\wedge T_{n+1})1_{\{\tau\wedge\tau_\varepsilon\geq T_n\}} &= [(T_n + R_n)\wedge T_{n+1}]1_{\{\tau\wedge\tau_\varepsilon\geq T_n\}}, \\ (\tau\wedge\tau_\varepsilon)1_{\{T_n\leq\tau\wedge\tau_\varepsilon<T_{n+1}\}} &= (T_n + R_n)1_{\{T_n\leq\tau\wedge\tau_\varepsilon<T_{n+1}\}}, \\ 1_{\{\tau\wedge\tau_\varepsilon\geq T_{n+1}\}} &= 1_{\{\tau\wedge\tau_\varepsilon\geq T_n\}}1_{\{T_n+R_n\geq T_{n+1}\}}, \end{aligned}$$

$$1_{\{T_n \le \tau\wedge\tau_\varepsilon < T_{n+1}\}} = 1_{\{\tau\wedge\tau_\varepsilon \ge T_n\}} 1_{\{T_n+R_n < T_{n+1}\}},$$

we can write $\mathbb{E}^\phi_\infty \left[1_{\{\tau\wedge\tau_\varepsilon\ge T_n\}}\left(M_{\tau\wedge\tau_\varepsilon\wedge T_{n+1}} - M_{T_n}\right)\right]$ as

$$\begin{aligned}
&\mathbb{E}^\phi_\infty \left[1_{\{\tau\wedge\tau_\varepsilon\ge T_n\}}\left(\int_{T_n}^{\tau\wedge\tau_\varepsilon\wedge T_{n+1}} e^{-\lambda t} g(\Phi_t)\mathrm{d}t + e^{-\lambda(\tau\wedge\tau_\varepsilon\wedge T_{n+1})} V(\Phi_{\tau\wedge\tau_\varepsilon\wedge T_{n+1}}) - e^{-\lambda T_n} V(\Phi_{T_n})\right)\right] \\
&= \mathbb{E}^\phi_\infty \Bigg[1_{\{\tau\wedge\tau_\varepsilon\ge T_n\}}\Bigg(\int_{T_n}^{\tau\wedge\tau_\varepsilon\wedge T_{n+1}} e^{-\lambda t} g(\Phi_t)\mathrm{d}t + 1_{\{\tau\wedge\tau_\varepsilon\ge T_{n+1}\}} e^{-\lambda T_{n+1}} V(\Phi_{T_{n+1}}) \\
&\qquad\qquad + 1_{\{\tau\wedge\tau_\varepsilon < T_{n+1}\}} e^{-\lambda(\tau\wedge\tau_\varepsilon)} V(\Phi_{\tau\wedge\tau_\varepsilon}) - e^{-\lambda T_n} V(\Phi_{T_n})\Bigg)\Bigg] \\
&= \mathbb{E}^\phi_\infty \Bigg[1_{\{\tau\wedge\tau_\varepsilon\ge T_n\}}\Bigg(\int_{T_n}^{(T_n+R_n)\wedge T_{n+1}} e^{-\lambda t} g(\Phi_t)\mathrm{d}t + 1_{\{T_n+R_n\ge T_{n+1}\}} e^{-\lambda T_{n+1}} V(\Phi_{T_{n+1}}) \\
&\qquad\qquad + 1_{\{T_n+R_n < T_{n+1}\}} e^{-\lambda(T_n+R_n)} V(\Phi_{T_n+R_n}) - e^{-\lambda T_n} V(\Phi_{T_n})\Bigg)\Bigg] \\
&= \mathbb{E}^\phi_\infty \Bigg[1_{\{\tau\wedge\tau_\varepsilon\ge T_n\}} e^{-\lambda T_n}\Bigg(\int_{T_n}^{(T_n+R_n)\wedge T_{n+1}} e^{-\lambda(t-T_n)} g(\varphi(t-T_n,\Phi_{T_n})\mathrm{d}t \\
&\qquad + 1_{\{R_n\ge T_{n+1}-T_n\}} e^{-\lambda(T_{n+1}-T_n)} V\left(\varphi(T_{n+1}-T_n,\Phi_{T_n})\frac{\lambda_1}{\lambda_0}\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(Z_{n+1})\right) \\
&\qquad\qquad + 1_{\{R_n < T_{n+1}-T_n\}} e^{-\lambda R_n} V(\varphi(R_n,\Phi_{T_n})) - V(\Phi_{T_n})\Bigg)\Bigg].
\end{aligned}$$

Because the random variables $1_{\{\tau\wedge\tau_\varepsilon\ge T_n\}}$, T_n, R_n, Φ_{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 λ_0 and with distribution ν_0 under $\mathbb{P}_\infty$, respectively, taking the conditional expectation with respect to $\mathcal{F}_{T_n}$ inside the above expectation gives

$$\begin{aligned}
&\mathbb{E}^\phi_\infty \Bigg[1_{\{\tau\wedge\tau_\varepsilon\ge T_n\}} e^{-\lambda T_n}\Bigg(\mathbb{E}^\phi_\infty\Bigg\{\int_0^{R_n\wedge(T_{n+1}-T_n)} e^{-\lambda t} g(\varphi(t,\Phi_{T_n}))\mathrm{d}t \\
&\qquad + 1_{\{R_n\ge T_{n+1}-T_n\}} e^{-\lambda(T_{n+1}-T_n)} V\left(\varphi(T_{n+1}-T_n,\Phi_{T_n})\frac{\lambda_1}{\lambda_0}\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(Z_{n+1})\right)\Bigg|\,\mathcal{F}_{T_n}\Bigg\} \\
&\qquad\qquad + \mathbb{P}^\phi_\infty\{T_{n+1}-T_n > R_n \mid \mathcal{F}_{T_n}\} e^{-\lambda R_n} V(\varphi(R_n,\Phi_{T_n})) - V(\Phi_{T_n})\Bigg)\Bigg], \\
&= \mathbb{E}^\phi_\infty \Bigg[1_{\{\tau\wedge\tau_\varepsilon\ge T_n\}} e^{-\lambda T_n}\Bigg(\int_0^{R_n} e^{-(\lambda+\lambda_0)t} g(\varphi(t,\Phi_{T_n}))\mathrm{d}t \\
&\qquad + \int_0^{R_n} \lambda_0 e^{-(\lambda+\lambda_0)t} \underbrace{\int_E V\left(\varphi(t,\Phi_{T_n})\frac{\lambda_1}{\lambda_0}\frac{\mathrm{d}\nu_1}{\mathrm{d}\nu_0}(z)\right)\nu_0(\mathrm{d}z)}_{(KV)(\varphi(t,\Phi_{T_n}))}\mathrm{d}t \\
&\qquad\qquad + e^{-(\lambda+\lambda_0)R_n} V(\varphi(R_n,\Phi_{T_n})) - V(\Phi_{T_n})\Bigg)\Bigg],
\end{aligned}$$

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

Because

$$
\begin{aligned}
(\tau\wedge\tau_\varepsilon)1_{\{T_n \le \tau\wedge\tau_\varepsilon < T_{n+1}\}} &= (T_n + R_n)1_{\{T_n \le \tau\wedge\tau_\varepsilon < T_{n+1}\}}, \\
\tau_\varepsilon 1_{\{T_n \le \tau_\varepsilon < T_{n+1}\}} &= [T_n + r_\varepsilon(\Phi_{T_n})]1_{\{T_n \le \tau_\varepsilon < T_{n+1}\}},
\end{aligned}
$$

where $r_\varepsilon(\cdot)$ is defined as in Lemma 3.5, and since $\tau\wedge\tau_\varepsilon \le \tau_\varepsilon$, we have $R_n \le r_\varepsilon(\Phi_{T_n})$ on $\{T_n \le \tau\wedge\tau_\varepsilon < T_{n+1}\}$. However, since R_n and $r_\varepsilon(\Phi_{T_n})$ are $\mathcal{F}_{T_n}$-measurable, we must also have $R_n \le r_\varepsilon(\Phi_{T_n})$ on $\{\tau\wedge\tau_\varepsilon \ge T_n\}$. Because $V(\varphi(s,\Phi_{T_n})) < -\varepsilon \le 0$ for every $0 \le s < r_\varepsilon(\Phi_{T_n})$, we also have $V(\varphi(s,\Phi_{T_n})) < -\varepsilon \le 0$ for every $0 \le s < R_n$ on $\{\tau\wedge\tau_\varepsilon \ge T_n\}$. Then Corollary 3.3 guarantees that $(J_{R_n}V)(\Phi_{T_n}) = V(\Phi_{T_n})$ on $\{\tau\wedge\tau_\varepsilon \ge T_n\}$ and we finally have

$$
\mathbb{E}^{\phi}_{\infty}\left[1_{\{\tau\wedge\tau_\varepsilon \ge T_n\}}\left(M_{\tau\wedge\tau_\varepsilon\wedge T_{n+1}} - M_{T_n}\right)\right] = \mathbb{E}^{\phi}_{\infty}\left[1_{\{\tau\wedge\tau_\varepsilon \ge T_n\}} e^{-\lambda T_n}\left((J_{R_n}V)(\Phi_{T_n}) - V(\Phi_{T_n})\right)\right] = 0,
$$

which completes the proof of Proposition 3.4. □

Proof of Corollary 3.4. Observe that since $\|V\| \le 1$ by Corollary 3.2, we have $|M_{t\wedge\tau_\varepsilon\wedge T_n}| \le 1 + \int_0^{T_n} e^{-\lambda u}\Phi^{(1)}_u \mathrm{d}u$ for every $t \ge 0$. On the other hand, since T_n is an $(\mathcal{F}_t)_{t\ge 0}$-stopping time, the derivations on page 4 show that

$$
\begin{aligned}
(1-p)\mathbb{E}_\infty\left[\int_0^{T_n} e^{-\lambda u}\Phi^{(1)}_u \mathrm{d}u\right] &= \mathbb{E}[f(T_n - \Theta)1_{\{T_n \ge \Theta\}}] - f(0)\mathbb{P}\{T_n \ge \Theta\} \\
&= \mathbb{E}\left[(T_n - \Theta)^m 1_{\{T_n \ge \Theta\}}\right] \le \mathbb{E}[T_n^m \mid T_n \ge \Theta]\mathbb{P}\{T_n \ge \Theta\} \le \mathbb{E}[T_n^m \mid T_n \ge \Theta] \\
&\qquad \le \mathbb{E}[(\Theta + \widetilde{T}_n)^m] \le 2^{m-1}\mathbb{E}[\Theta^m + \widetilde{T}_n^m] < \infty,
\end{aligned}
$$

where $\widetilde{T}_n$ is the n-th arrival time of a Poisson process with arrival rate λ_1 under $\mathbb{P}$, and the last inequality follows from that all of the moments of an Erlang distribution are finite. Therefore, $\sup_{t\ge 0}\|M_{t\wedge\tau_\varepsilon\wedge T_n}\|$ is bounded from above by the integrable random variable $1 + \int_0^{T_n} e^{-\lambda u}\Phi^{(1)}_u \mathrm{d}u$, and $\{M_{t\wedge\tau_\varepsilon\wedge T_n}, \mathcal{F}_t;\ t \ge 0\}$ is uniformly integrable under $\mathbb{P}_\infty$.

Fix any $0 \le s \le t$ and $F \in \mathcal{F}_s$. Let us define $\tau = t1_F + s1_{\Omega\setminus F}$. By Proposition 3.4,

$$
\begin{aligned}
\mathbb{E}_\infty[M_{s\wedge\tau_\varepsilon\wedge T_n}] = \mathbb{E}_\infty[M_0] = \mathbb{E}_\infty[M_{\tau\wedge\tau_\varepsilon\wedge T_n}] &= \mathbb{E}_\infty\left[M_{t\wedge\tau_\varepsilon\wedge T_n}1_F + M_{s\wedge\tau_\varepsilon\wedge T_n}1_{\Omega\setminus F}\right] \\
&= \mathbb{E}_\infty[M_{t\wedge\tau_\varepsilon\wedge T_n}1_F] + \mathbb{E}_\infty\left[M_{s\wedge\tau_\varepsilon\wedge T_n}1_{\Omega\setminus F}\right],
\end{aligned}
$$

which can be rearranged into $\mathbb{E}_\infty[M_{s\wedge\tau_\varepsilon\wedge T_n}1_F] = \mathbb{E}_\infty[M_{t\wedge\tau_\varepsilon\wedge T_n}1_F] = \mathbb{E}_\infty[\mathbb{E}_\infty(M_{t\wedge\tau_\varepsilon\wedge T_n} \mid \mathcal{F}_s)1_F]$. Because $F \in \mathcal{F}_s$ is arbitrary, we conclude that $\mathbb{E}_\infty[M_{t\wedge\tau_\varepsilon\wedge T_n} \mid \mathcal{F}_s] = M_{s\wedge\tau_\varepsilon\wedge T_n}$. □

REFERENCES

[1] M. Baron and A. G. Tartakovsky. Asymptotic optimality of change-point detection schemes in general continuous-time models. *Sequential Anal.*, 25(3):257–296, 2006.

[2] M. Basseville and I. V. Nikiforov. *Detection of abrupt changes: theory and application*. Prentice Hall Information and System Sciences Series. Prentice Hall Inc., Englewood Cliffs, NJ, 1993.

[3] E. Bayraktar and S. Dayanik. Poisson disorder problem with exponential penalty for delay. *Math. Oper. Res.*, 31(2):217–233, 2006.

[4] E. Bayraktar, S. Dayanik, and I. Karatzas. The standard Poisson disorder problem revisited. *Stochastic Process. Appl.*, 115(9):1437–1450, 2005.

[5] M. Beibel. A note on sequential detection with exponential penalty for the delay. *Ann. Statist.*, 28(6):1696–1701, 2000.

[6] P. Brémaud. *Point processes and queues*. Springer-Verlag, New York, 1981. Martingale dynamics, Springer Series in Statistics.

[7] E. A. Coddington and R. Carlson. *Linear ordinary differential equations*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[8] M. H. A. Davis. *Markov models and optimization*, volume 49 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1993.

[9] S. Dayanik, H. V. Poor, and S. O. Sezer. Multisource Bayesian sequential change detection. *Ann. Appl. Probab.*, 18(2):552–590, 2008.

[10] S. Dayanik and S. O. Sezer. Compound Poisson disorder problem. *Math. Oper. Res.*, 31(4):649–672, 2006.

[11] L. I. Galchuk and B. L. Rozovskii. The problem of "disorder" for a Poisson process. *Teor. Verojatnost. i Primenen.*, 16:729–734, 1971.

[12] P. V. Gapeev. The disorder problem for compound Poisson processes with exponential jumps. *Ann. Appl. Probab.*, 15(1A):487–499, 2005.

[13] U. S. Gugerli. Optimal stopping of a piecewise-deterministic Markov process. *Stochastics*, 19(4):221–236, 1986.

[14] O. Hadjiliadis. Optimality of the 2-CUSUM drift equalizer rules for detecting two-sided alternatives in the Brownian motion model. *J. Appl. Probab.*, 42(4):1183–1193, 2005.

[15] O. Hadjiliadis and V. Moustakides. Optimal and asymptotically optimal CUSUM rules for change point detection in the Brownian motion model with multiple alternatives. *Teor. Veroyatn. Primen.*, 50(1):131–144, 2005.

[16] R. S. Liptser and A. N. Shiryaev. *Statistics of random processes. II*, volume 6 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, expanded edition, 2001.

[17] G. Lorden. Procedures for reacting to a change in distribution. *Ann. Math. Statist.*, 42:1897–1908, 1971.

[18] G. V. Moustakides. Optimal stopping times for detecting changes in distributions. *Ann. Statist.*, 14(4):1379–1387, 1986.

[19] G. V. Moustakides. Optimality of the CUSUM procedure in continuous time. *Ann. Statist.*, 32(1):302–315, 2004.

[20] G. V. Moustakides. Sequential change detection revisited. *Ann. Statist.*, 36(2):787–807, 2008.

[21] G. Peskir and A. Shiryaev. *Optimal stopping and free-boundary problems*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2006.

[22] G. Peskir and A. N. Shiryaev. Solving the Poisson disorder problem. In *Advances in finance and stochastics*, pages 295–312. Springer, Berlin, 2002.

[23] M. Pollak. Optimal detection of a change in distribution. *Ann. Statist.*, 13(1):206–227, 1985.

[24] H. V. Poor. Quickest detection with exponential penalty for delay. *Ann. Statist.*, 26(6):2179–2205, 1998.

[25] H. V. Poor and O. Hadjiliadis. *Quickest Detection*. Cambridge University Press, 2008.

[26] A. N. Shiryaev. On optimum methods in quickest detection problems. *Theory of Probability and Its Applications*, (8):22–26, 1963.

[27] A. N. Shiryaev. On markov sufficient statistics in non-additive bayes problems of sequential analysis. *Theory of Probability and its Applications*, 9(4):604–618, 1964.

[28] A. N. Shiryaev. Minimax optimality of the method of cumulative sums (CUSUM) in the continuous time case. *Uspekhi Mat. Nauk*, 51(4(310)):173–174, 1996.

[29] A. N. Shiryaev. Generalized bayesian nonlinear quickest detection problems: On markov family of sufficient statistics. In A. G. M. G. M. Sarychev, A.; Shiryaev, editor, *Mathematical Control Theory and Finance*, pages 377–386. Springer, 2008.

[30] A. G. Tartakovsky. Multidecision quickest change-point detection: previous achievements and open problems. *Sequential Anal.*, 27(2):201–231, 2008.

[31] A. G. Tartakovsky and V. V. Veeravalli. General asymptotic Bayesian theory of quickest change detection. *Teor. Veroyatn. Primen.*, 49(3):538–582, 2004.