Compound Poisson disorder problem

Savas Dayanik
Princeton University, Department of Operations Research and Financial Engineering, and Bendheim Center for Finance, Princeton, NJ 08544
email: sdayanik@princeton.edu `http://www.princeton.edu/~sdayanik`

Semih Onur Sezer
Princeton University, Department of Operations Research and Financial Engineering, Princeton, NJ 08544
email: ssezer@princeton.edu `http://www.princeton.edu/~ssezer`

This paper is dedicated to our teacher and mentor,
Professor Erhan Çınlar, on the occasion of his 65th birthday.

In compound Poisson disorder problem, arrival rate and/or jump distribution of some compound Poisson process change suddenly at some unknown and unobservable time. The problem is to detect the change (or disorder) time as quickly as possible. A sudden regime-shift may require some counter-measures be taken promptly, and a quickest detection rule can help with those efforts. We describe complete solution of compound Poisson disorder problem with several standard Bayesian risk measures. Solution methods are feasible for numerical implementation and are illustrated on examples.

Key words: Poisson disorder problem; quickest detection; compound Poisson processes, optimal stopping

MSC2000 Subject Classification: Primary: 62L10; Secondary: 62L15, 62C10, 60G40

OR/MS subject classification: Primary: Statistics: Bayesian, Estimation; Secondary: Dynamic programming/optimal control: Applications

---

1. Introduction. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space hosting a compound Poisson process

$$X_t = X_0 + \sum_{k=1}^{N_t} Y_k, \qquad t \geq 0. \tag{1}$$

Jumps arrive according to a standard Poisson process $N = \{N_t;\, t \geq 0\}$ at some rate $\lambda_0 > 0$. The marks at each jump are i.i.d. $\mathbb{R}^d$-valued random variables $Y_1, Y_2, \ldots$ with some common distribution $\nu_0(\cdot)$ independent of the arrival process N. The process X may represent customer orders arriving in batches to a multi-product service system, claims of various sizes filed with an insurance company, or sizes of electronic files requested for download from a network server.

Suppose that, at an *unknown and unobservable* time θ, the initial arrival rate λ_0 and mark distribution $\nu_0(\cdot)$ of the process X change suddenly to λ_1 and $\nu_1(\cdot)$, respectively. This regime shift at the *disorder time* θ may become detrimental on the underlying system unless certain counter-measures are taken quickly. For example, optimal inventory levels, insurance premiums, or number of network servers may need to be revised as soon as the regime changes in order to maintain profitability, avoid bankruptcy, or ensure the network availability.

The objective of this paper is to *detect the disorder time θ as quickly as possible* in order to give decision makers an opportunity to react the regime change on a timely basis. We assume that λ_0, λ_1, $\nu_0(\cdot)$ and $\nu_1(\cdot)$ are known, and that the disorder time θ is a random variable whose prior distribution is

$$\mathbb{P}\{\theta = 0\} = \pi \qquad \text{and} \qquad \mathbb{P}\{\theta > t\} = (1-\pi)e^{-\lambda t}, \quad t \geq 0; \quad \pi \in [0,1),\ \lambda > 0.$$

The disorder time θ is still unobservable, and we need a quickest detection rule adapted to the history $\mathbb{F}$ of the observation process X in (1). More precisely, we would like to find a stopping time τ of the process X whose *Bayes risk*

$$R_\tau(\pi) \triangleq \mathbb{P}\{\tau < \theta\} + c\,\mathbb{E}(\tau - \theta)^+, \qquad \pi \in [0,1),\ \tau \in \mathbb{F} \tag{2}$$

is the smallest ($x^+ \triangleq \max\{x, 0\}$.) If an $\mathbb{F}$-stopping time τ attains the *minimum Bayes risk*

$$U(\pi) \triangleq \inf_{\tau \in \mathbb{F}} R_\tau(\pi), \qquad \pi \in [0,1), \tag{3}$$

then it is called a *Bayes-optimal alarm time* and solves optimally the tradeoff between the *false-alarm frequency* $\mathbb{P}\{\tau < \theta\}$ and the *expected detection delay cost* $c \cdot \mathbb{E}(\tau - \theta)^+$.

All of the early work has dealt with (simple) *Poisson disorder problem.* In that problem and in the notation above, the observation process was the counting process N whose rate changes at some unobservable time θ from some known constant λ_0 to some other λ_1. While the question was the same; namely, to detect the disorder time θ as quickly as possible, the information about marks $Y_1, Y_2, \ldots$ were ignored completely. This omission was understandable because of the difficulty of the problem: simple Poisson disorder problem was solved completely by Peskir and Shiryaev [13] only recently—more than thirty years after it was formulated by Galchuk and Rozovskii [9] for the first time. In the meantime, partial solutions and new insights were provided. Most notably, Davis [5] showed that quickest detection rules should not differ much if they are to minimize some "standard" Bayes risks; namely, one of $R^{(1)}$, $R^{(2)}$ (same as R of (2)), or $R^{(3)}$ in

$$\begin{aligned} &R^{(1)}_\tau(\pi) \triangleq \mathbb{P}\{\tau < \theta - \varepsilon\} + c\,\mathbb{E}(\tau-\theta)^+, \quad &R^{(2)}_\tau(\pi) \triangleq \mathbb{P}\{\tau < \theta\} + c\,\mathbb{E}(\tau-\theta)^+, \\ &R^{(3)}_\tau(\pi) \triangleq \mathbb{E}(\theta-\tau)^+ + c\,\mathbb{E}(\tau-\theta)^+, \quad &R^{(4)}_\tau(\pi) \triangleq \mathbb{P}\{\tau < \theta\} + c\,\mathbb{E}[e^{\alpha(\tau-\theta)^+} - 1], \end{aligned} \tag{4}$$

where ε, c, and α are some known positive constants (see also Shiryaev [17]). Recently, Bayraktar and Dayanik [1] solved simple Poisson disorder problem with Bayes risk $R^{(4)}$ in (4), whose exponential detection-delay penalty makes it more suitable for financial applications. Later, Bayraktar, Dayanik, and Karatzas [2] showed that the measure $R^{(4)}$ is also a "standard" Bayes risk (if the latter is redefined suitably) and gave a general solution method for standard problems.

For the first time, Gapeev [10] has recently succeeded to include the observed marks $Y_1, Y_2, \ldots$ into an optimal decision rule in order to detect the disorder time (more) quickly and accurately. He provided the full solution for the following *very special instance* of compound Poisson disorder problem: before and after the disorder time θ, real-valued marks $Y_1, Y_2, \ldots$ have exponential distributions, and the expected mark sizes are the same as the corresponding jump arrival rates. Namely, the mark distributions are

$$\nu_i(A) = \int_A \frac{1}{\lambda_i} \exp\left\{-\frac{1}{\lambda_i}\, y\right\} dy, \qquad A \in \mathcal{B}(\mathbb{R}_+),\ i = 0, 1, \tag{5}$$

where λ_0 and λ_1 are the arrival rates of jumps (i.e., the counting process N in (1)) before and after the disorder, respectively.

The main contribution of our paper is the *complete* solution of compound Poisson disorder problem *in its full generality.* For any pair of arrival rates λ_0 and λ_1 and mark distributions $\nu_0(\cdot)$ and $\nu_1(\cdot)$, we describe explicitly a quickest detection rule. These rules depend on the some $\mathbb{F}$-adapted *odds-ratio process* $\Phi = \{\Phi_t; t \geq 0\}$; see (11). At every $t \geq 0$, the random variable Φ_t is the conditional odds-ratio of the event $\{\theta \leq t\}$ that disorder has happened at or before time t given past and present observations $\mathcal{F}_t$ of the process X. For a suitable constant $\xi > 0$, the first crossing time $U_0 = \inf\{t \geq 0 : \Phi_t \geq \xi\}$ of the process Φ turns out to be a quickest detection rule: the Bayes risk R_{U_0} in (2) of U_0 is the smallest among all of the stopping times of the process X. The critical threshold ξ can be calculated numerically, and the quickest detection rule U_0 is suitable for online implementation since Φ_t, $t \geq 0$ can be updated by a recursive formula; see (13).

We also show that every compound Poisson disorder problem with one of "standard" Bayes risks in (4) can be solved in the same way.

Our probabilistic methods are different from the analytical methods of all previously cited work. The latter attacked Poisson disorder problems by studying analytical properties of related free-boundary integro-differential equations. Instead, we study very carefully sample-paths of the process Φ, which turn out to be piecewise deterministic and Markovian. General characterization of stopping times of jump processes allows us to approximate the minimum Bayes risk successively. This approximation is the key to our computational and theoretical results.

In the next section we give the precise description of compound Poisson disorder problem and show how to reduce it to an optimal stopping problem for a suitable Markov process. In Sections 3 and 4, we introduce successive approximations of the value function of the optimal stopping problem and establish key results for an efficient numerical method, which is presented in Section 5. We illustrate this method on several old and new examples and discuss briefly some extensions in Section 6. Finally, we establish in Section 7 the connection between our method and method of variational inequalities as applied to compound Poisson disorder problem. Appendix A contains some basic derivations and long proofs.

2. Model and problem description. Starting with a reference probability measure, we shall first construct a model containing all of the random elements of our problem with the correct probability laws.

Model. Let $(\Omega, \mathcal{F}, \mathbb{P}_0)$ be a probability space hosting the following *independent* stochastic elements:

(i) a standard Poisson process $N = \{N_t;\, t \geq 0\}$ with the arrival rate λ_0,

(ii) independent and identically distributed $\mathbb{R}^d$-valued random variables $Y_1, Y_2, \ldots$ with some common distribution $\nu_0(B) \triangleq \mathbb{P}_0\{Y_1 \in B\}$ for every set B in the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$ and $\nu_0(\{0\}) = 0$,

(iii) a random variable θ with the distribution

$$\mathbb{P}_0\{\theta = 0\} = \pi \in [0, 1) \quad \text{and} \quad \mathbb{P}_0\{\theta > t\} = (1 - \pi)e^{-\lambda t},\ t \geq 0,\ \lambda > 0. \tag{6}$$

Let $X = \{X_t;\, t \geq 0\}$ be the process defined by (1) with the jump times

$$\sigma_n \triangleq \inf\{t > \sigma_{n-1} : X_t \neq X_{t-}\}, \qquad n \geq 1 \ \ (\sigma_0 \equiv 0), \tag{7}$$

and $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ be the augmentation of its natural filtration $\sigma(X_s, s \leq t)$, $t \geq 0$ with $\mathbb{P}_0$-null sets. Then the process X is a $(\mathbb{P}_0, \mathbb{F})$-compound Poisson process with the arrival rate λ_0 and the jump distribution $\nu_0(\cdot)$.

Let $\lambda_1 > 0$ be a constant, and $\nu_1(\cdot)$ be a probability measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ *absolutely continuous* with respect to the distribution $\nu_0(\cdot)$. In general, every probability measure $\nu_1(\cdot)$ is the sum of two probability measures; one is singular, and the other is absolutely continuous with respect to $\nu_0(\cdot)$. If it is necessary, the distribution $\nu_1(\cdot)$ is replaced with its component which is absolutely continuous with respect to the measure $\nu_0(\cdot)$ *without loss of generality* as explained by Poor [14, pp. 269-271]. Then the Radon-Nikodym derivative

$$f(y) \triangleq \left.\frac{d\nu_1}{d\nu_0}\right|_{\mathcal{B}(\mathbb{R}^d)}(y), \quad y \in \mathbb{R}^d \tag{8}$$

of $\nu_1(\cdot)$ with respect to $\nu_0(\cdot)$ exists and is a ν_0-a.e. nonnegative Borel function.

We shall denote by $\mathbb{G} = \{\mathcal{G}_t\}_{t \geq 0}$ and $\mathcal{G}_t \triangleq \mathcal{F}_t \vee \sigma(\theta)$, $t \geq 0$ the enlargement of the filtration $\mathbb{F}$ with the sigma-algebra $\sigma(\theta)$ generated by θ. Let us define a *new probability measure* $\mathbb{P}$ on the measurable space $(\Omega, \vee_{s \geq 0}\mathcal{G}_s)$ locally in terms of the Radon-Nikodym derivatives

$$\left.\frac{d\mathbb{P}}{d\mathbb{P}_0}\right|_{\mathcal{G}_t} = Z_t \triangleq 1_{\{t < \theta\}} + 1_{\{t \geq \theta\}} e^{-(\lambda_1 - \lambda_0)(t - \theta)} \prod_{k = N_{\theta-}+1}^{N_t} \left[\frac{\lambda_1}{\lambda_0} f(Y_k)\right], \qquad t \geq 0, \tag{9}$$

where $N_{\theta-}$ is the number of arrivals in the time-interval $[0, \theta)$. If the disorder time θ is known, then each random variable Z_t is simply the likelihood ratio of the interarrival times $\sigma_1, \sigma_2 - \sigma_1, \ldots$ and the jump sizes $Y_1, Y_2, \ldots$ observed at or before time t. Under $\mathbb{P}$, the interarrival times and jump sizes are *conditionally* independent and have the desired conditional distributions *given* θ: the rate of exponentially distributed interarrival times and the distribution of the jump sizes change at time θ from λ_0 and $\nu_0(\cdot)$ to λ_1 and $\nu_1(\cdot)$, respectively. See also Appendix A.1 for another justification by using an absolutely continuous change of measure for point processes.

Finally, because $Z_0 = 1$ almost surely and the probability measures $\mathbb{P}_0$ and $\mathbb{P}$ coincide on $\mathcal{G}_0 = \sigma(\theta)$, the distribution of θ is the same under $\mathbb{P}$ and $\mathbb{P}_0$. Hence, under the probability measure $\mathbb{P}$ defined by (9), the process X and the random variable θ have the same properties as in the setup of the disorder problem described in the introduction.

Problem description. In the remainder, we shall work with the concrete model described above. The random variable θ is the unobservable disorder time and must be detected as quickly as possible as the history $\mathbb{F}$ of the observation process X is unfolded. The admissible detection rules are the stopping times of the filtration $\mathbb{F}$.

Our problem is to find the smallest Bayes risk $U(\cdot)$ in (3) by minimizing over all stopping rules τ of the filtration $\mathbb{F}$ the tradeoff $R_\tau(\cdot)$ in (2) between the false-alarm frequency and expected detection delay cost. If this infimum is attained, then we also want to describe explicitly a stopping rule with the minimum Bayes risk.

In the remainder of this section, we shall formulate the quickest-detection problem as an optimal stopping problem for a suitable Markov process; see (16) below. In later sections, we solve this optimal stopping problem completely and identify an optimal stopping rule.

One may check as in Bayraktar, Dayanik, and Karatzas [2, Proposition 2.1] that the Bayes risk in (2) can be expressed as

$$R_\tau(\pi) = (1-\pi) + c(1-\pi)\,\mathbb{E}_0\left[\int_0^\tau e^{-\lambda t}\left(\Phi_t - \frac{\lambda}{c}\right)dt\right], \qquad \pi\in[0,1),\ \tau\in\mathbb{F} \tag{10}$$

in terms of the $\mathbb{F}$-adapted *odds-ratio process*

$$\Phi_t \triangleq \frac{\mathbb{P}\{\theta\le t|\mathcal{F}_t\}}{\mathbb{P}\{\theta> t|\mathcal{F}_t\}}, \quad t\ge 0. \tag{11}$$

For every $t\ge 0$, the random variable Φ_t is the conditional odds-ratio of the event that the disorder happened at or before time t given the history $\mathcal{F}_t$ of the process X. In (10), the expectation $\mathbb{E}_0$ is taken with respect to $\mathbb{P}_0$, and the probability measure $\mathbb{P}$ in (11) is defined by the absolutely continuous change of measure in (9).

In Appendix A.2, we show that the process $\Phi=\{\Phi_t;\ t\ge 0\}$ in (11) is a *piecewise-deterministic Markov process* (Davis [6, 7]). If we define

$$\begin{aligned} &a \triangleq \lambda-\lambda_1+\lambda_0, \qquad \phi_d \triangleq \left\{\begin{matrix} -\lambda/a, & \text{if} \quad a\ne 0\\ -\infty, & \text{if} \quad a=0\end{matrix}\right\},\\ &x(t,\phi)\triangleq\left\{\begin{matrix}\phi_d+e^{at}\left[\phi-\phi_d\right], & a\ne 0\\ \phi+\lambda t, & a=0\end{matrix}\right\}, \qquad t\in\mathbb{R},\ \phi\in\mathbb{R}.\end{aligned} \tag{12}$$

and the σ_n, $n\ge 0$ are the jump times in (7) of the process X, then we get

$$\left\{\begin{matrix}\Phi_t = x\left(t-\sigma_{n-1},\Phi_{\sigma_{n-1}}\right), & t\in[\sigma_{n-1},\sigma_n)\\ \Phi_{\sigma_n} = \dfrac{\lambda_1}{\lambda_0}f(Y_n)\Phi_{\sigma_n-} & \end{matrix}\right\}, \quad n\ge 1. \tag{13}$$

Namely, the process Φ follows one of the deterministic curves $t\mapsto x(t,\phi)$, $\phi\in\mathbb{R}$ in (12) between consecutive jumps of X and is updated instantaneously at every jump of X as in (13); see also Figure 1 on page 9. The $(\mathbb{P}_0,\mathbb{F})$-infinitesimal generator of the process Φ coincides on the collection of continuously differentiable functions $h:\mathbb{R}_+\mapsto\mathbb{R}$ with the first-order integro-differential operator (see Appendix A.3)

$$\mathcal{A}h(x) = [\lambda+ax]\,h'(x) + \lambda_0\int_{y\in\mathbb{R}^d}\left[h\left(\frac{\lambda_1}{\lambda_0}f(y)\,x\right)-h(x)\right]\nu_0(dy), \qquad x\in\mathbb{R}_+. \tag{14}$$

Finally, the minimum Bayes risk in (3, 10) is given by

$$U(\pi) = (1-\pi)+c(1-\pi)\,V\left(\frac{\pi}{1-\pi}\right), \qquad \pi\in[0,1) \tag{15}$$

in terms of the value function

$$V(\phi)\triangleq\inf_{\tau\in\mathbb{F}}\ \mathbb{E}_0^\phi\left[\int_0^\tau e^{-\lambda t}g(\Phi_t)dt\right], \quad \phi\in\mathbb{R}_+ \tag{16}$$

of a discounted optimal stopping problem with the running cost

$$g(\phi)\triangleq\phi-\frac{\lambda}{c}, \qquad \phi\in\mathbb{R}_+ \tag{17}$$

and discount rate $\lambda>0$ for the piecewise-deterministic Markov process Φ in (13). In (16), the expectation $\mathbb{E}_0^\phi$ is taken with respect to the probability measure $\mathbb{P}_0$ and $\mathbb{P}_0\{\Phi_0=\phi\}=1$.

Thus, our problem becomes to calculate the value function $V(\cdot)$ in (16) and to find an optimal stopping rule if the infimum is attained. Our approach is *direct* and very suitable for piecewise-deterministic Markov processes. The solution is described in Section 5 in terms of *single-jump operators* after key results are established in Sections 3 and 4.

We adopt the direct approach instead of its widely-used alternative, namely, the *method of variational inequalities*. In the latter method, the value function $V(\cdot)$ in (16) is expected to satisfy the variational inequalities

$$\min\{(\mathcal{A}-\lambda)v(\phi)+g(\phi),-v(\phi)\}=0, \qquad \phi\in\mathbb{R}_+ \tag{18}$$

in some suitable sense and may be identified by solving (18) subject to certain boundary conditions. However, solving (18) is very difficult because of unfavorable analytical properties of the (singular) integro-differential operator $\mathcal{A}$ in (14). Our direct approach not only provides the complete solution of the original optimal stopping problem in (16), but also concludes as a by-product that $V(\cdot)$ is indeed the unique solution of the variational inequalities in (18); see Section 7.

3. A useful approximation and its single-jump analysis. Let us introduce the family of optimal stopping problems

$$V_n(\phi)\triangleq\inf_{\tau\in\mathbb{F}}\ \mathbb{E}_0^{\phi}\left[\int_0^{\tau\wedge\sigma_n}e^{-\lambda t}g(\Phi_t)dt\right], \qquad \phi\in\mathbb{R}_+,\ n\in\mathbb{N}, \tag{19}$$

obtained from (16) by stopping the odds-ratio process Φ at the nth jump time σ_n of the observation process X. Since the running cost $g(\cdot)$ in (17) is bounded from below by the constant $-\lambda/c$, the expectation in (19) is well-defined for every stopping time $\tau\in\mathbb{F}$. In fact, $-1/c\le V_n\le 0$ for every $n\in\mathbb{N}$. Since the sequence $(\sigma_n)_{n\ge1}$ of jump times of the process X is increasing almost surely, the sequence $(V_n)_{n\ge1}$ is decreasing. Therefore, $\lim_{n\to\infty}V_n$ exists everywhere. It is also obvious that $V_n\ge V$, $n\in\mathbb{N}$.

Proposition 3.1 *As $n\to\infty$, the sequence $V_n(\phi)$ converges to $V(\phi)$ uniformly in $\phi\in\mathbb{R}_+$. In fact, for every $n\in\mathbb{N}$ and $\phi\in\mathbb{R}_+$, we have*

$$-\frac{1}{c}\cdot\left(\frac{\lambda_0}{\lambda+\lambda_0}\right)^n\le V(\phi)-V_n(\phi)\le 0. \tag{20}$$

PROOF. Since $g(\phi)\ge-\lambda/c$ for every $\phi\ge0$, we have

$$\mathbb{E}_0^{\phi}\left[\int_0^{\tau}e^{-\lambda s}g(\Phi_s)ds\right]\ge\mathbb{E}_0^{\phi}\left[\int_0^{\tau\wedge\sigma_n}e^{-\lambda s}g(\Phi_s)ds\right]-\frac{1}{c}\cdot\mathbb{E}_0^{\phi}\left[e^{-\lambda\sigma_n}\right], \qquad \tau\in\mathbb{F},\ n\in\mathbb{N}.$$

Under $\mathbb{P}_0$, the nth jump time σ_n has Erlang distribution with parameters n and λ_0. Taking the infimum of both sides over $\tau\in\mathbb{F}$ gives the first inequality in (20). □

The uniform approximation in Proposition 3.1 is fast and accurate. On the other hand, the functions $V_n(\cdot)$ can be found easily by an iterative algorithm. We shall calculate the V_n's by adapting to our problem a method of Gugerli [11] and Davis [7, Chapter 5]. Developed for optimal stopping of general piecewise-deterministic Markov processes with an undiscounted terminal reward, the results of U. Gugerli and M. Davis do not apply here immediately. Since total discounted running cost over the infinite horizon has infinite expectation, an obvious transformation of our problem to those studied by U. Gugerli and M. Davis does not exist.

Let us start by defining the following operators acting on bounded Borel functions $w:\mathbb{R}_+\mapsto\mathbb{R}$:

$$Jw(t,\phi)\triangleq\mathbb{E}_0^{\phi}\left[\int_0^{t\wedge\sigma_1}e^{-\lambda u}g(\Phi_u)\,du+1_{\{t\ge\sigma_1\}}e^{-\lambda\sigma_1}w(\Phi_{\sigma_1})\right], \qquad t\in[0,\infty], \tag{21}$$

$$J_tw(\phi)\triangleq\inf_{u\in[t,\infty]}Jw(u,\phi), \qquad t\in[0,\infty]. \tag{22}$$

The special structure of the stopping times of jump processes (see Lemma A.1 below) implies

$$J_0w(\phi)=\inf_{\tau\in\mathbb{F}}\ \mathbb{E}_0^{\phi}\left[\int_0^{\tau\wedge\sigma_1}e^{-\lambda t}g(\Phi_t)\,dt+1_{\{\tau\ge\sigma_1\}}e^{-\lambda\sigma_1}w(\Phi_{\sigma_1})\right].$$

By relying on the strong Markov property of the process X at its first jump time σ_1, one expects that the value function V of (16) satisfies the equation $V=J_0V$. In Proposition 3.6 below, we show that this is indeed the case. In fact, if we define $v_n:\mathbb{R}_+\mapsto\mathbb{R}$, $n\in\mathbb{N}_0$, sequentially by

$$v_0\equiv0, \quad \text{and} \quad v_n\triangleq J_0v_{n-1} \quad \forall\,n\in\mathbb{N}, \tag{23}$$

then every v_n is bounded and identical to V_n of (19), $\lim_{n\to\infty} v_n$ exists and equals the value function V in (16); see Corollary 3.4 and Proposition 3.5.

Under $\mathbb{P}_0$, the first jump time σ_1 of the process X has exponential distribution with rate λ_0. Using the Fubini theorem and (13), we can write (21) as

$$Jw(t,\phi) = \int_0^t e^{-(\lambda+\lambda_0)u}\big(g+\lambda_0\cdot Sw\big)\big(x(u,\phi)\big)du, \quad t\in[0,\infty], \tag{24}$$

where the function $x(\cdot,\phi)$ is given by (12), and S is the linear operator

$$Sw(x) \triangleq \int_{\mathbb{R}^d} w\left(\frac{\lambda_1}{\lambda_0} f(y)\, x\right)\nu_0(dy), \qquad x\in\mathbb{R}, \tag{25}$$

defined on the collection of bounded functions $w:\mathbb{R}\mapsto\mathbb{R}$.

Remark 3.2 Using the explicit form of $x(u,\phi)$ in (12), it is easy to check that the integrand in (24) is absolutely integrable on $\mathbb{R}_+$. Therefore,

$$\lim_{t\to\infty} Jw(t,\phi) = Jw(\infty,\phi) < \infty,$$

and the mapping $t\mapsto Jw(t,\phi):[0,+\infty]\mapsto\mathbb{R}$ is continuous. Therefore, the infimum $J_t w(\phi)$ in (22) is attained for every $t\in[0,\infty]$.

Lemma 3.3 *For every bounded Borel function $w:\mathbb{R}_+\mapsto\mathbb{R}$, the mapping $J_0 w$ is bounded. If we define $||w|| \triangleq \sup_{\phi\in\mathbb{R}_+}|w(\phi)|<\infty$, then*

$$-\left(\frac{\lambda}{\lambda+\lambda_0}\cdot\frac{1}{c}+\frac{\lambda_0}{\lambda+\lambda_0}\cdot||w||\right)\le J_0 w(\phi)\le 0, \qquad \phi\in\mathbb{R}_+. \tag{26}$$

If the function $w(\cdot)$ is concave, then so is $J_0 w(\cdot)$. If $w_1(\cdot)\le w_2(\cdot)$ are real-valued and bounded Borel functions defined on $\mathbb{R}_+$, then $J_0 w_1(\cdot)\le J_0 w_2(\cdot)$. Namely, the operator J_0 preserves the boundedness, concavity, and monotonicity.

Proof. The lower bound in (26) follows from the lower bound $-\lambda/c$ on the running cost $g(\cdot)$ in (16). The concavity and the monotonicity can be checked directly. □

Corollary 3.4 *Every v_n, $n\in\mathbb{N}_0$ in (23) is bounded and concave, and $-1/c\le\ldots\le v_n\le v_{n-1}\le v_1\le v_0\equiv 0$. The limit*

$$v(\phi)\triangleq\lim_{n\to\infty} v_n(\phi), \quad \phi\in\mathbb{R}_+ \tag{27}$$

exists, and is bounded, concave, and nondecreasing. Both $v_n:\mathbb{R}_+\mapsto\mathbb{R}$, $n\in\mathbb{N}$ and $v:\mathbb{R}_+\mapsto\mathbb{R}$ are continuous and nondecreasing. Their left and right derivatives are bounded on every compact subset of $\mathbb{R}_+$.

Proof. By definition, $v_0\equiv 0$ is bounded, concave, and nondecreasing. By an induction argument on n, the conclusions follow from Lemma 3.3, the properties of concave functions, and the monotonicity of the functions $x(t,\cdot)$ for every fixed $t\in\mathbb{R}$ and $g(\cdot)$ in (12) and (17), and the operator S in (25). □

Next proposition describes some ε-optimal stopping rules for each problem in (19). In conjunction with Proposition 3.6 below, it is the basic block of the numerical scheme described in Section 5. Its proof is presented in Section A.4.

Proposition 3.5 *For every $n\in\mathbb{N}$, the functions v_n of (23) and V_n of (19) coincide. For every $\varepsilon\ge 0$, let*

$$r_n^\varepsilon(\phi)\triangleq\inf\left\{s\in(0,\infty]: Jv_n\big(s,\phi\big)\le J_0 v_n(\phi)+\varepsilon\right\}, \quad n\in\mathbb{N}_0,\ \phi\in\mathbb{R}_+,$$
$$S_1^\varepsilon\triangleq r_0^\varepsilon\big(\Phi_0\big)\wedge\sigma_1, \quad \text{and} \quad S_{n+1}^\varepsilon\triangleq\begin{Bmatrix} r_n^{\varepsilon/2}\big(\Phi_0\big), & \text{if } \sigma_1 > r_n^{\varepsilon/2}\big(\Phi_0\big) \\ \sigma_1+S_n^{\varepsilon/2}\circ\theta_{\sigma_1}, & \text{if } \sigma_1\le r_n^{\varepsilon/2}\big(\Phi_0\big)\end{Bmatrix}, \quad n\in\mathbb{N}, \tag{28}$$

where θ_s is the shift-operator on Ω: $X_t\circ\theta_s = X_{s+t}$. Then

$$\mathbb{E}_0^\phi\left[\int_0^{S_n^\varepsilon} e^{-\lambda t} g\big(\Phi_t\big)dt\right]\le v_n(\phi)+\varepsilon, \quad \forall\, n\in\mathbb{N},\ \forall\,\varepsilon\ge 0. \tag{29}$$

Proposition 3.6 *We have $v(\phi) = V(\phi)$ for every $\phi \in \mathbb{R}_+$. Moreover, V is the largest nonpositive solution U of the equation $U = J_0 U$.*

Proof. Corollary 3.4 and Propositions 3.5 and 3.1 imply that $v(\phi) = \lim_{n\to\infty} v_n(\phi) = \lim_{n\to\infty} V_n(\phi) = V(\phi)$ for every $\phi \in \mathbb{R}_+$. Next, let us show that $V = J_0 V$. Since $(v_n)_{n\geq 1}$ and $(Jv_n)_{n\geq 1}$ are decreasing, the bounded convergence theorem gives

$$V(\phi) = \lim_{n\to\infty} v_n(\phi) = \inf_{n\geq 1} J_0 v_{n-1}(\phi) = \inf_{t\in[0,\infty]} \lim_{n\to\infty} Jv_{n-1}(t,\phi) = \inf_{t\in[0,\infty]} Jv(t,\phi) = J_0 v(\phi).$$

If $U = J_0 U$ and $U \leq 0 \equiv v_0$, then repeated applications of J_0 to both sides of the last inequality and the monotonicity of J_0 (see Lemma 3.3) imply $U \leq V$. ☐

The next lemma and its immediate corollary below characterize the *smallest (deterministic) optimal stopping times* $r_n^0(\cdot)$, $n \in \mathbb{N}$ of Proposition 3.5 in a way familiar from the general theory of optimal stopping: $r_n^0(\phi)$ is the first time when the continuous path $t \mapsto x(t,\phi)$ enters the *stopping region* $\{x \in \mathbb{R}_+ : V_{n+1}(x) = 0\}$.

Lemma 3.7 *Let $w : \mathbb{R}_+ \mapsto \mathbb{R}$ be a bounded function. For every $t \in \mathbb{R}_+$ and $\phi \in \mathbb{R}_+$,*

$$J_t w(\phi) = Jw(t,\phi) + e^{-(\lambda+\lambda_0)t}\, J_0 w\big(x(t,\phi)\big). \tag{30}$$

Corollary 3.8 *Let*

$$r_n(\phi) = \inf\big\{s \in (0,\infty] : Jv_n\big(s,\phi\big) = J_0 v_n(\phi)\big\} \tag{31}$$

be the same as $r_n^\varepsilon(\phi)$ in Proposition 3.5 with $\varepsilon = 0$. Then

$$r_n(\phi) = \inf\big\{t > 0 : v_{n+1}\big(x(t,\phi)\big) = 0\big\} \qquad (\inf\varnothing \equiv \infty). \tag{32}$$

Remark 3.9 For every $t \in [0, r_n(\phi)]$, we have $J_t v_n(\phi) = J_0 v_n(\phi) = v_{n+1}(\phi)$. Then substituting $w(\cdot) = v_n(\cdot)$ in (30) gives the "dynamic programming equation" for the family $\{v_k(\cdot)\}_{k\in\mathbb{N}_0}$: for every $\phi \in \mathbb{R}_+$ and $n \in \mathbb{N}_0$

$$v_{n+1}(\phi) = Jv_n(t,\phi) + e^{-(\lambda+\lambda_0)t} v_{n+1}(x(t,\phi)), \quad t \in [0, r_n(\phi)].$$

Remark 3.10 Since $V(\cdot)$ is bounded, and $V = J_0 V$ by Proposition 3.6, Lemma 3.7 gives

$$J_t V(\phi) = JV(t,\phi) + e^{-(\lambda+\lambda_0)t}\, V\big(x(t,\phi)\big), \quad t \in \mathbb{R}_+ \tag{33}$$

for every $\phi \in \mathbb{R}_+$. If we define

$$r(\phi) \triangleq \inf\{t > 0 : JV(t,\phi) = J_0 V(\phi)\}, \qquad \phi \in \mathbb{R}_+,$$

then same arguments as in the proof of Corollary 3.8 with obvious changes and (33) give

$$r(\phi) = \inf\{t > 0 : V(x(t,\phi)) = 0\}, \qquad \phi \in \mathbb{R}_+, \tag{34}$$

$$V(\phi) = JV(t,\phi) + e^{-(\lambda+\lambda_0)t} V(x(t,\phi)), \quad t \in [0, r(\phi)]. \tag{35}$$

Let us define the $\mathbb{F}$-stopping times

$$U_\varepsilon \triangleq \inf\{t \geq 0 : V(\Phi_t) \geq -\varepsilon\}, \qquad \varepsilon \geq 0. \tag{36}$$

Next proposition shows that for the problem in (16) the stopping time $U_0 = \inf\{t \geq 0 : V(\Phi_t) = 0\}$ is optimal, and the stopping times U_ε in (36), $\varepsilon \geq 0$ are ε-optimal as in (37).

Proposition 3.11 *For every $\varepsilon \geq 0$, the stopping time U_ε in (36) is an ε-optimal stopping time for the optimal stopping problem (16), i.e.,*

$$\mathbb{E}_0^\phi\left[\int_0^{U_\varepsilon} e^{-\lambda s} g(\Phi_s)ds\right] \leq V(\phi) + \varepsilon, \quad \textit{for every} \quad \phi \in \mathbb{R}_+. \tag{37}$$

The proof in Section A.4 makes use of the local martingales described by the next proposition, which will be needed also in Section 7, where we show that the value function $V(\cdot)$ is the unique solution of variational equations in (18).

Proposition 3.12 *The process*

$$M_t \triangleq e^{-\lambda t} V(\Phi_t) + \int_0^t e^{-\lambda s} g(\Phi_s) ds, \qquad t \geq 0. \tag{38}$$

is a $(\mathbb{P}_0, \mathbb{F})$*-local martingale. For every* $n \in \mathbb{N}$*,* $\varepsilon \geq 0$*, and* $\phi \in \mathbb{R}_+$*, we have* $\mathbb{E}_0^\phi[M_0] = \mathbb{E}_0^\phi[M_{U_\varepsilon \wedge \sigma_n}]$*, i.e.,*

$$V(\phi) = \mathbb{E}_0^\phi \left[e^{-\lambda(U_\varepsilon \wedge \sigma_n)} V(\Phi_{U_\varepsilon \wedge \sigma_n}) + \int_0^{U_\varepsilon \wedge \sigma_n} e^{-\lambda s} g(\Phi_s) ds \right]. \tag{39}$$

4. Sample paths and bounds on the optimal alarm time. A brief study of sample paths of the sufficient statistic Φ in (11-13) gives simple lower and upper bounds on the optimal alarm time U_0 in (36). In several special cases, the lower bound becomes optimal. On the other hand, the upper bound has always finite Bayes risk.

Recall from Section 2 that the sufficient statistic Φ follows the deterministic curves $t \mapsto x(t, \phi), \phi \in \mathbb{R}_+$ in (12) when the observation process X does not jump. At every jump of the process X, the motion of Φ restarts on a different curve. Between jumps, the process Φ reverts to the mean-level ϕ_d if ϕ_d is positive, and grows unboundedly otherwise; see Figure 1. A jump at time t of the process Φ is in the forward direction if $f(Y_{N_t})(\lambda_1/\lambda_0) \geq 1$ and in the backward direction otherwise.

Since the running cost $g(\phi) = \phi - \lambda/c$ in (17) is negative on the interval $\phi \in [0, \lambda/c)$, the maximum $\tau \vee \underline{\tau}$ of any stopping rule τ and

$$\underline{\tau} \triangleq \inf\{t \geq 0 : \Phi_t \geq \lambda/c\} \tag{40}$$

gives a lower expected discounted total running cost than τ does:

$$\begin{aligned} \mathbb{E}_0^\phi \left[\int_0^{\tau \vee \underline{\tau}} e^{-\lambda t} g(\Phi_t) dt \right] &= \mathbb{E}_0^\phi \left[\int_0^{\tau} e^{-\lambda t} g(\Phi_t) dt \right] + \mathbb{E}_0^\phi \left[1_{\{\underline{\tau} > \tau\}} \int_\tau^{\underline{\tau}} e^{-\lambda t} g(\Phi_t) dt \right] \\ &\leq \mathbb{E}_0^\phi \left[\int_0^{\tau} e^{-\lambda t} g(\Phi_t) dt \right], \qquad \text{for every} \quad \phi \in \mathbb{R}_+. \end{aligned}$$

Therefore, the infimum in (16) can be taken over the stopping times $\{\tau \in \mathbb{F} : \tau \geq \underline{\tau}\}$ without any loss, and $\underline{\tau}$ in (40) is a *lower bound* on the optimal alarm time.

Proposition 4.1 *Suppose that* $f(y)(\lambda_1/\lambda_0) \geq 1$ *for every* $y \in \mathbb{R}^d$*. If* $\phi_d < 0$ *or* $0 < \lambda/c \leq \phi_d$ *in (12), then the stopping rule* $\underline{\tau}$ *of (40) is optimal for the problem (16).*

By Proposition 3.11, the stopping time $U_0 = \inf\{t \geq 0 : V(\Phi_t) = 0\}$ is always optimal for the problem (16). Next we show that U_0 is bounded almost surely between $\underline{\tau}$ in (40) and

$$\begin{aligned} &\overline{\tau} \triangleq \inf\{t \geq 0 : \Phi_t \geq \overline{\xi}\} \\ \text{with} \quad &\overline{\xi} \triangleq \max\left\{ \frac{\lambda + \lambda_0}{c}, \left[\frac{\lambda + \lambda_0}{c} - \phi_d \right] \left(\frac{\lambda_1}{\lambda + \lambda_0} \right) + \phi_d \right\} > \frac{\lambda}{c}. \end{aligned} \tag{41}$$

Proposition 4.2 *We always have* $U_0 \in [\underline{\tau}, \overline{\tau}]$ *almost surely and*

$$[\lambda/c, \infty) \supseteq \{\phi \in \mathbb{R}_+ : v_1(\phi) = 0\} \supseteq \{\phi \in \mathbb{R}_+ : V(\phi) = 0\} \supseteq [\overline{\xi}, \infty). \tag{42}$$

From (10), we find that the Bayes risk of the upper bound $\overline{\tau}$ in (41)

$$R_{\overline{\tau}}(\pi) = 1 - \pi + c(1 - \pi)\, \mathbb{E}_0 \left[\int_0^{\overline{\tau}} e^{-\lambda t} \left(\Phi_t - \frac{\lambda}{c} \right) dt \right] \leq 1 - \pi + c(1 - \pi) \left(\overline{\xi} - \frac{\lambda}{c} \right) \frac{1}{\lambda}$$

is finite. Since $\mathbb{E}[\overline{\tau}] \leq \mathbb{E}[(\overline{\tau} - \theta)^+] + \mathbb{E}[\theta] < (1/c) R_{\overline{\tau}}(\pi) + (1/\lambda) < \infty$, the stopping time $\overline{\tau}$ is finite $\mathbb{P}$-almost surely.

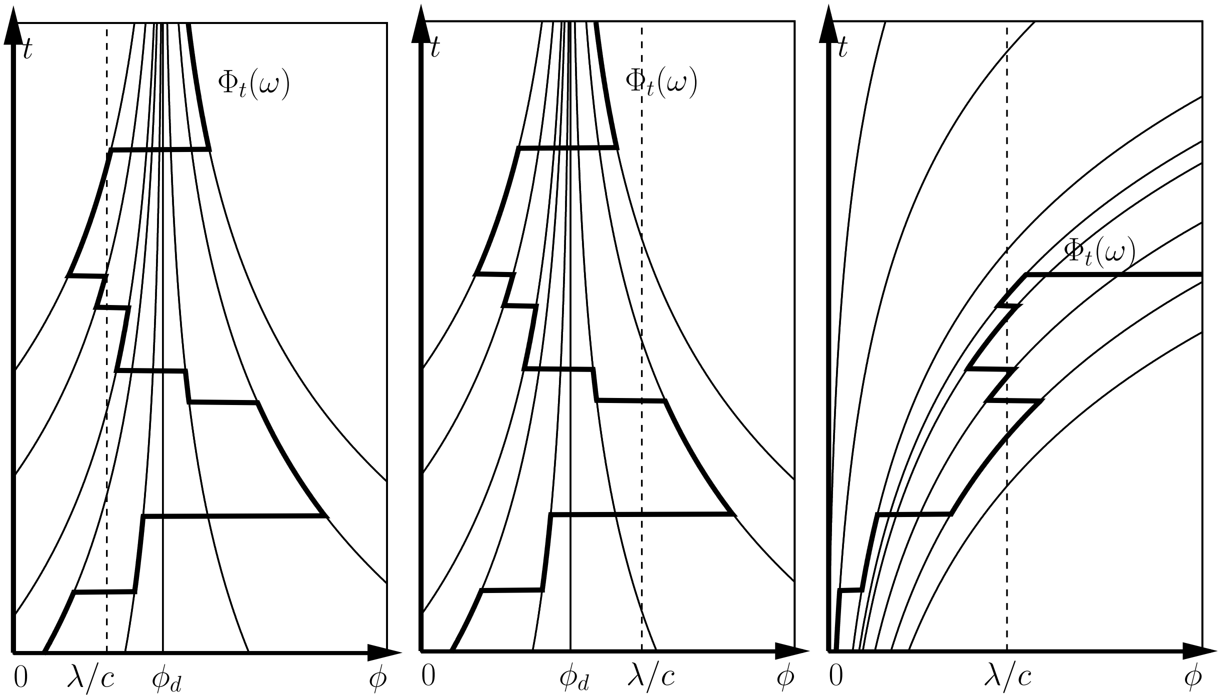

Figure 1: The sample paths of the process Φ in (11-13). If the quantity ϕ_d in (12) is positive, then it is the mean-reversion level for the process Φ: between successive jumps, the process reverts to the level ϕ_d as in (a). If however $\phi_d < 0$, then the process increases unboundedly between jumps as in (c). In general, the process Φ may jump in both directions in both cases (compare this with the sample paths of a similar statistic in the *standard Poisson disorder problem*; see Bayraktar, Dayanik, and Karatzas [2]).

5. The solution. By Propositions 3.1 and 3.5, the value function $V(\cdot)$ of the optimal stopping problem in (16) is approximated uniformly with a decreasing sequence of functions $\{v_n(\cdot)\}_{n\ge 0}$ defined sequentially by $v_0 \equiv 0$, and

$$v_{n+1}(\phi) = \inf_{t\in[0,\infty]} \quad Jv_n(t,\phi) = \int_0^t e^{-(\lambda+\lambda_0)u} \left[g + \lambda_0 \cdot Sv_n\right](x(u,\phi))du, \quad n \ge 0, \tag{43}$$

where S is the operator in (25). The sequence $\{v_n(\cdot)\}_{n\ge 1}$ converges to $V(\cdot)$ pointwise at an exponential rate, and the explicit bound in (20) determines the number n of iterations of (43) needed in order to achieve any desired accuracy: for any given $\varepsilon > 0$, we have

$$\frac{1}{c}\left(\frac{\lambda}{\lambda+\lambda_0}\right)^{n+1} < \varepsilon \qquad \Longrightarrow \qquad 0 \le V(\phi) - v_{n+1}(\phi) < \varepsilon \quad \text{for every } \phi \in \mathbb{R}_+. \tag{44}$$

For every integer $n+1$ as in (44), the stopping rule $S_{n+1} \equiv S^0_{n+1}$ of Proposition 3.5 is ε-optimal for the problem in (16):

$$0 \le V(\phi) - \mathbb{E}^{\phi}_0\left[\int_0^{S_{n+1}} e^{-\lambda t} g(\Phi_t)dt\right] < \varepsilon \qquad \text{for every } \phi \in \mathbb{R}_+.$$

The stopping time S_{n+1} is determined collectively by the jump times $\sigma_1, \ldots, \sigma_{n+1}$ of the observation process X and the *smallest* minimizers $r_n(\cdot), r_{n-1}(\cdot), \ldots, r_0(\cdot)$ of the deterministic optimization problems in (43); see (28) and (31): We wait until the earliest of the first jump at σ_1 and the time $r_n(\Phi_0)$. If $r_n(\Phi_0)$ occurs first, then we stop; otherwise, we reset the clock and continue to wait until the earliest of the next jump at $\sigma_2 - \sigma_1 = \sigma_1 \circ \theta_{\sigma_1}$ and the time $r_{n-1}(\Phi_{\sigma_1})$. If $r_{n-1}(\Phi_{\sigma_1})$ occurs first, then we stop; otherwise, we reset the clock and continue to wait until the earliest of the next jump at $\sigma_3 - \sigma_2 = \sigma_1 \circ \theta_{\sigma_2}$ and the time $r_{n-2}(\Phi_{\sigma_2})$, and so on. We stop at the $(n+1)$st jump time σ_{n+1} if we have not stopped yet.

The original definition of the time $r_n(\cdot)$, $n \geq 0$ in (28) obscures its simple meaning. Let us introduce the *stopping* and *continuation regions*,

$$\begin{bmatrix} \mathbf{\Gamma}_n \triangleq \{\phi \in \mathbb{R}_+ : v_n(\phi) = 0\}, \quad n \geq 1 \\ \mathbf{\Gamma} \triangleq \{\phi \in \mathbb{R}_+ : v(\phi) = 0\} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbf{C}_n \triangleq \mathbb{R}_+ \setminus \mathbf{\Gamma}_n, \quad n \geq 1 \\ \mathbf{C} \triangleq \mathbb{R}_+ \setminus \mathbf{\Gamma} \end{bmatrix}, \tag{45}$$

respectively. By Corollary 3.8, the deterministic time

$$r_n(\phi) = \inf\{t > 0 : x(t, \phi) \in \mathbf{\Gamma}_{n+1}\}, \qquad n \geq 0 \tag{46}$$

is the *first return time* of the continuous and deterministic path $t \mapsto x(t, \phi)$ in (12) to the stopping region $\mathbf{\Gamma}_{n+1}$.

Clearly, a concrete characterization of the stopping regions $\mathbf{\Gamma}_{n+1}$, $n \geq 0$ will ease the calculation of the return times $r_n(\cdot)$, $n \geq 0$ and an ε-optimal alarm time S_{n+1} as described above. Moreover, the function $v_{n+1}(\cdot)$ is already known on the set $\mathbf{\Gamma}_{n+1}$ (it equals zero identically), so the location and shape of the region $\mathbf{C}_{n+1} = \mathbb{R}_+ \setminus \mathbf{\Gamma}_{n+1}$ help a better implementation of (43). Since the sequence of nonpositive functions $\{v_n(\cdot)\}_{n \geq 0}$ decreases to $v(\cdot)$, Proposition 4.2 implies that

$$\begin{aligned} [\lambda/c, \infty) \supseteq \mathbf{\Gamma}_1 \supseteq \mathbf{\Gamma}_2 \supseteq \cdots \supseteq \mathbf{\Gamma}_{n+1} \supseteq \cdots \supseteq \mathbf{\Gamma} \supseteq [\overline{\xi}, \infty), \\ [0, \lambda/c) \subseteq \mathbf{C}_1 \subseteq \mathbf{C}_2 \subseteq \cdots \subseteq \mathbf{C}_{n+1} \subseteq \cdots \subseteq \mathbf{C} \subseteq [0, \overline{\xi}), \end{aligned} \tag{47}$$

where $\overline{\xi}$ is the explicit threshold in (41) for the upper bound on the optimal alarm time U_0. Therefore, the deterministic problems in (43) should be solved only for $\phi \in [0, \overline{\xi}]$. The *smallest* infimum $r_n(\phi)$ in (46) of the problem (43) is less than or equal to

$$\overline{r}_n(\phi) \triangleq \inf\{t > 0 : x(t, \phi) \geq \overline{\xi}\},$$

and the infimum in (43) may be taken only over the interval $t \in [0, \overline{r}_n(\phi)]$ without any loss. Let us define

$$\xi_n \triangleq \inf\{\phi \in \mathbb{R}_+ : v_n(\phi) = 0\}, \quad n \geq 1 \qquad \text{and} \qquad \xi \triangleq \inf\{\phi \in \mathbb{R}_+ : v(\phi) = 0\}. \tag{48}$$

Proposition 5.1 *We have* $\lambda/c \leq \xi_1 \leq \xi_2 \leq \cdots \leq \xi_n \leq \cdots \leq \xi \leq \overline{\xi}$, *and*

$$\mathbf{\Gamma}_n = [\xi_n, \infty), \quad n \geq 1 \qquad \text{and} \qquad \mathbf{\Gamma} = [\xi, \infty). \tag{49}$$

Moreover, $\xi_n \nearrow \xi$ *as* $n \to \infty$. *The functions* $v_n(\cdot)$, $n \geq 1$ *and* $v(\cdot)$ *are strictly increasing on* $\mathbf{C}_n = [0, \xi_n)$, $n \geq 1$ *and* $\mathbf{C} = [0, \xi)$, *respectively.*

Proof. By (47), we have $\lambda/c \leq \xi_n \leq \xi \leq \overline{\xi}$ for every $n \geq 1$, and the sequence $(\xi_n)_{n \geq 1}$ is increasing.

Since the nonpositive functions $v_n(\cdot)$, $n \geq 1$ and $v(\cdot)$ are increasing and continuous by Corollary 3.4, the identities in (49) follow. Because the functions are also concave, they are *strictly* increasing on the corresponding continuation regions.

Because $(\xi_n)_{n \geq 1}$ is increasing, we have $\xi \geq \xi^* \triangleq \lim_{n \to \infty} \xi_n \in \mathbf{\Gamma}_k$ and $v_k(\xi^*) = 0$ for every $k \geq 1$. Therefore, $v(\xi^*) = \lim_{k \to \infty} v_n(\xi^*) = 0$ and $\xi^* \in \mathbf{\Gamma}$, i.e., $\xi^* \geq \xi$. Hence $\xi = \xi^* \equiv \lim_{n \to \infty} \xi_n$. □

The structure of the problems in (43) helps to lay out a concrete iterative solution algorithm; see Figure 2. Suppose that $v_n(\cdot)$ is already calculated for some $n \geq 0$, and $v_{n+1}(\cdot)$ is the next. The infimum in (43) is not reached before the curve $t \mapsto x(t, \phi)$ leaves the region

$$A_n \triangleq \{\phi \in \mathbb{R}_+ : [g + \lambda_0 \cdot S v_n](\phi) < 0\} = [0, \alpha_n), \qquad n \geq 0, \tag{50}$$

where the boundary point

$$\alpha_n \triangleq \inf\{x \in \mathbb{R}_+ : [g + \lambda_0 \cdot S v_n](x) = 0\} \tag{51}$$

can be calculated immediately since $v_n(\cdot)$ is known. In (50), the identity $A_n = [0, \alpha_n)$ follows from that the mapping $x \mapsto [g + \lambda_0 \cdot S v_n](x) : \mathbb{R}_+ \mapsto \mathbb{R}$ is *strictly* increasing and continuous with limits $-\lambda/r + v_n(0) < 0$ and $+\infty$ as x goes to 0 and $+\infty$, respectively. Now the unknown boundary ξ_{n+1} of the continuation region $\mathbf{C}_{n+1} = [0, \xi_{n+1})$ and the function $v_{n+1}(\phi)$ for $\phi \in \mathbf{C}_{n+1}$ can be found from the relation between the known α_n in (51) and ϕ_d in (12):

Step 0 Set $n = 0$ and $v_0 \equiv 0$. Calculate $\overline{\xi}$ of (41).

Step 1 Find α_n of (51) by a bisection search in $[0, \overline{\xi}]$.

- If $\phi_d \notin [0, \alpha_n)$, then set ξ_{n+1} to α_n, and calculate on $\mathbb{R}_+$ the function

$$v_{n+1}(\phi) = \left\{ \begin{array}{ll} Jv_n(r_n(\phi), \phi), & \phi < \xi_{n+1} \\ 0, & \phi \geq \xi_{n+1} \end{array} \right\}, \qquad r_n(\phi) \equiv \left\{ \begin{array}{ll} \dfrac{1}{a} \ln \dfrac{\xi_{n+1} - \phi_d}{\phi - \phi_d}, & a \neq 0 \\ \dfrac{\xi_{n+1} - \phi}{\lambda}, & a = 0 \end{array} \right\}.$$

- If $\phi_d \in [0, \alpha_n)$, then set ξ_{n+1} to the unique root of the strictly increasing mapping $\phi \mapsto Jv_n(\infty, \phi)$ of (52). The root can be found by another bisection search in $[\alpha_n, \overline{\xi}]$. Calculate on $\mathbb{R}_+$ the function

$$v_{n+1}(\phi) = \left\{ \begin{array}{ll} Jv_n(\infty, \phi), & \phi < \xi_{n+1} \\ 0, & \phi \geq \xi_{n+1} \end{array} \right\}.$$

Step 2 For the $(n+1)$st problem in (19) set the stopping region $\mathbf{\Gamma}_{n+1}$ to $[\xi_{n+1}, \infty)$ and the value function V_{n+1} to v_{n+1}. Increase n by one and go to **Step 1**.

Figure 2: The solution of (16) by iterative approximations. In Step 2, the relation (44) may be used as a stopping rule to obtain arbitrarily close approximations $V_{n+1}(\cdot)$ for the value function $V(\cdot)$ of (16).

Case I: $\phi_d \notin [0, \alpha_n)$: the curve $t \mapsto x(t, \phi)$, $\phi \in \mathbb{R}_+$ leaves the interval $[0, \alpha_n)$ and never comes back; see (12) and Figure 1 on page 9 . Therefore, $\mathbf{C}_{n+1} = A_n$ (i.e., $\xi_{n+1} = \alpha_n$) and

$$v_{n+1}(\phi) = Jv_n(r_n(\phi), \phi) = \int_0^{r_n(\phi)} e^{-(\lambda+\lambda_0)u} \left[g + \lambda_0 \cdot Sv_n\right](x(u, \phi))du,$$

where $r_n(\phi)$ in (46) becomes the first exit time of $t \mapsto x(t, \phi)$ from $A_n = [0, \alpha_n)$.

Case II: $\phi_d \in [0, \alpha_n)$: as $t \to +\infty$, we have $x(t, \phi) \to \phi_d$ *monotonically.* Therefore, the infimum in (43) is attained at either $t = 0$ or $t = +\infty$. The continuous function

$$\phi \mapsto Jv_n(+\infty, \phi) = \int_0^\infty e^{-(\lambda+\lambda_0)t}[g + \lambda_0 \cdot Sv_n](x(t, \phi))dt : \mathbb{R}_+ \mapsto \mathbb{R} \tag{52}$$

is strictly increasing and $Jv_n(+\infty, \alpha_n) < 0 < \lim_{\phi \mapsto \infty} Jv_n(+\infty, \phi) = +\infty$. Therefore, the mapping $\phi \mapsto Jv_n(+\infty, \phi)$ has unique root, and this root is at $\xi_{n+1} > \alpha_n$, since $\min\{0, Jv_n(\infty, \phi)\} = v_{n+1}(\phi)$ is negative at $\phi \in [0, \xi_{n+1})$ and zero at $\phi \in [\xi_{n+1}, \infty)$. The algorithm is summarized in Figure 2. It is implemented to solve several numerical examples in Section 6.

We shall close this section with a summary of the discussions above. The following corollary will be needed later as we describe how smooth the value function $V(\cdot)$ is. Below (i) is proved while discussing **Case I** and **Case II** above. The proof of (ii) is very similar.

Corollary 5.2 *Recall that the continuation regions* $\{\mathbf{C}_n\}_{n\geq1}$ *and* $\mathbf{C}$*, the sets* $\{A_n\}_{n\geq1}$*, the numbers* $\{\xi_n\}_{n\geq1}$*,* ξ*,* $\{\alpha_n\}_{n\geq1}$*, and* α *are defined as in (45), (50), (48), and (51), respectively. Analogously, let us introduce*

$$\alpha \triangleq \inf\{x \in \mathbb{R}_+ : [g + \lambda_0 \cdot SV](x) = 0\},$$
$$A \triangleq \{\phi \in \mathbb{R}_+ : [g + \lambda_0 \cdot SV](\phi) < 0\} = [0, \alpha).$$

The identity $A = [0, \alpha)$ *follows from that the mapping* $x \mapsto [g+\lambda_0 \cdot SV](x) : \mathbb{R}_+ \mapsto \mathbb{R}$ *is* strictly *increasing and continuous with limits* $-\lambda/r + V(0) < 0$ *and* $+\infty$ *as* x *goes to* 0 *and* $+\infty$*, respectively. Moreover, the followings hold:*

(i) If $\phi_d \notin \mathbf{C}_{n+1} = [0, \xi_{n+1})$*, then* $\mathbf{C}_{n+1} = A_n = [0, \alpha_n)$ *and* $[g+\lambda_0 \cdot SV](\xi_{n+1}) = [g+\lambda_0 \cdot SV](\alpha_n) = 0$*. If* $\phi_d \in \mathbf{C}_{n+1} = [0, \xi_{n+1})$*, then* $A_n \subsetneqq \mathbf{C}_{n+1}$ *and*

$$v_{n+1}(\phi) = Jv_n(+\infty, \phi) = \int_0^\infty e^{-(\lambda+\lambda_0)t}[g + \lambda_0 \cdot Sv_n](x(t, \phi))dt, \quad \phi \in \mathbf{C}_{n+1}.$$

(ii) If $\phi_d \notin \mathbf{C} = [0, \xi)$, then $\mathbf{C} = A = [0, \alpha)$ and $[g + \lambda_0 \cdot SV](\xi) = [g + \lambda_0 \cdot SV](\alpha) = 0$. If $\phi_d \in \mathbf{C} = [0, \xi)$, then $A \subsetneqq \mathbf{C}$ and

$$V(\phi) = JV(+\infty, \phi) = \int_0^\infty e^{-(\lambda+\lambda_0)t}[g + \lambda_0 \cdot SV](x(t, \phi))dt, \quad \phi \in \mathbf{C}.$$

6. Examples and extensions. In Section 6.1, we provide numerical examples with discrete and absolutely continuous jump distributions.

The methods of previous sections apply to quickest detection problems with other "standard" Bayes risk measures. A few necessary minor changes are explained, and numerical examples are given in Section 6.2. Finally, we revisit in Section 6.4 Gapeev's [10] very special compound Poisson disorder problem.

6.1 Numerical examples. In the first example, jump sizes are *discrete*. The jump distributions before and after the disorder are

$$\nu_0 = \left(\frac{1}{15}, \frac{5}{15}, \frac{4}{15}, \frac{3}{15}, \frac{2}{15}\right) \quad \text{and} \quad \nu_1 = \left(\frac{2}{15}, \frac{3}{15}, \frac{4}{15}, \frac{5}{15}, \frac{1}{15}\right) \tag{53}$$

on the set $\{1, 2, 3, 4, 5\}$, respectively; see the upper left panel in Figure 3 on page 13. The jump distribution is right skewed before the disorder (histogram with heavy outline in the background) and left skewed after the disorder (histogram with filled bars in the foreground). The mode of the jump distribution increases after the disorder.

After having set the parameters c (cost per unit delay time), λ (disorder arrival rate), λ_0 (arrival rate of observations before the disorder), the quickest-detection problem has been solved for three different arrival rates λ_1 of observations after the disorder; see the upper panels (b)-(d) in Figure 3: (b) $\lambda_1 = \lambda_0/2$ (observations arrive at a lower rate after the disorder), (c) $\lambda_1 = \lambda_0$ (arrival rate does not change), and (d) $\lambda_1 = 2\lambda_0$ (observations arrive at a higher rate after the disorder).

In each panel (b)-(d) are the successive approximations $V_1(\cdot), V_2(\cdot), \ldots$ of the value function $V(\cdot)$ of (16) drawn. The successive approximations $V_1(\cdot), V_2(\cdot), \ldots$ are the same as the functions in (19) and are calculated iteratively by using the algorithm in Figure 2. The algorithm is terminated after 13, 14, and 17 iterations, respectively, for (b), (c), and (d), when the largest difference between most recent two approximations becomes negligible. The functions $V_{13}(\cdot)$ in (b), $V_{14}(\cdot)$ in (c), and $V_{17}(\cdot)$ in (d) are the approximations of $V(\cdot)$. In (b), the relation $V_{13}(\cdot) \approx V(\cdot)$ implies that the disorder time will be spotted as closely as possible by the arrival of the 13th observation with a negligible sacrifice from the optimal Bayes risk; see also (20). Similar conclusions are true in (c) and (d).

Given that everything else is the same, we expect that the minimum Bayes risk is smaller when pre- and post-disorder arrival rates of observations are different than when they are the same. Intuitively, if the arrival rates before and after the disorder are different, then the interarrival times between observations carry useful information for the quickest detection of the disorder time. In the light of the relation in (15) between the Bayes risk $U(\cdot)$ and the value function $V(\cdot)$, this intuitive remark is confirmed empirically by a comparison of the case (c) with (b) and (d). The value functions in cases (b) and (d) (where $\lambda_1 \neq \lambda_0$) are smaller than that in case (c) (where $\lambda_1 = \lambda_0$). The difference is more striking between (d) and (c) than between (b) and (c). This is perhaps because case (b) (unlike case (d)) is deprived of useful additional information about the jump-sizes due to slow arrival rate of observations after the disorder.

Finally, the rightmost vertical bar at the edge of each panel marks the critical threshold ξ in (49) which determines the optimal alarm time: declare an alarm as soon as the odds-ratio process Φ in (11-13) leaves the interval $[0, \xi)$.

In the second example, jump-size distributions before and after the disorder are *absolutely continuous*. Before the disorder, jump sizes are exponentially distributed with some rate μ. After the disorder, they have gamma distribution with scale parameter μ—the same as the rate of the exponential distribution. For three different shape parameters—2, 3, and 6, the quickest detection problem is solved; in Figure 3, see panel (e) for the comparisons of probability density functions and panels (f)-(h) for successive approximations $V_1(\cdot), V_2(\cdot), \ldots$ for each of three cases.

In all of the cases, the arrival rate of observations before and after the disorder is kept the same (i.e., $\lambda_0 = \lambda_1$); thus, only observed jump sizes contain useful information to detect quickly the disorder time.

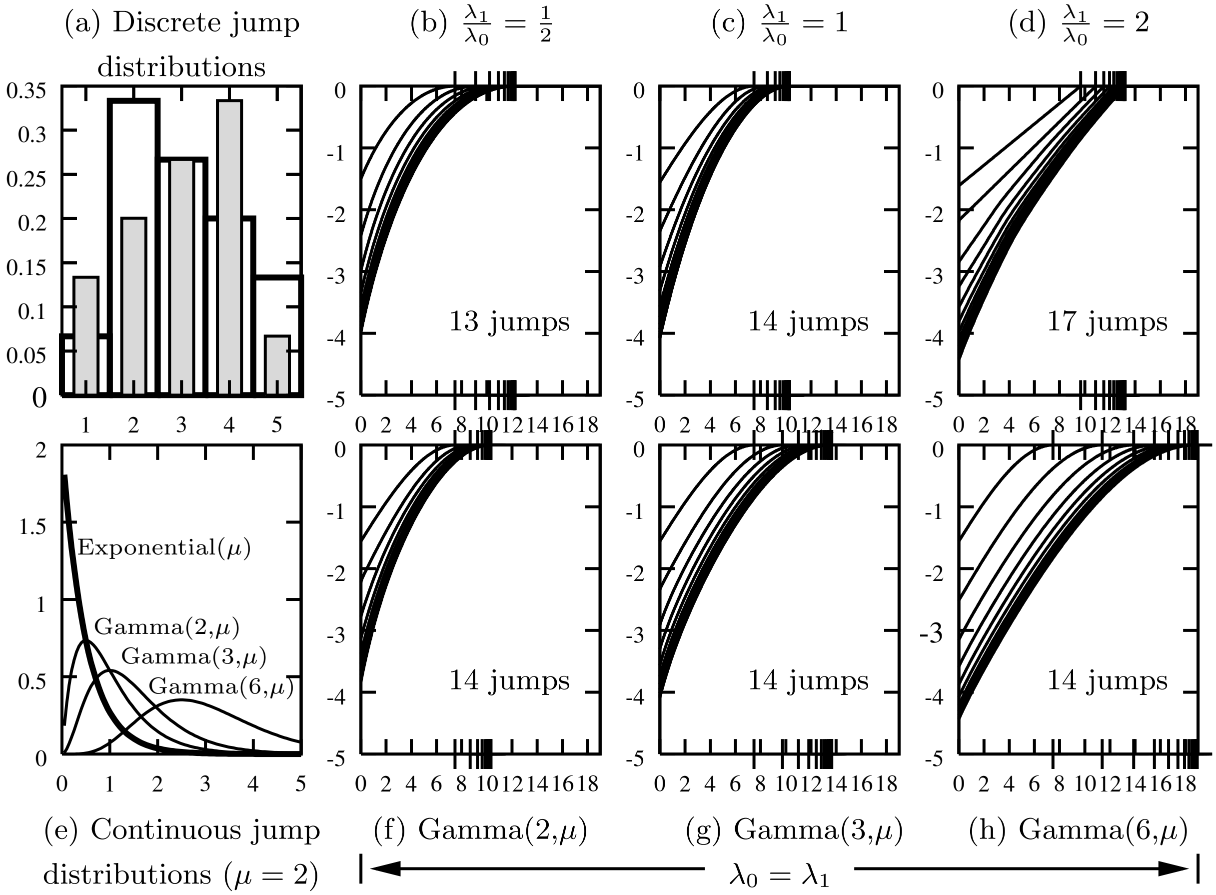

Figure 3: The solutions of the compound Poisson disorder problems with the Bayes risk in (2) ($c = 0.2$, $\lambda = 1.5$, $\lambda_0 = 3$). **Top row:** The jump distributions before and after the disorder are discrete. In (a), their probability mass functions are sketched (shaded is the post-disorder probability mass function). The number of iterations (jumps) and the successive approximations $v_n(\cdot)$ are reported when the ratio λ_1/λ_0 equals (b) 1/2, (c) 1, and (d) 2. **Bottom row:** Before the disorder, the jumps are exponentially distributed with rate $\mu = 2$. After the disorder, the jumps have (f) Gamma(2,μ), (g) Gamma(3,μ), and (h) Gamma(6,μ) distributions; see (e) for the sketches of their probability density functions. In all of the cases, $\lambda_1 = \lambda_0$. The optimal thresholds are indicated by the vertical bars at upper and lower edges of the panels; see also Figure 4.

Intuitively, if jump distributions before and after the disorder concentrate *more* on distinct/disjoint subsets, then the disorder can be spotted more accurately, and the Bayes risk becomes smaller. The numerical results (e)-(h) confirm our expectation. As the shape parameter increases, the post-disorder jump distribution shifts to the right—away from the pre-disorder jump distribution. At the same time, the value function $V(\cdot)$ (and the Bayes risk $U(\cdot)$ thanks to (15)) gets uniformly smaller.

6.2 Standard Poisson disorder problems. The Bayes risk of (2) is the second of four "standard" Bayes risks in (4). The risk measures in (4) are called "standard" by Bayraktar et. al. [2] following Davis [5] since they have essentially the same representation

$$\mathfrak{R}_\tau(\pi \,|\, \alpha, k, \gamma(\cdot), \beta(\cdot)\,) \triangleq \gamma(\pi) + \beta(\pi)\mathbb{E}_0\left[\int_0^\tau e^{-\lambda t}\Big(\Phi_t^{(\alpha)} - k\Big)\right] dt, \qquad \pi \in [0,1) \tag{54}$$

for some known constants $\alpha \geq 0$, $k > 0$ and functions $\gamma(\cdot), \beta(\cdot)$ from $[0,1)$ into $\mathbb{R}_+$. The *generalized* odds-ratio process

$$\Phi_t^{(\alpha)} \triangleq \frac{\mathbb{E}\left[e^{\alpha(t-\theta)} 1_{\{\theta \leq t\}} | \mathcal{F}_t\right]}{\mathbb{P}\{\theta > t | \mathcal{F}_t\}}, \qquad t \geq 0,\ \alpha \geq 0 \tag{55}$$

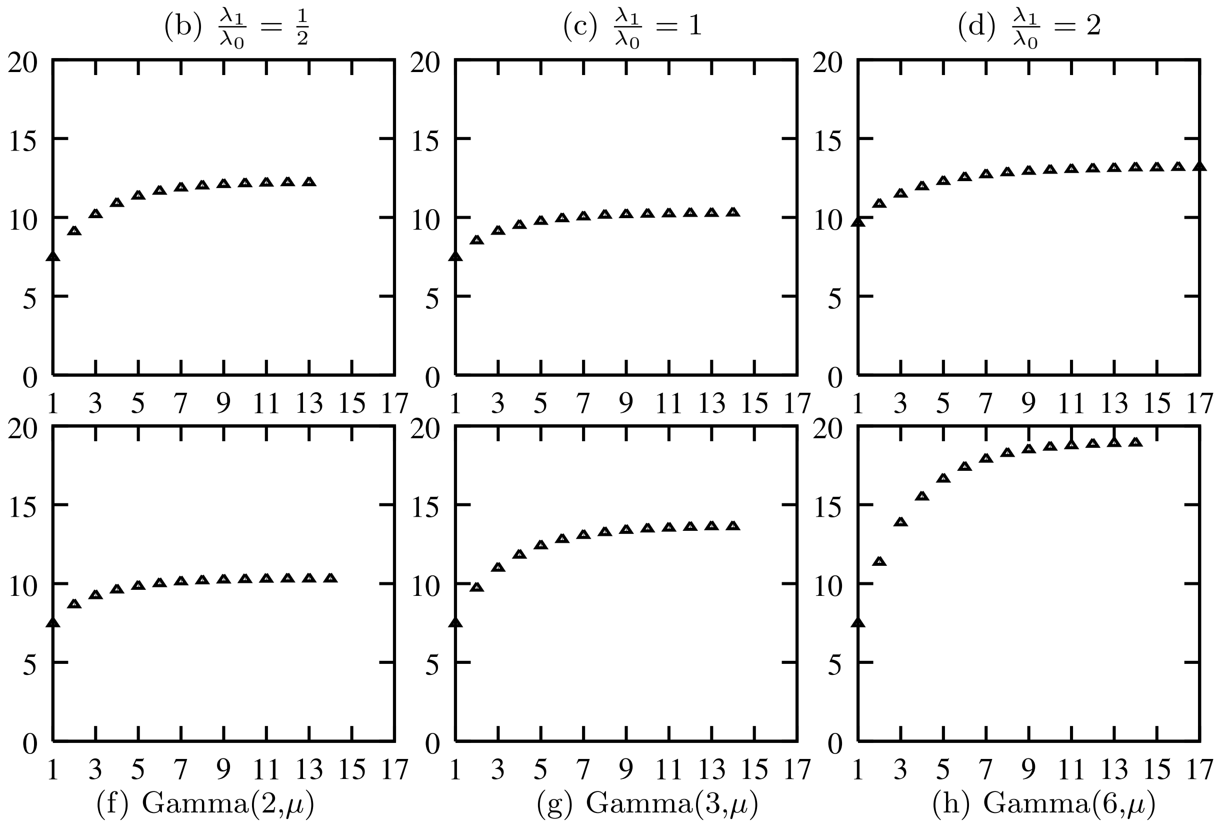

Figure 4: The critical thresholds ξ_n, $n = 1, 2, \ldots$ for the compound Poisson disorder problems considered in Figure 3

becomes the same as the odds-ratio process Φ_t, $t \geq 0$ in (11) when $\alpha = 0$. If we redefine the parameter a in (12) by

$$a \triangleq \lambda + \alpha - \lambda_1 + \lambda_0,$$

then the process $\Phi^{(\alpha)} = \{\Phi^{(\alpha)}_t; t \geq 0\}$ has the same dynamics as in (13) for every $\alpha \geq 0$:

$$\left\{ \begin{array}{c} \Phi^{(\alpha)}_t = x\left(t - \sigma_{n-1}, \Phi^{(\alpha)}_{\sigma_{n-1}}\right), \quad t \in [\sigma_{n-1}, \sigma_n) \\ \Phi^{(\alpha)}_{\sigma_n} = \dfrac{\lambda_1}{\lambda_0} f(Y_n) \Phi^{(\alpha)}_{\sigma_n -} \end{array} \right\}, \quad n \geq 1.$$

See Bayraktar et. al. [2, Proposition 2.1] for the proof of the following result.

Proposition 6.1 *For every $\pi \in [0, 1)$ and stopping time $\tau \in \mathbb{F}$, we have*

$$R^{(i)}_\tau(\pi) = \mathfrak{R}_\tau(\pi | \alpha_i, k_i, \gamma_i(\cdot), \beta_i(\cdot)), \qquad \textit{for every } i = 1, 2, 3, 4, \tag{56}$$

where $\alpha_1 = \alpha_2 = \alpha_3 = 0$, $\alpha_4 = \alpha$; $k_1 = (\lambda/c)e^{-\varepsilon\lambda}$, $k_2 = \lambda/c$, $k_3 = 1/c$, $k_4 = \lambda/(c\alpha)$; and

$$\gamma_1(\pi) = (1-\pi)e^{-\lambda\varepsilon}, \qquad \gamma_2(\pi) = 1 - \pi, \qquad \gamma_3(\pi) = \frac{1-\pi}{\lambda}, \qquad \gamma_4(\pi) = 1 - \pi$$
$$\beta_1(\pi) = c(1-\pi), \qquad \beta_2(\pi) = c(1-\pi), \qquad \beta_3(\pi) = c(1-\pi), \qquad \beta_4(\pi) = c\alpha(1-\pi).$$

For $i = 2$ the identity in (54, 56) is the same as the representation (10) which was the key for the solution. Therefore, the solution of the compound Poisson disorder problem with any "standard" Bayes risk in (4,56) remains the same after a few obvious changes.

The minimum Bayes risk $U(\pi) = \inf_{\tau \in \mathbb{F}} \mathfrak{R}(\pi | \alpha, k, \gamma(\cdot), \beta(\cdot))$, $\pi \in [0, 1)$ is given by

$$U(\pi) = \gamma(\pi) + \beta(\pi)\, V\left(\frac{\pi}{1-\pi}\right), \qquad \pi \in [0, 1)$$

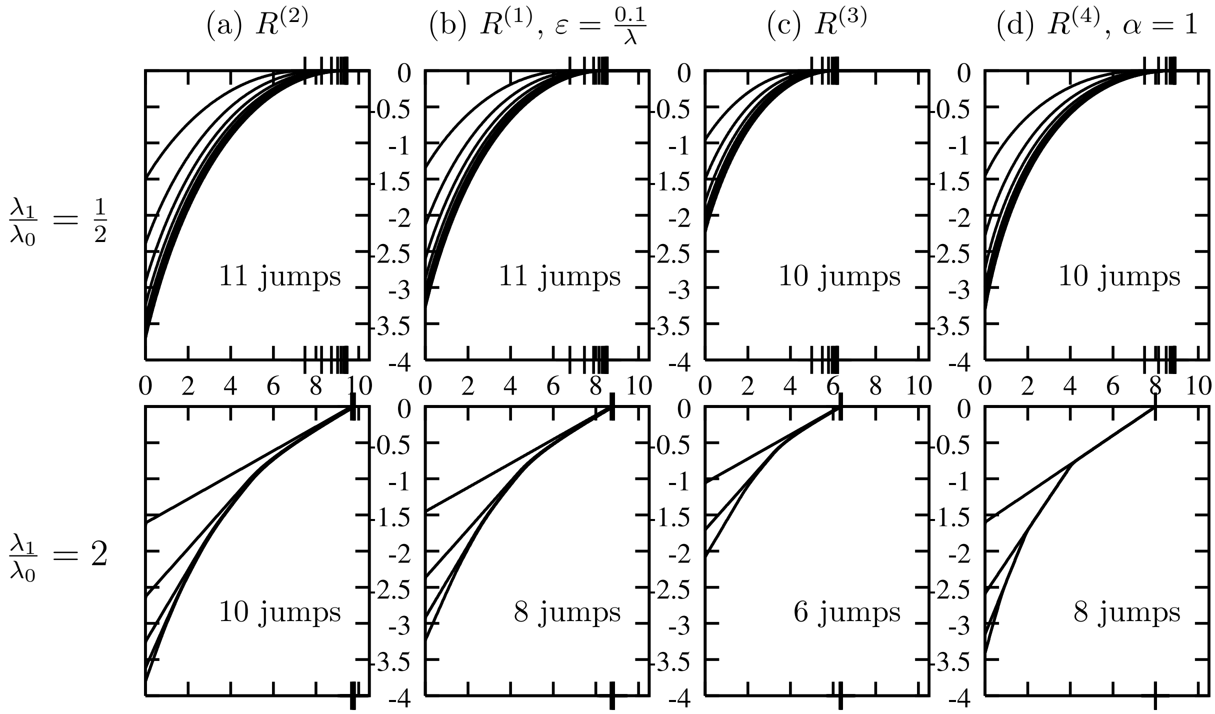

Figure 5: As in Bayraktar, Dayanik, and Karatzas [2], we take $c = 0.2$, $\lambda = 1.5$, $\lambda_0 = 3$ and $\nu_0(\cdot) \equiv \nu_1(\cdot) \equiv \delta_{\{1\}}(\cdot)$. For each case, the rate λ_1 is determined according to the ratio λ_1/λ_0 at the beginning of the same row. In every column, the disorder problem is solved for one of four penalties (linear, ε, expected miss, and exponential). The number of iterations (jumps) before convergence and the successive approximations $v_n(\cdot)$ of the value function $V(\cdot)$ are displayed for eight cases. In every case, the optimal threshold ξ_n for each subproblem $v_n(\cdot)$ is indicated by a vertical bar on both top and bottom edges of the panels; see also Figure 6.

in terms of the value function

$$V(\phi) \triangleq \inf_{\tau \in \mathbb{F}} \mathbb{E}_0^{\phi} \left[\int_0^{\tau} e^{-\lambda t} g(\Phi_t^{(\alpha)}) dt \right], \quad \phi \in \mathbb{R}_+$$

of a discounted optimal stopping problem with the running cost

$$g(\phi) \triangleq \phi - k, \qquad \phi \in \mathbb{R}_+$$

and discount rate $\lambda > 0$ for the piecewise-deterministic Markov process $\Phi^{(\alpha)}$ in (55). The successive approximations $\{V_n(\cdot)\}_{n \geq 1}$ in (19) of the value function $V(\cdot)$ are uniformly decreasing; and since $g(\cdot) \geq -k$, we have

$$-\frac{k}{\lambda} \cdot \left(\frac{\lambda_0}{\lambda + \lambda_0} \right)^n \leq V(\phi) - V_n(\phi) \leq 0.$$

The results of Sections 3-5 remain valid in this general case.

Figure 5 illustrates solutions of some Poisson disorder problem for each of four "standard" Bayes risk measures in (4). For comparison the parameters are chosen the same as in Bayraktar et. al. [2, Table 1], whose methods are unable to detect the change in the jump-size distribution, and therefore, can only use the count data on the number of arrivals to detect the disorder. On the other hand, the method of Sections 3-5 can be told to ignore completely the jump-size information (and to use number of arrivals only) by setting the density function $f(\cdot)$ in (8) and (13) identically to one (more precisely, the jump-distributions $\nu_0(\cdot)$ and $\nu_1(\cdot)$ are replaced with the Dirac measure $\delta_{\{1\}}(\cdot)$ at one on $\mathbb{R}_+$, so that the process X is the same as the counting process N in (1)). In Figure 5, the rightmost vertical bars at the edge of panels mark the critical thresholds of the quickest alarm rules and agree with those reported by Bayraktar et. al. [2, Table 1].

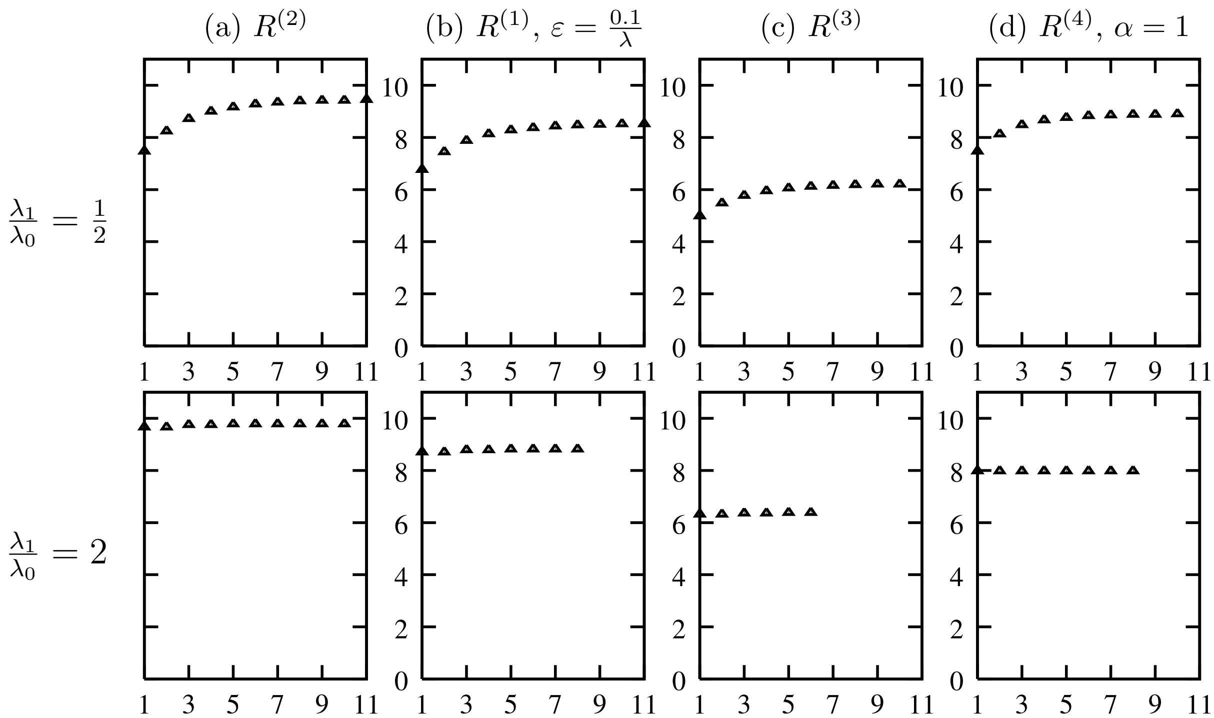

Figure 6: The critical thresholds ξ_n, $n = 1, 2, \ldots$ for the standard Poisson disorder problems considered in Figure 5.

6.3 Reducing the Bayes-risk by observing marks in addition to arrival times. Suppose that in the examples (b)-(d) of Figure 3 the observations of marks are unavailable, and one has to use only the data on the arrival times in order to detect the disorder time. How do the optimal Bayes risks and optimal strategies differ?

For different values of the ratio λ_1/λ_0, the value function of (16) is calculated in the presence and the absence of the mark data and displayed in the first row of Figure 7. In the absence of the mark data, compound Poisson disorder problem reduces to standard Poisson disorder problem, and the solutions of the latter are recalled from Figure 5(a) for $\lambda_1/\lambda_0 = 1/2$ and 2.

If $\lambda_1/\lambda_0 = 1$, and the mark data is absent, then (i) the sufficient statistic Φ in (11-13) becomes the *increasing* deterministic process

$$\Phi_t = x(t, \Phi_0) = -1 + e^{\lambda t}[\Phi_0 + 1], \quad t \geq 0 \qquad (\lambda_0 = \lambda_1,\ f(\cdot) \equiv 1), \tag{57}$$

(ii) following from (16, 17, 19)), the optimal thresholds in (49) become $\xi_1 = \xi_2 = \ldots = \xi = \lambda/c$, (iii) the optimal alarm time $t^*(\Phi_0) = \inf\{t \geq 0 : \Phi_t \geq \lambda/c\}$ is also deterministic,

$$t^*(\Phi_0) = \left[\frac{1}{\lambda}\ln\left(\frac{1+(\lambda/c)}{1+\Phi_0}\right)\right]^+ \quad \text{and} \quad V(\phi) = \frac{1+\phi}{\lambda}\left[1 + \ln\left(\frac{1+(\lambda/c)}{1+\phi}\right)\right] - \frac{\lambda+c}{c\lambda}.$$

The latter expression is used to draw the graph in Figure 7(b) of the value function $V(\cdot)$ of (16) corresponding to the case without mark observations.

The first row of Figure 7 shows that the reduction in the Bayes risk obtained by using the observations of the marks (in addition to those of the arrival times) can be significant. Moreover, this reduction tends to grow as the number of arrivals (hence the additional information carried by the accompanying mark data) increases with the increasing rate λ_1 for fixed λ_0. Finally, observe from (57) that arrival times carry no information about the disorder time if the arrival rate is not expected to change (i.e., $\lambda_0 = \lambda_1$), and the observations of marks become more crucial for early detection of the disorder and for lower Bayes risk; see Figure 7(b).

Since every stopping time of the arrival process $N = \{N_t;\, t \geq 0\}$ in (1) is also a stopping time of $X = \{X_t;\, t \geq 0\}$, the value function $V(\cdot)$ of (19) is always at least as small in the presence of

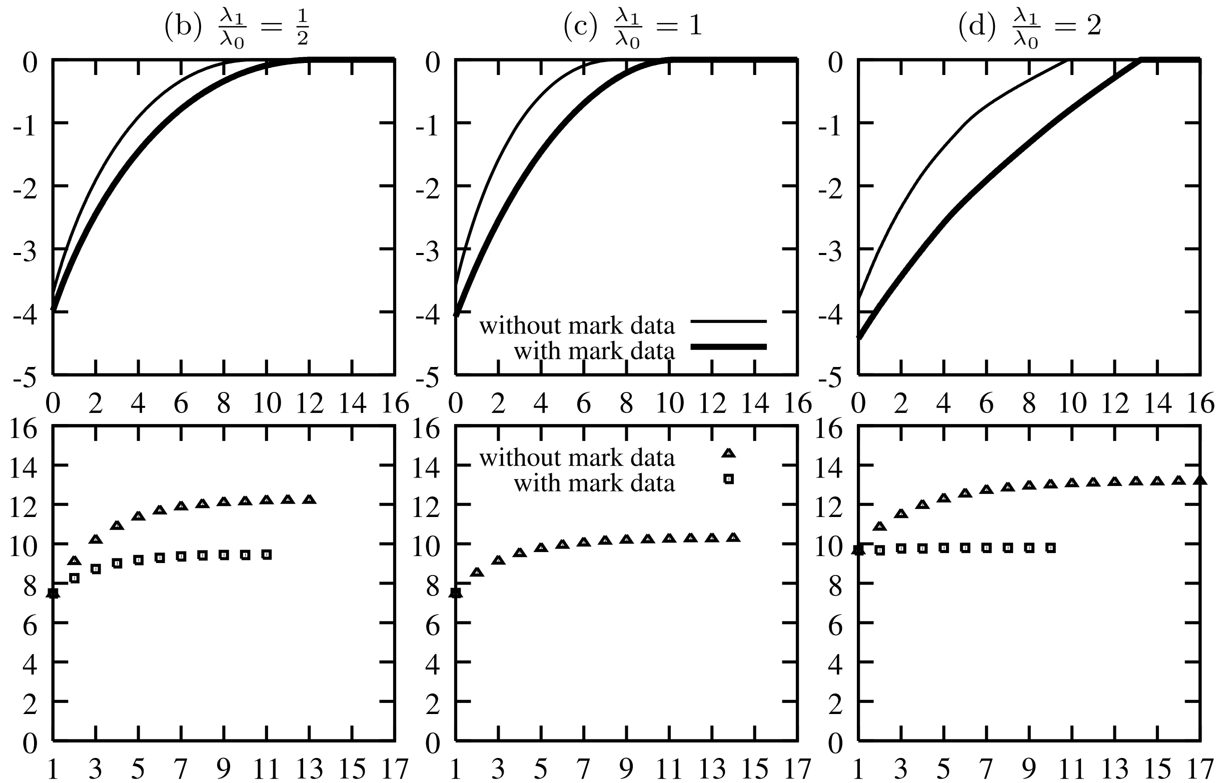

Figure 7: In the presence and the absence of mark observations, the value function $V(\cdot)$ of (16) and the thresholds $\{\xi_n;\, n = 1, 2, \ldots\}$ of (48) (until the termination of the algorithm in Figure 2) are displayed, respectively, in the first and second rows. The data are the same as those of Figure 3 (b), (c), and (d): $c = 0.2$, $\lambda = 1.5$, $\lambda_0 = 3$, and the discrete mark distributions $\nu_0(\cdot)$, $\nu_1(\cdot)$ are as in (53).

mark observations as the same function in the absence of mark observations. Therefore, the thresholds $\{\xi_n; n = 1, 2, \ldots\}$ and ξ in (48) are always at least as large in the presence of mark observations as those in the absence of mark observations. This fact is confirmed by the illustrations in the second row of Figure 7, where the thresholds $\{\xi_n; n = 1, 2, \ldots\}$ are displayed for each case before the algorithm in Figure 2 terminates. Note that this fact does not imply that an optimal alarm in the presence of mark observations is given always earlier than that in the absence of mark observations: not only the critical thresholds ξ but also the dynamics of the sufficient statistic Φ in (12, 13) are different in the presence (i.e., nontrivial $f(\cdot)$) and in the absence (i.e., $f(\cdot) \equiv 1$) of the mark observations. Therefore, the relation between optimal alarm times is not obvious.

6.4 Compound Poisson disorder problem with exponential jumps. Gapeev [10] recently solved fully a very special compound Poisson disorder problem: before and after the disorder, the jump sizes are exponentially distributed, and their common expected values are the same as the arrival rates of jumps in corresponding regimes. Namely, jump-size distributions are as in (5), and the Radon-Nikodym derivative in (8) becomes

$$f(y) = \frac{d\nu_1}{d\nu_0}\bigg|_{\mathcal{B}(\mathbb{R}_+)} = \frac{\lambda_0}{\lambda_1} \exp\left\{ -\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_0} \right) y \right\}. \tag{58}$$

Below are Gapeev's [10, Theorem 4.1] conclusions obtained by using general methods of this paper.

If $\lambda_0 < \lambda_1$ and $-a \equiv \lambda_1 - \lambda_0 - \lambda \le c$, then $f(\cdot)(\lambda_1/\lambda_0) \ge 1$ and either $\phi_d < 0$ or $0 < \lambda/c \le \phi_d$. Therefore, Proposition 4.1 applies, and the stopping time $\underline{\tau}$ in (40) is optimal. P. Gapeev works with the posterior probability process

$$\Pi_t \triangleq \mathbb{P}\{\theta \le t | \mathcal{F}_t\} \equiv \frac{\Phi_t}{1 + \Phi_t}, \qquad t \ge 0,$$

and the optimal stopping rule $\underline{\tau}$ can be rewritten as

$$\underline{\tau} = \inf\left\{t \geq 0 : \Pi_t \geq \frac{\lambda}{\lambda + c}\right\}.$$

If either "$\lambda_0 < \lambda_1$ and $-a > c$" or $\lambda_0 > \lambda_1$, then the stopping rule $U_0 = \inf\{t \geq 0 : \Phi_t \geq \xi\} = \inf\{t \geq 0 : \Pi_t \geq \xi/(1+\xi)\}$ in (36, 45, 48, 49) is optimal by Propositions 3.11 and 5.1. If $\lambda_0 > \lambda_1$, then $\phi_d < 0$ and the value function $V(\cdot)$ in (16) is continuously differentiable on $\mathbb{R}_+$ by Lemma 7.1 below, and $V'(\xi) = 0$.

7. Differentiability and variational inequalities. In this final section, smoothness of the value function $V(\cdot)$ in (16) is studied. The function $V(\cdot)$ is shown to be piecewise continuously differentiable and unique bounded solution of the variational inequalities in (18); see Lemma 7.1 and Proposition 7.3 on pages 20 and 21, respectively.

7.1 Differentiability of the value function. Since $V(\cdot) \equiv 0$ on the stopping region $\mathbf{\Gamma} = [\xi, \infty)$ by (45, 48, 49), it is obviously continuously differentiable on (ξ, ∞). Its smoothness on $[0, \xi]$ is investigated below separately in two cases due to different behavior of functions $t \mapsto x(t, \phi)$, $\phi \in \mathbb{R}_+$ of (12) for $\phi_d \notin (0, \xi]$ and $\phi_d \in (0, \xi]$. We summarize our conclusions in Lemma 7.1.

In both case, it will be very useful to recall from Remark 3.10 and (34, 35, 45, 49) that the value function $V(\cdot)$ satisfies some form of dynamic programming equation; namely,

$$\begin{aligned} V(\phi) &= JV(t, \phi) + e^{-(\lambda+\lambda_0)t} V(x(t, \phi)), \quad t \in [0, r(\phi)], \\ r(\phi) &= \inf\{t > 0 : x(t, \phi) \geq \xi\}, \qquad \phi \in \mathbb{R}_+ \end{aligned} \tag{59}$$

Case I: $\boldsymbol{\phi_d \notin (0, \xi]}$. Let us fix some $\phi \in [0, \xi)$ and define for every $0 < h < \xi - \phi$ that

$$T(h, \phi) \triangleq \inf\{t \geq 0 : x(t, \phi) \geq \phi + h\} = \begin{cases} \dfrac{1}{a} \cdot \ln\left(\dfrac{\phi + h - \phi_d}{\phi - \phi_d}\right), & a \neq 0 \\ h/\lambda, & a = 0 \end{cases}.$$

The second equality follows from (12). Because $T(h, \phi) \leq r(\phi)$, replacing $T(h, \phi)$ with t in (59) gives

$$V(\phi) = \int_0^{T(h,\phi)} e^{-(\lambda+\lambda_0)u} \left[g + \lambda_0 \cdot SV\right](x(u, \phi))du + e^{-(\lambda+\lambda_0)T(h,\phi)} V(\phi + h). \tag{60}$$

Subtracting $V(\phi + h)$ from each side and dividing by $-1/h$ give

$$\frac{V(\phi + h) - V(\phi)}{h} = -\frac{1}{h} \int_0^{T(h,\phi)} e^{-(\lambda+\lambda_0)u} \left[g + \lambda_0 \cdot SV\right](x(u, \phi))du - \frac{1}{h}\left[e^{-(\lambda+\lambda_0)T(h,\phi)} - 1\right] V(\phi + h).$$

Since $V(\cdot)$ is concave by Corollary 3.4 and Proposition 3.6, it has right derivatives everywhere. As h decreases to 0, we obtain

$$\lim_{h \to 0^+} \frac{V(\phi + h) - V(\phi)}{h} = -\Big(g(\phi) + \lambda_0 \cdot SV(\phi) - (\lambda + \lambda_0)V(\phi)\Big) \cdot \left(\left.\frac{\partial T(h, \phi)}{\partial h}\right|_{h=0}\right), \tag{61}$$

since the functions $V(\cdot)$ and $SV(\cdot)$ are bounded and continuous (by bounded convergence theorem). Because

$$\left.\frac{\partial T(h, \phi)}{\partial h}\right|_{h=0} = \begin{cases} 1/[a(\phi - \phi_d)], & a \neq 0 \\ 1/\lambda, & a = 0 \end{cases} = \frac{1}{\lambda + a\phi} \tag{62}$$

is a continuous function of $\phi \in [0, \xi)$ (recall that $\phi_d \notin [0, \xi)$, so the denominator is bounded away from zero on $\phi \in [0, \xi)$), the right derivative of $V(\cdot)$ in (61) is continuous on $\phi \in [0, \xi)$. Since $V(\cdot)$ is concave, this implies that $V(\cdot)$ is continuously differentiable on $[0, \xi)$, and (61, 62) give the derivative

$$V'(\phi) = \frac{g(\phi) + \lambda_0 \cdot SV(\phi) - (\lambda + \lambda_0)V(\phi)}{\lambda + a\phi}, \qquad \phi \in [0, \xi). \tag{63}$$

Finally, $V'(\xi-) = 0 = V'(\xi+)$ since $V(\cdot) \equiv 0$ on $[\xi, \infty)$ and $[g + \lambda_0 \cdot SV](\xi) = 0$ because of Corollary 5.2(ii) and $\phi_d \notin [0, \xi)$. The concavity of $V(\cdot)$ implies again that $V'(\xi)$ exists and equals zero. Hence the function $V(\cdot)$ is continuously differentiable everywhere on $\mathbb{R}_+$ if $\phi_d \notin [0, \xi)$.

Case II: $\phi_d \in (0, \xi]$. For every $\phi \in \mathbb{R}_+$, the function $x(t,\phi)$ converges monotonically to ϕ_d as t increases to infinity. For every $\phi \in [0,\xi)$, we have $V(\phi) = JV(\infty,\phi)$ by Corollary 5.2(ii).

If we redefine $T(h,\phi) \triangleq \inf\{t \geq 0 : |x(t,\phi) - \phi| \geq h\}$ for every $\phi \in [0,\xi)$ and $h > 0$, then the same arguments as in the previous case show that $V(\cdot)$ is continuously differentiable (with the same derivative $V'(\phi)$ as in (63)) on $\phi \in [0,\xi) \setminus \{\phi_d\}$.

Let us show now that the function $V(\phi)$ is *not differentiable* at $\phi = \xi$. In terms of

$$W(\phi) \triangleq [g + \lambda_0 \cdot SV](\phi), \qquad \phi \in \mathbb{R}_+,$$

one can write using Corollary 5.2(ii) that

$$\frac{V(\xi) - V(\xi - h)}{h} = \int_0^\infty e^{-(\lambda+\lambda_0)u} \left[\frac{W\big(x(u,\xi)\big) - W\big(x(u,\xi-h)\big)}{x(u,\xi) - x(u,\xi-h)} \right] \cdot \left[\frac{x(u,\xi) - x(u,\xi-h)}{h} \right] du.$$

Since the functions $g(\cdot)$ and $V(\cdot)$ are increasing, so are $SV(\cdot)$ of (25) and $W(\cdot)$. Therefore,

$$\frac{W\big(x(u,\xi)\big) - W\big(x(u,\xi-h)\big)}{x(u,\xi) - x(u,\xi-h)} \geq 0 \quad \text{and} \quad \frac{x(u,\xi) - x(u,\xi-h)}{h} = e^{au};$$

and Fatou's Lemma gives

$$\varliminf_{h\to 0^+} \frac{V(\xi) - V(\xi-h)}{h} \geq \int_0^\infty e^{-\lambda_1 u} \left[1 + \lambda_0 \varliminf_{h\to 0^+} \frac{SV\big(x(u,\xi)\big) - SV\big(x(u,\xi-h)\big)}{x(u,\xi) - x(u,\xi-h)} \right] du.$$

Since $SV(\cdot)$ is increasing, the limit infimum above is non-negative and

$$\varliminf_{h\to 0^+} \frac{V(\xi) - V(\xi-h)}{h} \geq \int_0^\infty e^{-\lambda_1 u} du = \frac{1}{\lambda_1} > 0 = \lim_{h\to 0^+} \frac{V(\xi+h) - V(\xi)}{h}. \tag{64}$$

Hence, the lefthand and righthand derivatives of $V(\cdot)$ are unequal at $\phi = \xi$, and $V(\cdot)$ is not differentiable at $\phi = \xi$.

On the other hand, the function $V(\cdot)$ may or may not be differentiable at $\phi = \phi_d$. Since $x(t,\phi_d) = \phi_d$ for every $t \geq 0$ by (12), Corollary 5.2(ii) gives

$$\frac{V(\phi_d + h) - V(\phi_d)}{h} = \frac{1}{\lambda_1} + \lambda_0 \int_0^\infty e^{-(\lambda+\lambda_0)u} \frac{SV(\phi_d + e^{au}h) - SV(\phi_d)}{h} du.$$

Because $SV(\cdot)$ is nondecreasing, Fatou's lemma gives

$$\varliminf_{h\to 0^+} \frac{V(\phi_d + h) - V(\phi_d)}{h} \geq \frac{1}{\lambda_1} + \lambda_0 \int_0^\infty e^{-(\lambda+\lambda_0)u} \Big[\varliminf_{h\to 0^+} \frac{SV(\phi_d + e^{au}h) - SV(\phi_d)}{h} \Big] du. \tag{65}$$

We shall calculate the limit infimum on the righthand side. In terms of the sets

$$A \triangleq \Big\{ y \in \mathbb{R}^d; \quad f(y)\frac{\lambda_1}{\lambda_0} = 1 \Big\} \qquad \text{and} \qquad B \triangleq \Big\{ y \in \mathbb{R}^d; \quad f(y)\frac{\lambda_1}{\lambda_0}\phi_d = \xi \Big\}, \tag{66}$$

the definition in (25) of $SV(\cdot)$ implies

$$\begin{aligned} \frac{SV(\phi_d + e^{au}h) - SV(\phi_d)}{h} &= \int_{\mathbb{R}^d \setminus (A\cup B)} \nu_0(dy) \left[\frac{V\big(f(y)\frac{\lambda_1}{\lambda_0}(\phi_d + e^{au}h)\big) - V\big(f(y)\frac{\lambda_1}{\lambda_0}\phi_d\big)}{h} \right] \\ &\quad + \int_A \nu_0(dy) \left[\frac{V(\phi_d + e^{au}h) - V(\phi_d)}{h} \right] + \int_B \nu_0(dy) \left[\frac{V(\xi + (\xi/\phi_d)e^{au}h) - V(\xi)}{h} \right]. \end{aligned}$$

The last integral is equal to 0 because $V(\phi) = 0$ for every $\phi \geq \xi$. Since the concave and increasing function $V(\cdot)$ has bounded right derivatives by Corollary 3.4 and is continuously differentiable on $\mathbb{R}_+ \setminus \{\phi_d, \xi\}$, the dominated convergence theorem implies that

$$\begin{aligned} \varliminf_{h\to 0^+} \frac{SV(\phi_d + e^{au}h) - SV(\phi_d)}{h} &= e^{au}\frac{\lambda_1}{\lambda_0} \int_{\mathbb{R}^d \setminus (A\cup B)} \nu_1(dy) V'\Big(f(y)\frac{\lambda_1}{\lambda_0}\phi_d \Big) \\ &\quad + e^{au}\nu_0(A). \varliminf_{h\to 0^+} \frac{V(\phi_d + h) - V(\phi_d)}{h} \qquad \text{for every } u \in \mathbb{R}_+. \end{aligned} \tag{67}$$

Inside the integral above, we used the relation $f(y)\nu_0(dy) = \nu_1(dy)$. After plugging (67) into (65), rearrangement of the terms gives

$$\left[\underline{\lim}_{h\to 0^+} \frac{V(\phi_d+h)-V(\phi_d)}{h}\right]\cdot\left(1-\frac{\lambda_0}{\lambda_1}\nu_0(A)\right) \geq \frac{1}{\lambda_1} + \int_{\mathbb{R}^d\setminus(A\cup B)} \nu_1(dy) V'\left(f(y)\frac{\lambda_1}{\lambda_0}\,\phi_d\right).$$

Similar arguments will also give

$$\left[\overline{\lim}_{h\to 0^+} \frac{V(\phi_d+h)-V(\phi_d)}{h}\right]\cdot\left(1-\frac{\lambda_0}{\lambda_1}\nu_0(A)\right) \leq \frac{1}{\lambda_1} + \int_{\mathbb{R}^d\setminus(A\cup B)} \nu_1(dy) V'\left(f(y)\frac{\lambda_1}{\lambda_0}\,\phi_d\right).$$

Since $\phi_d \in (0,\xi]$ in Case II, we have $a<0$, $\lambda_0<\lambda_1$, and $[1-(\lambda_0/\lambda_1)\nu_0(A)]>0$. By the last two displayed inequalities, the righthand derivative $D^+V(\phi)$ of $V(\cdot)$ at $\phi=\phi_d$ becomes

$$D^+V(\phi_d) = \left(1-\frac{\lambda_0}{\lambda_1}\nu_0(A)\right)^{-1}\left[\frac{1}{\lambda_1} + \int_{\mathbb{R}^d\setminus(A\cup B)} \nu_1(dy) V'\Big(f(y)\frac{\lambda_1}{\lambda_0}\,\phi_d\Big)\right].$$

By following the same arguments, one can show that the lefthand derivative $D^-V(\phi)$ of $V(\cdot)$ at $\phi=\phi_d$ becomes

$$D^-V(\phi_d) = D^+V(\phi_d) + \left(1-\frac{\lambda_0}{\lambda_1}\nu_0(A)\right)^{-1}\left[\frac{\lambda_0\xi}{\lambda_1\phi_d}\,\nu_0(B)\,D^-V(\xi)\right].$$

Since the derivative $D^-V(\xi)$ on the right does not vanish by (64), this equality shows that the value function $V(\cdot)$ is differentiable at $\phi=\phi_d$ (i.e., $D^-V(\phi_d)=D^+V(\phi_d)$) if and only if $\nu_0(B)=0$ for the set B defined in (66). Next lemma summarizes main conclusions.

Lemma 7.1 *Recall from (45, 48, 49) that the optimal continuation region for the problem (16) is in the form of* $\mathbf{C}=[0,\xi)$ *for some* $\xi>0$*, and the constant* ϕ_d *is given by (12).*

(i) If $\phi_d \notin \mathbf{C}=[0,\xi)$*, then the value function* $V(\cdot)$ *in (16) is continuously differentiable on* $\mathbb{R}_+$*.*

(ii) If $\phi_d \in \mathbf{C}=[0,\xi)$*, then* $V(\cdot)$ *is continuously differentiable on* $\mathbb{R}_+\setminus\{\phi_d,\xi\}$*. It is not differentiable at* ξ*. It is differentiable at* ϕ_d *if and only if*

$$\nu_0\left(\left\{y\in\mathbb{R}^d;\quad f(y)\frac{\lambda_1}{\lambda_0}\,\phi_d=\xi\right\}\right)=0.$$

Remark 7.2 If $\phi_d\in\mathbf{C}$, then $V(\phi_d)<0$. The local martingale in (38) and optional sampling imply that

$$V(\phi_d) = \mathbb{E}_0^{\phi_d}\left[\int_0^{\sigma_1} e^{-\lambda t} g(\Phi)dt + e^{\lambda\sigma_1}V(\Phi_{\sigma_1})\right] = \frac{g(\phi_d)}{\lambda+\lambda_0} + \frac{\lambda_0}{\lambda+\lambda_0}\cdot SV(\phi_d),$$

since the process Φ does not leave ϕ_d until the first jump time σ_1 if it starts initially at ϕ_d. This relation of $SV(\phi_d)$ of (25) to $V(\phi_d)$ and Lemma 7.1(ii) suggest that the lack of smoothness of $V(\cdot)$ at ϕ_d can occur if and only if this "ill" behavior can be "transmitted" from ξ. Alternatively, the function $V(\cdot)$ is not differentiable at ϕ_d if and only if the process Φ may jump before the disorder from ϕ_d to ξ with positive probability.

7.2 Unique solution of variational inequalities. The value function $V(\cdot)$ satisfies the variational inequalities in (18) wherever $V(\cdot)$ is differentiable. By Proposition 5.1,

$$V<0 \quad \text{on} \quad \mathbf{C}=[0,\xi) \qquad\qquad \text{and} \qquad\qquad V=0 \quad \text{on} \quad \mathbf{\Gamma}=\mathbb{R}_+\setminus\mathbf{C}. \tag{68}$$

The function $V(\cdot)$ is piecewise continuously differentiable by Lemma 7.1. The derivative $V'(\cdot)$ exists and equals zero on the stopping region $\mathbf{\Gamma}$. Since $V(\cdot)\equiv 0$ on $\mathbf{\Gamma}$ and $A\triangleq\{x\in\mathbb{R}_+ : [g+\lambda_0\cdot SV](x)<0\}\subseteq\mathbf{C}$ by Lemma 5.2, we have

$$(\mathcal{A}-\lambda)V(\phi)+g(\phi) = [g+\lambda_0\cdot SV](\phi)\geq 0 \qquad \text{on} \quad \phi\in\mathbf{\Gamma}=[\xi,\infty).$$

The above inequality is strict in the interior of $\mathbf{\Gamma}$ because $\phi\mapsto[g+\lambda_0\cdot SV](\phi)$ is strictly increasing. At every point $\phi\in\mathbf{C}$ where the derivative $V'(\phi)$ exists (see Lemma 7.1 above), it is given by (63) which can be rearranged as

$$\begin{aligned}
0 &= (\lambda+a\phi)V'(\phi)+\lambda_0 SV(\phi)-(\lambda+\lambda_0)V(\phi)+g(\phi)\\
&= [\lambda+a\phi]\,V'(x)+\lambda_0\int_{y\in\mathbb{R}^d}\left[V\left(\frac{\lambda_1}{\lambda_0}f(y)\,\phi\right)-V(\phi)\right]\nu_0(dy)-\lambda V(\phi)+g(\phi)\\
&= (\mathcal{A}-\lambda)V(\phi)+g(\phi).
\end{aligned}\tag{69}$$

It is easy to see from (68-69) that $V(\cdot)$ satisfies the variational inequalities in (18) wherever the derivative $V'(\cdot)$ exists. Next result shows that $V(\cdot)$ is the unique piecewise continuously differentiable bounded solution of (18).

Proposition 7.3 *Suppose that $U : \mathbb{R}_+ \mapsto \mathbb{R}$ is a continuous and bounded function which is continuously differentiable except at most finitely many points and satisfies (18) everywhere except at those points. Then $U = V$ on $\mathbb{R}_+$.*

Proof. Let T be any $\mathbb{F}$-stopping time and $t \geq 0$ be any constant. As in Appendix A.3,

$$e^{-\lambda(T\wedge t)}U(\Phi_{T\wedge t}) - U(\Phi_0) = \int_0^{T\wedge t} e^{-\lambda s}(\mathcal{A}-\lambda)U(\Phi_{s-})ds + \int_{(0,T\wedge t]\times\mathbb{R}^d} e^{-\lambda s}\left[U\left(\frac{\lambda_1}{\lambda_0}f(y)\Phi_{s-}\right) - U(\Phi_{s-})\right] q_0(dsdy).$$

Since $U(\cdot)$ is bounded, the integrand of the last integral on the right is absolutely integrable with respect to the $(\mathbb{P}_0, \mathbb{F})$-compensator measure $p_0(dsdy) = \lambda_0 ds\, \nu_0(dy)$. Therefore, the last integral on the right is a martingale, and its $\mathbb{P}_0$-expectation equals zero. Taking the $\mathbb{P}_0$-expectations of both sides and using the inequality $(\mathcal{A}-\lambda)U + g \geq 0$ give

$$U(\phi) \leq \mathbb{E}_0^{\phi}\left[e^{-\lambda(T\wedge t)}U(\Phi_{T\wedge t})\right] + \mathbb{E}_0^{\phi}\left[\int_0^{T\wedge t} e^{-\lambda s}g(\Phi_s)ds\right]. \tag{70}$$

Since $U(\cdot)$ is bounded and $g(\cdot) + \lambda/c \geq 0$, the bounded convergence and monotone convergence theorems give $U(\phi) \leq \mathbb{E}_0^{\phi}\left[\int_0^T e^{-\lambda s} g(\Phi_s)\right]$ when we take limit of both sides as t goes to infinity. Since $\mathbb{F}$-stopping time T is arbitrary, this implies $U \leq V$.

For the opposite inequality, let $T_* \triangleq \inf\{t \geq 0 : U(\Phi_t) = 0\}$. Then $[(\mathcal{A}-\lambda)U + g](\Phi_s) = 0$ for $s < T_*$, and (70) holds with equality. When we take the limits as before, we obtain $U(\phi) = \mathbb{E}_0^{\phi}\left[\int_0^{T_*} e^{-\lambda s} g(\Phi_s)\right] \geq V(\phi)$ for every $\phi \in \mathbb{R}_+$. □

Acknowledgments. We are grateful to the referee, the associate and area editors for their detailed suggestions, which improved our presentation. The work is supported by the National Science Foundation, under grant NSF-DMI-04-23327.

Appendix.

A.1 Absolutely continuous change of measure. The process X in (1) can also be expressed as the integral

$$X_t = X_0 + \int_{(0,t]\times\mathbb{R}^d} y\, p(dsdy), \qquad t \geq 0 \tag{A.71}$$

with the respect to the point process

$$p((0,t]\times A) \triangleq \sum_{k=1}^{\infty} 1_{\{\sigma_k \leq t\}} 1_{\{Y_k \in A\}}, \qquad t \geq 0,\ A \in \mathcal{B}(\mathbb{R}^d) \tag{A.72}$$

on $\left(\mathbb{R}_+ \times \mathbb{R}^d, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}^d)\right)$. Let $\mathbb{P}_0$ be the probability measure described in Section 2 and define

$$h(t,y) \triangleq 1_{\{t<\theta\}} + 1_{\{t\geq\theta\}}\frac{\lambda_1}{\lambda_0}f(y), \qquad t \in \mathbb{R}_+,\ y \in \mathbb{R}^d.$$

Since θ is $\mathcal{G}_0$-measurable, the process $\{h(t,y);\ t \geq 0\}$ is $\mathbb{G}$-predictable for every $y \in \mathbb{R}^d$. Therefore, the process

$$Z_t \triangleq \exp\left\{\int_{(0,t]\times\mathbb{R}^d} [\ln h(s,y)] p(dsdy) - \int_{(0,t]\times\mathbb{R}^d} [h(s,y)-1]\lambda_0\, ds\, \nu_0(dy))\right\}, \qquad t \geq 0$$

is a $(\mathbb{P}_0, \mathbb{G})$-martingale and induces a new probability measure $\mathbb{P}$ on the measurable space $(\Omega, \vee_{s\geq 0}\mathcal{G}_s)$ in terms of the Radon-Nikodym derivatives (9). The exponential formula for Z_t above also simplifies to that in (9). Girsanov theorem for the point processes (Jacod and Shiryaev [12, Chapter III], Cont and Tankov [4, p. 305]) guarantees that, under the new probability measure $\mathbb{P}$, the process X has the desired finite-dimensional distribution described in the introduction.

A.2 The dynamics of the odds-ratio process Φ in (11). The Radon-Nikodym derivative Z_t in (9) of the restriction to $\mathcal{G}_t$ of the probability measure $\mathbb{P}$ with respect to that of $\mathbb{P}_0$ can be written as

$$Z_t = 1_{\{\theta>t\}} + 1_{\{\theta\le t\}}\frac{L_t}{L_\theta}, \qquad \text{where} \quad L_t \triangleq e^{-(\lambda_1-\lambda_0)t}\prod_{k=1}^{N_t}\left[\frac{\lambda_1}{\lambda_0}f(Y_k)\right], \quad t\ge 0 \tag{A.73}$$

is the likelihood ratio process. The process $L=\{L_t;\ t\ge 0\}$ is the unique locally bounded solution of the differential equation (Elliott [8, p. 155])

$$dL_t = L_{t-}\left[-(\lambda_1-\lambda_0)dt + \int_{y\in\mathbb{R}^d}\left(\frac{\lambda_1}{\lambda_0}f(y)-1\right)p(dtdy)\right], \quad t\ge 0 \quad (L_0=1), \tag{A.74}$$

where $p(\cdot)$ is the point process in (A.72). Since the random variable θ is independent of the process X and has the exponential distribution in (6) under $\mathbb{P}_0$, the generalized Bayes theorem (Shiryaev [16, pp. 230-231]) and (A.73) give

$$\begin{aligned}\Phi_t &= \frac{\mathbb{E}[1_{\{\theta\le t\}}|\mathcal{F}_t]}{\mathbb{P}\{\tau>\theta|\mathcal{F}_t\}} = \frac{\mathbb{E}_0[Z_t 1_{\{\theta\le t\}}|\mathcal{F}_t]}{\mathbb{E}_0[Z_t|\mathcal{F}_t]}\left(\frac{\mathbb{E}_0[Z_t 1_{\{\theta>t\}}|\mathcal{F}_t]}{\mathbb{E}_0[Z_t|\mathcal{F}_t]}\right)^{-1} = \frac{e^{\lambda t}}{1-\pi}\,\mathbb{E}_0\left[\frac{L_t}{L_\theta}1_{\{\theta\le t\}}\middle|\,\mathcal{F}_t\right]\\ &= \frac{e^{\lambda t}}{1-\pi}\left[\pi L_t + (1-\pi)\int_0^t \lambda e^{-\lambda u}\frac{L_t}{L_u}\,du\right] \equiv \frac{\pi}{1-\pi}\,U_t + V_t\end{aligned}$$

in terms of

$$U_t \triangleq e^{\lambda t}L_t \qquad \text{and} \qquad V_t \triangleq \int_0^t \lambda e^{\lambda(t-u)}\frac{L_t}{L_u}du.$$

Using the change-of-variable formula (Protter [15, p. 78], Jacod and Shiryaev [12, p. 57], Cont and Tankov [4, p. 277]) and the dynamics of the process L in (A.74) give

$$\begin{aligned}dU_t &= \quad U_{t-}\left[(\lambda-\lambda_1+\lambda_0)dt + \int_{y\in\mathbb{R}^d}\left(\frac{\lambda_1}{\lambda_0}f(y)-1\right)p(dtdy)\right], \qquad U_0=1,\\ dV_t &= \lambda dt + V_{t-}\left[(\lambda-\lambda_1+\lambda_0)dt + \int_{y\in\mathbb{R}^d}\left(\frac{\lambda_1}{\lambda_0}f(y)-1\right)p(dtdy)\right], \qquad V_0=0.\end{aligned}$$

Therefore, the dynamics of the process $\Phi = [\pi/(1-\pi)]\cdot U + V$ are

$$\begin{aligned}d\Phi_t &= [\lambda + (\lambda-\lambda_1+\lambda_0)\Phi_t]\,dt + \Phi_{t-}\int_{y\in\mathbb{R}^d}\left[\frac{\lambda_1}{\lambda_0}f(y)-1\right]p(dtdy), \qquad t\ge 0,\\ \Phi_0 &= \frac{\pi}{1-\pi}.\end{aligned} \tag{A.75}$$

The stochastic differential equation in (A.75) can be solved pathwise and explicitly for Φ. Let the parameters $a=\lambda-\lambda_1+\lambda_0$, $\phi_d=-\lambda/a$ be defined as in (12) and let $x(\cdot,\phi)=\{x(t,\phi);\ t\in\mathbb{R}\}$, $\phi\in\mathbb{R}$ be the unique solution (given explicitly in (12)) of the ordinary differential equation

$$\frac{d}{dt}x(t,\phi) = \lambda + ax(t,\phi),\ t\in\mathbb{R} \qquad \text{and} \qquad x(0,\phi)=\phi. \tag{A.76}$$

As clearly seen from (A.75), the process Φ follows the integral curves of the differential equation in (A.76) between consecutive jumps of X and is updated instantaneously at every jump of X as summarized in (13).

A.3 The infinitesimal generator of the process Φ. The dynamics of Φ in (A.75) and the $(\mathbb{P}_0,\mathbb{F})$-compensator measure $\widetilde{p}_0(dsdy)=\lambda_0 ds\nu_0(dy)$ of the point process $p(\cdot)$ in (A.72) determine the $(\mathbb{P}_0,\mathbb{F})$-*infinitesimal generator* of the process Φ. Let $a=\lambda-\lambda_1+\lambda_0$ as in (12), and $h:\mathbb{R}_+\mapsto\mathbb{R}$ be any locally bounded continuously differentiable function. Then

$$\begin{aligned} h(\Phi_t)-h(\Phi_0) &= \int_0^t (\lambda+a\Phi_{s-})\,h'(\Phi_{s-})ds + \int_{(0,t]\times\mathbb{R}^d}\left[h\left(\frac{\lambda_1}{\lambda_0}f(y)\Phi_{s-}\right)-h(\Phi_{s-})\right]p(dsdy) \\ &= \int_0^t (\lambda+a\Phi_{s-})\,h'(\Phi_{s-})ds + \int_{(0,t]\times\mathbb{R}^d}\left[h\left(\frac{\lambda_1}{\lambda_0}f(y)\Phi_{s-}\right)-h(\Phi_{s-})\right]\widetilde{p}_0(dsdy) \\ &+ \int_{(0,t]\times\mathbb{R}^d}\left[h\left(\frac{\lambda_1}{\lambda_0}f(y)\Phi_{s-}\right)-h(\Phi_{s-})\right]q_0(dsdy) \\ &= \int_0^t \left.\left\{(\lambda+ax)\,h'(x)+\lambda_0\int_{\mathbb{R}^d}\left[h\left(\frac{\lambda_1}{\lambda_0}f(y)x\right)-h(x)\right]\nu_0(dy)\right\}\right|_{x=\Phi_{s-}} ds \\ &+ \int_{(0,t]\times\mathbb{R}^d}\left[h\left(\frac{\lambda_1}{\lambda_0}f(y)\Phi_{s-}\right)-h(\Phi_{s-})\right]q_0(dsdy). \end{aligned}$$

The last integral with respect to the compensated random measure $q_0(\cdot)=p_0(\cdot)-\widetilde{p}_0(\cdot)$ is a $(\mathbb{P}_0,\mathbb{F})$-local martingale. Therefore, the integrand of the last Lebesgue integral equals the $(\mathbb{P}_0,\mathbb{F})$-infinitesimal generator $(\mathcal{A}h)\circ\Phi_{s-}$ composed with the process Φ; see (14).

A.4 Long proofs. For the proof of Proposition 3.5, we shall need the following result on the characterization of $\mathbb{F}$-stopping times; see Brémaud [3, Theorem T33, p. 308], Davis [7, Lemma A2.3, p. 261].

Lemma A.1 *For every $\mathbb{F}$-stopping time τ and every $n\in\mathbb{N}_0$, there is an $\mathcal{F}_{\sigma_n}$-measurable random variable $R_n:\Omega\mapsto[0,\infty]$ such that $\tau\wedge\sigma_{n+1}=(\sigma_n+R_n)\wedge\sigma_{n+1}$ $\mathbb{P}_0$-a.s. on $\{\tau\ge\sigma_n\}$.*

Proof of Proposition 3.5. First, we shall establish the inequality

$$\mathbb{E}_0^\phi\int_0^{\tau\wedge\sigma_n} e^{-\lambda t}g\big(\Phi_t\big)dt \ge v_n(\phi), \qquad \tau\in\mathbb{F},\ \phi\in\mathbb{R}_+ \tag{A.77}$$

for every $n\in\mathbb{N}_0$, by proving inductively on $k=1,\dots,n+1$ that

$$\begin{aligned} &\mathbb{E}_0^\phi\left[\int_0^{\tau\wedge\sigma_n} e^{-\lambda t}g\big(\Phi_t\big)\right]dt \\ &\qquad \ge \mathbb{E}_0^\phi\left[\int_0^{\tau\wedge\sigma_{n-k+1}} e^{-\lambda t}g\big(\Phi_t\big)dt + 1_{\{\tau\ge\sigma_{n-k+1}\}}e^{-\lambda\sigma_{n-k+1}}v_{k-1}\big(\Phi_{\sigma_{n-k+1}}\big)\right] =: RHS_{k-1}. \end{aligned} \tag{A.78}$$

Observe that (A.77) follows from (A.78) when we set $k=n+1$.

If $k=1$, then the inequality (A.78) is satisfied as an equality since $v_0\equiv 0$. Suppose that (A.78) holds for some $1\le k<n+1$. We shall prove that it must also hold when k is replaced with $k+1$. Let us denote the righthand side of (A.78) by RHS_{k-1}, and rewrite it as

$$\begin{aligned} RHS_{k-1} &= RHS_{k-1}^{(1)}+RHS_{k-1}^{(2)} \triangleq \mathbb{E}_0^\phi\left[\int_0^{\tau\wedge\sigma_{n-k}} e^{-\lambda t}g\big(\Phi_t\big)dt\right] \\ &+ \mathbb{E}_0^\phi\left[1_{\{\tau\ge\sigma_{n-k}\}}\left(\int_{\sigma_{n-k}}^{\tau\wedge\sigma_{n-k+1}} e^{-\lambda t}g\big(\Phi_t\big)dt + 1_{\{\tau\ge\sigma_{n-k+1}\}}e^{-\lambda\sigma_{n-k+1}}v_{k-1}\big(\Phi_{\sigma_{n-k+1}}\big)\right)\right] \end{aligned} \tag{A.79}$$

where we used $\int_0^{\tau\wedge\sigma_{n-k+1}} = \int_0^{\tau\wedge\sigma_{n-k}} + 1_{\{\tau\ge\sigma_{n-k}\}}\int_{\tau\wedge\sigma_{n-k}}^{\tau\wedge\sigma_{n-k+1}}$, as well as $1_{\{\tau\ge\sigma_{n-k}\}}1_{\{\tau\ge\sigma_{n-k+1}\}} = 1_{\{\tau\ge\sigma_{n-k+1}\}}$. By Lemma A.1, there is an $\mathcal{F}_{\sigma_{n-k}}$-measurable random variable R_{n-k} such that $\tau\wedge\sigma_{n-k+1} = (\sigma_{n-k}+R_{n-k})\wedge\sigma_{n-k+1}$ $\mathbb{P}_0$-almost surely on $\{\tau\ge\sigma_{n-k}\}$. Therefore, the second expectation, denoted by $RHS_{k-1}^{(2)}$, in (A.79) becomes

$$\begin{aligned} &\mathbb{E}_0^\phi\left\{1_{\{\tau\ge t\}}\left.\left[\int_t^{(t+R_{n-k})\wedge s} e^{-\lambda t}g\big(\Phi_t\big)dt + 1_{\{t+R_{n-k}\ge s\}}e^{-\lambda s}v_{k-1}\big(\Phi_s\big)\right]\right|_{\substack{t=\sigma_{n-k}\\ s=\sigma_{n-k+1}}}\right\} \\ &\qquad\qquad = \mathbb{E}_0^\phi\left\{1_{\{\tau\ge\sigma_{n-k}\}}e^{-\lambda\sigma_{n-k}}f_{n-k}(R_{n-k},\Phi_{\sigma_{n-k}})\right\} \end{aligned}$$

by the strong Markov property of X, where $f_{k-1}(r,\phi)$ equals

$$\mathbb{E}_0^\phi\left[\int_0^{r\wedge\sigma_1} e^{-\lambda t} g(\Phi_t)dt + 1_{\{r\ge\sigma_1\}} e^{-\lambda\sigma_1} v_{k-1}(\Phi_{\sigma_1})\right] = Jv_{k-1}(r,\phi) \geq J_0 v_{k-1}(\phi) = v_k(\phi).$$

The (in)equalities follow from (21), (22) and (23), respectively. Thus

$$RHS_{k-1}^{(2)} \geq \mathbb{E}_0^\phi\left[1_{\{\tau\ge\sigma_{n-k}\}} e^{-\lambda\sigma_{n-k}} v_k(\Phi_{\sigma_{n-k}})\right].$$

From (A.78) and (A.79), we finally obtain

$$\begin{aligned}\mathbb{E}_0^\phi\left[\int_0^{\tau\wedge\sigma_n} e^{-\lambda t} g(\Phi_t)dt\right] &\geq RHS_{k-1} = \mathbb{E}_0^\phi\left[\int_0^{\tau\wedge\sigma_{n-k}} e^{-\lambda t} g(\Phi_t)dt\right] + RHS_{k-1}^{(2)} \\ &\geq \mathbb{E}_0^\phi\left[\int_0^{\tau\wedge\sigma_{n-k}} e^{-\lambda t} g(\Phi_t)dt + 1_{\{\tau\ge\sigma_{n-k}\}} e^{-\lambda\sigma_{n-k}} v_k(\Phi_{\sigma_{n-k}})\right] = RHS_k.\end{aligned}$$

This completes the proof of (A.78) by induction on k, and (A.77) follows by setting $k=n+1$ in (A.78). When we take the infimum of both sides in (A.77), we obtain $V_n \geq v_n$, $n\in\mathbb{N}$.

The opposite inequality $V_n \le v_n$, $n\in\mathbb{N}$ follows immediately from (29) since every $\mathbb{F}$-stopping time S_n^ε is less than or equal to σ_n, $\mathbb{P}_0$-a.s. by construction. Therefore, we only need to establish (29). We will prove it by induction on $n\in\mathbb{N}$. For $n=1$, the lefthand side of (29) becomes $\mathbb{E}_0^\phi\left[\int_0^{S_1^\varepsilon} e^{-\lambda t} g(\Phi_t)dt\right] = \mathbb{E}_0^\phi\left[\int_0^{r_0^\varepsilon(\phi)\wedge\sigma_1} e^{-\lambda t} g(\Phi_t)dt\right] = Jv_0(r_0^\varepsilon(\phi),\phi)$. Since $Jv_0(r_0^\varepsilon(\phi),\phi) \le J_0v_0(\phi)+\varepsilon$ by Remark 3.2, the inequality (29) holds for $n=1$.

Suppose that (29) holds for every $\varepsilon>0$ for some $n\in\mathbb{N}$. We will prove that it also holds when n is replaced with $n+1$. Since $S_{n+1}^\varepsilon\wedge\sigma_1 = r_n^{\varepsilon/2}(\Phi_0)\wedge\sigma_1$, $\mathbb{P}_0$-a.s., we have

$$\begin{aligned}\mathbb{E}_0^\phi\left[\int_0^{S_{n+1}^\varepsilon} e^{-\lambda t} g(\Phi_t)dt\right] &= \mathbb{E}_0^\phi\left[\int_0^{S_{n+1}^\varepsilon\wedge\sigma_1} e^{-\lambda t} g(\Phi_t)dt + 1_{\{S_{n+1}^\varepsilon\ge\sigma_1\}}\int_{\sigma_1}^{S_{n+1}^\varepsilon} e^{-\lambda t} g(\Phi_t)dt\right] \\ &= \mathbb{E}_0^\phi\left[\int_0^{r_n^{\varepsilon/2}(\phi)\wedge\sigma_1} e^{-\lambda t} g(\Phi_t)dt + 1_{\{r_n^{\varepsilon/2}(\phi)\ge\sigma_1\}}\int_{\sigma_1}^{\sigma_1+S_n^{\varepsilon/2}\circ\theta_{\sigma_1}} e^{-\lambda t} g(\Phi_t)dt\right] \\ &= \mathbb{E}_0^\phi\left[\int_0^{r_n^{\varepsilon/2}(\phi)\wedge\sigma_1} e^{-\lambda t} g(\Phi_t)dt\right] + \mathbb{E}_0^\phi\left[1_{\{r_n^{\varepsilon/2}(\phi)\ge\sigma_1\}} e^{-\lambda\sigma_1} f_n(\Phi_{\sigma_1})\right]\end{aligned}$$

by strong Markov property of X, where $f_n(\phi)\triangleq\mathbb{E}_0^\phi\left[\int_0^{S_n^{\varepsilon/2}} e^{-\lambda t} g(\Phi_t)dt\right] \le v_n(\phi)+\varepsilon/2$. by the induction hypothesis. Therefore, $\mathbb{E}_0^\phi\left[\int_0^{S_{n+1}^\varepsilon} e^{-\lambda t} g(\Phi_t)dt\right]$ is less than or equal to

$$\mathbb{E}_0^\phi\left[\int_0^{r_n^{\varepsilon/2}(\phi)\wedge\sigma_1} e^{-\lambda t} g(\Phi_t)dt + 1_{\{r_n^{\varepsilon/2}(\phi)\ge\sigma_1\}} e^{-\lambda\sigma_1} v_n(\Phi_{\sigma_1})\right] + \frac{\varepsilon}{2} = Jv_n(r_n^{\varepsilon/2}(\phi),\phi) + \frac{\varepsilon}{2}.$$

However, $Jv_n(r_n^{\varepsilon/2}(\phi),\phi) \le v_{n+1}(\phi)+\varepsilon/2$ by Remark 3.2. The last two inequalities prove (29) when n is replaced with $n+1$. □

Proof of Lemma 3.7. Let us fix a constant $u\ge t$ and $\phi\in\mathbb{R}_+$. Then

$$Jw(u,\phi) = \mathbb{E}_0^\phi\left[\left(\int_0^{t\wedge\sigma_1} + 1_{\{\sigma_1>t\}}\int_t^{u\wedge\sigma_1}\right) e^{-\lambda s} g(\Phi_s)ds + 1_{\{u\ge\sigma_1\}} e^{-\lambda\sigma_1} w(\Phi_{\sigma_1})\right]. \tag{A.80}$$

On the event $\{\sigma_1>t\}$, we have $u\wedge\sigma_1 = [t+(u-t)]\wedge[t+\sigma_1\circ\theta_t] = t+[(u-t)\wedge\sigma_1\circ\theta_t]$. Therefore, the strong Markov property of X gives

$$\begin{aligned}\mathbb{E}_0^\phi\left[1_{\{\sigma_1>t\}}\int_t^{u\wedge\sigma_1} e^{-\lambda s} g(\Phi_s)ds\right] &= \mathbb{E}_0^\phi\left[1_{\{\sigma_1>t\}} e^{-\lambda t}\mathbb{E}_0^{\Phi_t}\left[\int_0^{(u-t)\wedge\sigma_1} e^{-\lambda s} g(\Phi_s)ds\right]\right] \\ &= \mathbb{E}_0^\phi\left[1_{\{\sigma_1>t\}} e^{-\lambda t}\Big(Jw(u-t,\Phi_t) - \mathbb{E}_0^{\Phi_t}\left[1_{\{u-t\ge\sigma_1\}} e^{-\lambda\sigma_1} w(\Phi_{\sigma_1})\right]\Big)\right] \\ &= e^{-(\lambda+\lambda_0)t}\, Jw\big(u-t, x(t,\phi)\big) - \mathbb{E}_0^\phi\left[1_{\{\sigma_1>t\}} 1_{\{u\ge\sigma_1\}} e^{-\lambda\sigma_1} w(\Phi_{\sigma_1})\right].\end{aligned} \tag{A.81}$$

The second equality follows from the definition of Jw in (21), and the last from (13) and the strong Markov property. Substituting (A.81) into (A.80) gives

$$\begin{aligned} Jw(u,\phi) &= \mathbb{E}_0^{\phi}\left[\int_0^{t\wedge\sigma_1} e^{-\lambda s} g(\Phi_s)ds + 1_{\{t\geq\sigma_1\}} e^{-\lambda\sigma_1} w(\Phi_{\sigma_1})\right] + e^{-(\lambda+\lambda_0)t}\, Jw\big(u-t, x(t,\phi)\big) \\ &= Jw(t,\phi) + e^{-(\lambda+\lambda_0)t}\, Jw\big(u-t, x(t,\phi)\big). \end{aligned}$$

Finally, taking the infimum of both sides over $u \in [t,+\infty]$ proves (30). □

Proof of Proposition 3.11. Note that the sequence of random variables

$$\int_0^{U_\varepsilon\wedge\sigma_n} e^{-\lambda s} g(\Phi_s)ds + e^{-\lambda(U_\varepsilon\wedge\sigma_n)} V(\Phi_{U_\varepsilon\wedge\sigma_n}) \geq -2\int_0^\infty e^{-\lambda s}\,\frac{\lambda}{c}\,ds = -\frac{2}{c}$$

is bounded from below; see (16). By (39) and Fatou's Lemma, we have

$$\begin{aligned} V(\phi) &\geq \mathbb{E}_0^{\phi}\left[\varliminf_{n\to\infty}\left(\int_0^{U_\varepsilon\wedge\sigma_n} e^{-\lambda s} g(\Phi_s)ds + e^{-\lambda(U_\varepsilon\wedge\sigma_n)} V(\Phi_{U_\varepsilon\wedge\sigma_n})\right)\right] \\ &= \mathbb{E}_0^{\phi}\left[\int_0^{U_\varepsilon} e^{-\lambda s} g(\Phi_s)ds + 1_{\{U_\varepsilon<\infty\}} e^{-\lambda U_\varepsilon} V(\Phi_{U_\varepsilon})\right] \\ &\geq \mathbb{E}_0^{\phi}\left[\int_0^{U_\varepsilon} e^{-\lambda s} g(\Phi_s)ds\right] - \varepsilon\, \mathbb{E}_0^{\phi}\left[1_{\{U_\varepsilon<\infty\}} e^{-\lambda U_\varepsilon}\right] \geq \mathbb{E}_0^{\phi}\left[\int_0^{U_\varepsilon} e^{-\lambda s} g(\Phi_s)ds\right] - \varepsilon \end{aligned}$$

for every $\phi \in \mathbb{R}_+$. This concludes the proof. □

Proof of Proposition 3.12. First, let us show (39) for $n = 1$. Fix $\varepsilon \geq 0$ and $\phi \in \mathbb{R}_+$. By Lemma A.1, there exists a constant $u \in [0,\infty]$ such that $U_\varepsilon \wedge \sigma_1 = u \wedge \sigma_1$. Then

$$\begin{aligned} \mathbb{E}_0^{\phi}[M_{U_\varepsilon\wedge\sigma_1}] &= \mathbb{E}_0^{\phi}\left[\int_0^{u\wedge\sigma_1} e^{-\lambda s} g(\Phi_s)ds + 1_{\{u\geq\sigma_1\}} e^{-\lambda\sigma_1} V(\Phi_{\sigma_1})\right] \\ &\quad + \mathbb{E}_0^{\phi}\left[1_{\{u<\sigma_1\}} e^{-\lambda u} V(\Phi_u)\right] = JV(u,\phi) + e^{-(\lambda+\lambda_0)u} V\big(x(u,\phi)\big) = J_u V(\phi), \end{aligned} \tag{A.82}$$

where the second equality follows from (21) and (13), and the last from (33).

Fix any $t \in [0,u)$. By (33) and (13),

$$\begin{aligned} JV(t,\phi) &= J_t V(\phi) - e^{-(\lambda+\lambda_0)t} V\left(x(t,\phi)\right) \\ &\geq J_0 V(\phi) - e^{-(\lambda+\lambda_0)t} V\left(x(t,\phi)\right) = J_0 V(\phi) - \mathbb{E}_0^{\phi}\left[1_{\{\sigma_1>t\}} e^{-\lambda t} V(\Phi_t)\right]. \end{aligned}$$

On the event $\{\sigma_1 > t\}$, we have $U_\varepsilon > t$ (otherwise, $U_\varepsilon \leq t < \sigma_1$ would imply $U_\varepsilon = u \leq t$, which contradicts with our initial choice of $t < u$). Thus, $V(\Phi_t) < -\varepsilon$ on $\{\sigma_1 > t\}$. Hence, $JV(t,\phi) > J_0V(\phi) + \varepsilon\,\mathbb{E}_0^{\phi}\left[1_{\{\sigma_1>t\}} e^{-\lambda t}\right] = J_0V(\phi) + \varepsilon\, e^{-(\lambda+\lambda_0)u} \geq J_0V(\phi)$ for every $t \in [0,u)$. Therefore, $J_0V(\phi) = J_uV(\phi)$, and (A.82) implies $\mathbb{E}_0^{\phi}[M_{U_\varepsilon\wedge\sigma_1}] = J_uV(\phi) = J_0V(\phi) = V(\phi) = \mathbb{E}_0^{\phi}[M_0]$. This completes the proof of (39) for $n = 1$.

Now suppose that (39) holds for some $n \in \mathbb{N}$, and let us show the same equality for $n+1$. Note that $\mathbb{E}_0^{\phi}[M_{U_\varepsilon\wedge\sigma_{n+1}}] = \mathbb{E}_0^{\phi}[1_{\{U_\varepsilon<\sigma_1\}} M_{U_\varepsilon}] + \mathbb{E}_0^{\phi}[1_{\{U_\varepsilon\geq\sigma_1\}} M_{U_\varepsilon\wedge\sigma_{n+1}}]$ equals

$$\begin{aligned} &\mathbb{E}_0^{\phi}[1_{\{U_\varepsilon<\sigma_1\}} M_{U_\varepsilon}] + \mathbb{E}_0^{\phi}\left[1_{\{U_\varepsilon\geq\sigma_1\}} \int_0^{\sigma_1} e^{-\lambda s} g(\Phi_s)ds\right] \\ &\qquad + \mathbb{E}_0^{\phi}\left[1_{\{U_\varepsilon\geq\sigma_1\}}\left\{\int_{\sigma_1}^{U_\varepsilon\wedge\sigma_{n+1}} e^{-\lambda s} g(\Phi_s)ds + e^{-\lambda(U_\varepsilon\wedge\sigma_{n+1})} V(\Phi_{U_\varepsilon\wedge\sigma_{n+1}})\right\}\right]. \end{aligned}$$

Since $U_\varepsilon \wedge \sigma_{n+1} = \sigma_1 + [(U_\varepsilon \wedge \sigma_n) \circ \theta_{\sigma_1}]$ on the event $\{U_\varepsilon \geq \sigma_1\}$, the strong Markov property of Φ at the stopping time σ_1 implies that $\mathbb{E}_0^\phi[M_{U_\varepsilon \wedge \sigma_{n+1}}]$ equals

$$\begin{aligned}
&\mathbb{E}_0^\phi[1_{\{U_\varepsilon<\sigma_1\}} M_{U_\varepsilon}] + \mathbb{E}_0^\phi\left[1_{\{U_\varepsilon\geq\sigma_1\}}\int_0^{\sigma_1} e^{-\lambda s} g(\Phi_s)ds\right] \\
&\qquad + \mathbb{E}_0^\phi\left[1_{\{U_\varepsilon\geq\sigma_1\}} e^{-\lambda\sigma_1}\, \underbrace{\mathbb{E}_0^{\Phi_{\sigma_1}}\left[\int_0^{U_\varepsilon\wedge\sigma_n} e^{-\lambda s} g(\Phi_s)ds + e^{-\lambda(U_\varepsilon\wedge\sigma_n)} V(\Phi_{U_\varepsilon\wedge\sigma_n})\right]}_{\text{is equal to } V(\Phi_t) \text{ by the induction hypothesis}}\right] \\
&= \mathbb{E}_0^\phi[1_{\{U_\varepsilon<\sigma_1\}} M_{U_\varepsilon}] + \mathbb{E}_0^\phi\left[1_{\{U_\varepsilon\geq\sigma_1\}}\left(\int_0^{\sigma_1} e^{-\lambda s} g(\Phi_s)ds + e^{-\lambda\sigma_1} V(\Phi_{\sigma_1})\right)\right] \\
&= \mathbb{E}_0^\phi[1_{\{U_\varepsilon<\sigma_1\}} M_{U_\varepsilon}] + \mathbb{E}_0^\phi[1_{\{U_\varepsilon\geq\sigma_1\}} M_{\sigma_1}] = \mathbb{E}^\phi[M_{U_\varepsilon\wedge\sigma_1}] = \mathbb{E}^\phi[M_0],
\end{aligned}$$

where the last equality was proved above. This concludes the proof of the induction step. □

Proof of Proposition 4.1. The hypotheses guarantee that the process Φ always jumps forward and does not return to $[0, \lambda/c)$ after it leaves the same interval at time $\underline{\tau}$. Therefore, for every stopping time $\tau \in \mathbb{F}$,

$$\begin{aligned}
\mathbb{E}_0^\phi\left[\int_0^\tau e^{-\lambda t} g(\Phi_t)dt\right] &\geq \mathbb{E}_0^\phi\left[\int_0^{\tau\vee\underline{\tau}} e^{-\lambda t} g(\Phi_t)dt\right] \\
&= \mathbb{E}_0^\phi\left[\int_0^{\underline{\tau}} e^{-\lambda t} g(\Phi_t)dt\right] + \mathbb{E}_0^\phi\left[1_{\{\tau\geq\underline{\tau}\}}\int_{\underline{\tau}}^\tau e^{-\lambda t} g(\Phi_t)dt\right] \geq \mathbb{E}_0^\phi\left[\int_0^{\underline{\tau}} e^{-\lambda t} g(\Phi_t)dt\right].
\end{aligned}$$

□

Proof of Proposition 4.2. Proposition 3.6 and Corollary 3.4 imply for every $\phi \in [0, \lambda/c)$ that

$$V(\phi) = v(\phi) \leq v_1(\phi) = J_0 v_0(\phi) = \inf_{t\in[0,\infty]} \int_0^t e^{-(\lambda+\lambda_0)u} g(x(u,\phi))du < 0,$$

since the continuous curve $t \mapsto x(t,\phi)$ in (12) stays in the interval $[0, \lambda/c) = \{x \in \mathbb{R}_+ : g(x) < 0\}$ for some positive amount of time. Therefore, we have $[0, \lambda/c) \subseteq \{\phi \in \mathbb{R}_+ : v_1(\phi) < 0\} \subseteq \{\phi \in \mathbb{R}_+ : V(\phi) < 0\}$, and the first inclusion in (42) follows.

For the proof of the last inclusion, let us fix any stopping time $\tau \in \mathbb{F}$. By Lemma A.1, there exists some constant $t \in [0,\infty]$ such that $\tau \wedge \sigma_1 = t \wedge \sigma_1$ almost surely, and

$$\begin{aligned}
\mathbb{E}_0^\phi\left[\int_0^\tau e^{-\lambda u} g(\Phi_u)du\right] &= \mathbb{E}_0^\phi\left[\int_0^{\tau\wedge\sigma_1} e^{-\lambda u} g(\Phi_u)du\right] + \mathbb{E}_0^\phi\left[1_{\{\tau>\sigma_1\}}\int_{\sigma_1}^\tau e^{-\lambda u} g(\Phi_u)du\right] \\
&\geq \mathbb{E}_0^\phi\left[\int_0^t 1_{\{u<\sigma_1\}} e^{-\lambda u} g(x(u,\phi))du\right] - \frac{1}{c}\cdot\mathbb{E}_0^\phi\left[1_{\{t>\sigma_1\}} e^{-\lambda\sigma_1}\right] \\
&= \int_0^t e^{-(\lambda+\lambda_0)u}\left[x(u,\phi) - \frac{\lambda+\lambda_0}{c}\right] du.
\end{aligned}$$

The inequality and the last equality follow from that $g(\phi) = \phi - \lambda/c \geq -\lambda/c$ for every $\phi \in \mathbb{R}_+$, and that the first jump time σ_1 of the observation process X has exponential distribution with rate λ_0 under $\mathbb{P}_0$, respectively. Now the infimum of both sides gives

$$0 \geq V(\phi) \geq \inf_{t\in[0,\infty]} h(t,\phi) \triangleq \int_0^t e^{-(\lambda+\lambda_0)u}\left[x(u,\phi) - \frac{\lambda+\lambda_0}{c}\right] du. \tag{A.83}$$

The solution of the deterministic optimization problem on the right depends on ϕ_d in (12).

Case I: $\phi_d \notin [0, (\lambda+\lambda_0)/c)$. Then $x(u,\phi) \geq (\lambda+\lambda_0)/c$ for every $u \geq 0$ and $\phi \geq (\lambda+\lambda_0)/c$. Therefore, the infimum in (A.83) is attained at $t = 0$ and $V(\phi) = 0$ if $\phi \geq (\lambda+\lambda_0)/c$.

Case II: $\phi_d \in [0, (\lambda+\lambda_0)/c)$. As t tends to infinity, the *monotone* function $t \mapsto x(t,\phi)$ converges to $\phi_d \in [0, (\lambda+\lambda_0)/c)$. Therefore, the infimum in (A.83) is attained at $t = 0$ if $h(\infty,\phi) \geq 0$, and at $t = \infty$ otherwise. However, the affine function $\phi \mapsto h(\infty,\phi) : \mathbb{R}_+ \mapsto \mathbb{R}$ increases to $+\infty$ with ϕ, is negative at $\phi = (\lambda+\lambda_1)/c > \phi_d$ and has unique zero at $\phi = \overline{\xi} > (\lambda+\lambda_0)/c$ of (41). Thus, the infimum in (A.83) is attained at $t = 0$ and $V(\phi) = 0$ for every $\phi \geq \overline{\xi}$. Finally, both cases imply together the second inclusion in (42). □

References

[1] Erhan Bayraktar and Savas Dayanik, *Poisson disorder problem with exponential penalty for delay*, *Mathematics of Operations Research*. To appear, 2005.

[2] Erhan Bayraktar, Savas Dayanik, and Ioannis Karatzas, *The standard Poisson disorder problem revisited*, Stochastic Process. Appl. **115** (2005), no. 9, 1437–1450. MR MR2158013

[3] Pierre Brémaud, *Point Processes and Queues*, Springer-Verlag, New York, 1981, Martingale dynamics, Springer Series in Statistics. MR 82m:60058

[4] Rama Cont and Peter Tankov, *Financial Modelling with Jump Processes*, Chapman & Hall/CRC Financial Mathematics Series, Chapman & Hall/CRC, Boca Raton, FL, 2004. MR MR2042661

[5] M. H. A. Davis, *A note on the Poisson disorder problem*, Banach Center Publ. **1** (1976), 65–72.

[6] ______, *Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models*, J. Roy. Statist. Soc. Ser. B **46** (1984), no. 3, 353–388, With discussion. MR 87g:60062

[7] ______, *Markov Models and Optimization*, Monographs on Statistics and Applied Probability, vol. 49, Chapman & Hall, London, 1993. MR 96b:90002

[8] Robert J. Elliott, *Stochastic Calculus and Applications*, Applications of Mathematics (New York), vol. 18, Springer-Verlag, New York, 1982. MR MR678919 (85b:60059)

[9] L. I. Galchuk and B. L. Rozovskii, *The disorder problem for a Poisson process*, Theory of Prob. and Appl. **16** (1971), 729–734.

[10] Pavel V. Gapeev, *The disorder problem for compound Poisson processes with exponential jumps*, Ann. Appl. Probab. **15** (2005), no. 1A, 487–499. MR MR2115049

[11] U. S. Gugerli, *Optimal stopping of a piecewise-deterministic Markov process*, Stochastics **19** (1986), no. 4, 221–236. MR 88j:60084

[12] Jean Jacod and Albert N. Shiryaev, *Limit Theorems for Stochastic Processes*, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer-Verlag, Berlin, 2003. MR 2003j:60001

[13] Goran Peskir and Albert N. Shiryaev, *Solving the Poisson disorder problem*, Advances in finance and stochastics, Springer, Berlin, 2002, pp. 295–312. MR 2003i:60071

[14] H. Vincent Poor, *An Introduction to Signal Detection and Estimation*, second ed., Springer Texts in Electrical Engineering, Springer-Verlag, New York, 1994, A Dowden & Culver Book. MR MR1270019 (95c:94004)

[15] Philip E. Protter, *Stochastic Integration and Differential Equations*, second ed., Applications of Mathematics (New York), vol. 21, Springer-Verlag, Berlin, 2004, Stochastic Modelling and Applied Probability. MR MR2020294

[16] A. N. Shiryaev, *Probability*, Graduate Texts in Mathematics, vol. 95, Springer-Verlag, New York, 1996. MR 97c:60003

[17] Albert N. Shiryaev, *A remark on the quickest detection problems*, Statist. Decisions **22** (2004), no. 1, 79–82. MR MR2065992