SEQUENTIAL MULTI-HYPOTHESIS TESTING FOR COMPOUND POISSON PROCESSES

SAVAS DAYANIK, H. VINCENT POOR, AND SEMIH O. SEZER

Abstract. Suppose that there are finitely many simple hypotheses about the unknown arrival rate and mark distribution of a compound Poisson process, and that exactly one of them is correct. The objective is to determine the correct hypothesis with minimal error probability and as soon as possible after the observation of the process starts. This problem is formulated in a Bayesian framework, and its solution is presented. Provably convergent numerical methods and practical near-optimal strategies are described and illustrated on various examples.

1. Introduction

Let X be a compound Poisson process defined as

$$X_t = X_0 + \sum_{i=1}^{N_t} Y_i, \qquad t \geq 0, \tag{1.1}$$

where N is a simple Poisson process with some arrival rate λ, and $Y_1, Y_2, \ldots$ are independent $\mathbb{R}^d$-valued random variables with some common distribution $\nu(\cdot)$ such that $\nu(\{0\}) = 0$. Suppose that the characteristics (λ, ν) of the process X are unknown, and exactly one of M distinct hypotheses,

$$H_1 : (\lambda, \nu) = (\lambda_1, \nu_1), \quad H_2 : (\lambda, \nu) = (\lambda_2, \nu_2), \quad \ldots\ , \quad H_M : (\lambda, \nu) = (\lambda_M, \nu_M), \tag{1.2}$$

is correct. Let Θ be the index of correct hypothesis, and $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$ without loss of generality.

At time $t = 0$, the hypothesis H_i, $i \in I \triangleq \{1, \ldots, M\}$ is correct with prior probability $\mathbb{P}^{\vec{\pi}}\{\Theta = i\} = \pi_i$ for some $\vec{\pi} \in E \triangleq \{(\pi_1, \ldots, \pi_M) \in [0,1]^M : \sum_{k=1}^{M} \pi_k = 1\}$, and we start observing the process X. The objective is to determine the correct hypothesis as quickly as possible with minimal probability of making a wrong decision. We are allowed to observe the process X before the final decision as long as we want. Postponing the final decision improves the odds of its correctness but increases the sampling and lost opportunity costs.

2000 *Mathematics Subject Classification.* Primary 62L10; Secondary 62L15, 62C10, 60G40.

Key words and phrases. Sequential hypothesis testing, compound Poisson processes, optimal stopping.

Therefore, a good strategy must resolve optimally the trade-off between the costs of waiting and a wrong terminal decision.

A strategy (τ, d) consists of a sampling termination time τ and a terminal decision rule d. At time τ, we stop observing the process X and select the hypothesis H_i on the event $\{d = i\}$ for $i \in I$. A strategy (τ, d) is *admissible* if τ is a stopping time of the process X, and if the value of d is determined completely by the observations of X until time τ.

Suppose that $a_{ij} \geq 0$, $i, j \in I$ is the cost of deciding on H_i when H_j is correct, $a_{ii} = 0$ for every $i \in I$. Moreover, $\rho > 0$ is a discount rate, and $f : \mathbb{R}^d \mapsto \mathbb{R}$ is a Borel function such that its negative part $f^-(\cdot)$ is ν_i-integrable for every $i \in I$; i.e., $\int_{\mathbb{R}^d} f^-(y)\nu_i(dy) < \infty$, $i \in I$. The compromise achieved by an admissible strategy (τ, d) between sampling cost and cost of wrong terminal decision may be captured by one of three alternative Bayes risks,

$$R(\vec{\pi}, \tau, d) \triangleq \mathbb{E}^{\vec{\pi}} \left[\tau + \mathbf{1}_{\{\tau<\infty\}} \sum_{i=1}^{M} \sum_{j=1}^{M} a_{ij} \cdot \mathbf{1}_{\{d=i,\, \Theta=j\}} \right], \tag{1.3}$$

$$R'(\vec{\pi}, \tau, d) \triangleq \mathbb{E}^{\vec{\pi}} \left[N_\tau + \mathbf{1}_{\{\tau<\infty\}} \sum_{i=1}^{M} \sum_{j=1}^{M} a_{ij} \cdot \mathbf{1}_{\{d=i,\, \Theta=j\}} \right], \tag{1.4}$$

$$R''(\vec{\pi}, \tau, d) \triangleq \mathbb{E}^{\vec{\pi}} \left[\sum_{k=1}^{N_\tau} e^{-\rho\sigma_k} f(Y_k) + e^{-\rho\tau} \sum_{i=1}^{M} \sum_{j=1}^{M} a_{ij} \cdot \mathbf{1}_{\{d=i,\, \Theta=j\}} \right], \tag{1.5}$$

and our objective will be (i) to calculate the minimum Bayes risk

$$V(\vec{\pi}) \triangleq \inf_{(\tau,d)\in\mathcal{A}} R(\vec{\pi}, \tau, d), \qquad \vec{\pi} \in E, \quad \text{(similarly, } V'(\cdot) \text{ and } V''(\cdot)\text{)} \tag{1.6}$$

over the collection $\mathcal{A}$ of all admissible strategies and (ii) to find an admissible decision rule that attains the infimum for every $\vec{\pi} \in E$, if such a rule exists.

The Bayes risk R penalizes the waiting time at a constant rate regardless of the true identity of the system and is suitable if major running costs are wages and rents paid at the same rates per unit time during the study. Bayes risks R' and R'' reflect the impact of the unknown system identity on the waiting costs and account better for lost opportunity costs in a business setting, for example, due to unknown volume of customer arrivals or unknown amount of (discounted) cash flows brought by the same customers.

Sequential Bayesian hypothesis testing problems under Bayes risk R of (1.3) were studied by Wald and Wolfowitz (1950), Blackwell and Girshick (1954), Zacks (1971), Shiryaev (1978). Peskir and Shiryaev (2000) solved this problem for testing two simple hypotheses about the arrival rate of a simple Poisson process. For a compound Poisson process whose marks are exponentially distributed with mean the same as their arrival rate, Gapeev (2002) derived under the same Bayes risk optimal sequential tests for two simple hypotheses about both mark distribution and arrival rate. The solution for general mark distributions has been obtained by Dayanik and Sezer (2005), who formulated the problem under an auxiliary probability measure as the optimal stopping of an $\mathbb{R}_+$-valued likelihood ratio process. Unfortunately, that approach for $M = 2$ fails when $M \geq 3$, since optimal continuation region in the latter case is not always bounded, and good convergence results do not exist anymore.

In this paper, we solve the Bayesian sequential multi-hypothesis testing problems in (1.6) for every $M \geq 2$ without restrictions on the arrival rate and mark distribution of the compound Poisson process. We describe an optimal admissible strategy $(U_0, d(U_0))$ in terms of an E-valued Markov process $\Pi(t) \triangleq (\Pi_1(t), \ldots, \Pi_M(t))$, $t \geq 0$, whose ith coordinate is the posterior probability $\Pi_i(t) = \mathbb{P}^{\vec{\pi}}\{\Theta = i \mid \mathcal{F}_t\}$ that hypothesis H_i is correct given the past observations $\mathcal{F}_t = \sigma\{X_s;\ 0 \leq s \leq t\}$ of the process X. The stopping time U_0 of the observable filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ is the hitting time of the process $\Pi(t)$ to some closed subset Γ_∞ of E. The stopping region Γ_∞ can be partitioned into M closed convex sets $\Gamma_{\infty,1}, \ldots, \Gamma_{\infty,M}$, and $d(U_0)$ equals $i \in I$ on the event $\{U_0 < \infty, \Pi(U_0) \in \Gamma_{\infty,i}\}$. Namely, the optimal admissible strategy $(U_0, d(U_0))$ is fully determined by the collection of subsets $\Gamma_{\infty,1}, \ldots, \Gamma_{\infty,M}$, and they will depend of course on the choice of Bayes risks R, R', and R''.

In plain words, one observes the process X until the uncertainty about true system identity is reduced to a low level beyond which waiting costs are no longer justified. At this time, the observer stops collecting new information and selects the least-costly hypothesis.

The remainder of the paper is organized as follows. In Sections 2 through 5, we discuss exclusively the problem with the first Bayes risk, R. We start with a precise problem description in Section 2 and show that (1.6) is equivalent to optimal stopping of the posterior probability process Π. After the successive approximations of the latter problem are studied in Section 3, the solution is described in Section 4. A provably-convergent numerical algorithm that gives nearly-optimal admissible strategies is also developed. The algorithm is illustrated on several examples in Section 5. Finally, we mention in Section 6 the necessary changes to previous arguments in order to obtain similar results for the alternative Bayes risks R' and R'' in (1.4). Long derivations and proofs are deferred to the appendix.

2. Problem description

Let $(\Omega, \mathcal{F})$ be a measurable space hosting a counting process N, a sequence of $\mathbb{R}^d$-valued random variables $(Y_n)_{n\geq 1}$, and a random variable Θ taking values in $I \triangleq \{1, \ldots, M\}$. Suppose that $\mathbb{P}_1, \ldots, \mathbb{P}_M$ are probability measures on this space such that the process X in (1.1) is a compound Poisson process with the characteristics $(\lambda_i, \nu_i(\cdot))$ under $\mathbb{P}_i$, and

$$\mathbb{P}_i\{\Theta = j\} = \left\{ \begin{matrix} 1, & \text{if } j = i \\ 0, & \text{if } j \neq i \end{matrix} \right\} \qquad \text{for every} \quad i \in I.$$

Then for every $\vec{\pi} \in E \triangleq \{(\pi_1, \ldots, \pi_M) \in [0,1]^M : \sum_{i=1}^M \pi_i = 1\}$,

$$\mathbb{P}^{\vec{\pi}}(A) \triangleq \sum_{i=1}^{M} \pi_i \mathbb{P}_i(A), \qquad A \in \mathcal{F} \tag{2.1}$$

is a probability measure on $(\Omega, \mathcal{F})$, and for $A = \{\Theta = i\}$ we obtain $\mathbb{P}^{\vec{\pi}}\{\Theta = i\} = \pi_i$. Namely, $\vec{\pi} = (\pi_1, \ldots, \pi_M)$ gives the prior probabilities of the hypotheses in (1.2), and $\mathbb{P}_i(A)$ becomes the conditional probability $\mathbb{P}^{\vec{\pi}}(A|\Theta = i)$ of the event $A \in \mathcal{F}$ given that the correct hypothesis is $\Theta = i$ for every $i \in I$.

Let $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ be the natural filtration of the process X, and $\mathcal{A}$ be the collection of the pairs (τ, d) of an $\mathbb{F}$-stopping time τ and an I-valued $\mathcal{F}_\tau$-measurable random variable d. We would like to calculate the minimum Bayes risk $V(\cdot)$ in (1.6) and find a strategy $(\tau, d) \in \mathcal{A}$ that has the smallest Bayes risk $R(\vec{\pi}, \tau, d)$ of (1.3) for every $\vec{\pi} \in E$ (Bayes risks R' and R'' in (1.4) and (1.5) are discussed in Section 6). Next proposition, whose proof is in Appendix A, identifies the form of the optimal terminal decision rule d and reduces the original problem to the optimal stopping of the posterior probability process

$$\Pi(t) = (\Pi_1(t), \ldots, \Pi_M(t)),\ t \geq 0, \qquad \text{where} \qquad \Pi_i(t) = \mathbb{P}^{\vec{\pi}}\{\Theta = i | \mathcal{F}_t\},\ i \in I. \tag{2.2}$$

Proposition 2.1. *The smallest Bayes risk $V(\cdot)$ of (1.6) becomes*

$$V(\vec{\pi}) = \inf_{\tau \in \mathbb{F}} \mathbb{E}^{\vec{\pi}} \left[\tau + \mathbf{1}_{\{\tau < \infty\}} h\left(\Pi(\tau)\right) \right], \quad \vec{\pi} = (\pi_1, \ldots, \pi_M) \in E, \tag{2.3}$$

where $h : E \mapsto \mathbb{R}_+$ is the optimal terminal decision cost function given by

$$h(\vec{\pi}) \triangleq \min_{i \in I} h_i(\vec{\pi}) \quad \text{and} \quad h_i(\vec{\pi}) \triangleq \sum_{j=1}^{M} a_{ij} \pi_j. \tag{2.4}$$

Let $d(t) \in \operatorname{argmin}_{i \in I} \sum_{j=1}^M a_{ij} \Pi_j(t)$, $t \geq 0$. If an $\mathbb{F}$-stopping time τ^* solves the problem in (2.3), then $(\tau^*, d(\tau^*)) \in \mathcal{A}$ is an optimal admissible strategy for $V(\cdot)$ in (1.6).

Thus, the original Bayesian sequential multi-hypothesis testing simplifies to the optimal stopping problem in (2.3). To solve it, we shall first study sample paths of the process Π.

For every $i \in I$, let $p_i(\cdot)$ be the density of the probability distribution ν_i with respect to some σ-finite measure μ on $\mathbb{R}^d$ with Borel σ-algebra $\mathscr{B}(\mathbb{R}^d)$. For example, one can take $\mu = \nu_1 + \cdots + \nu_M$. In the appendix, we show that

$$\Pi_i(t) = \frac{\pi_i\, e^{-\lambda_i t} \prod_{k=1}^{N_t} \lambda_i\, p_i(Y_k)}{\sum_{j=1}^M \pi_j\, e^{-\lambda_j t} \prod_{k=1}^{N_t} \lambda_j\, p_j(Y_k)}, \quad t \geq 0 \quad \mathbb{P}^{\vec{\pi}}\text{-almost surely}, \; \vec{\pi} \in E, \; i \in I. \tag{2.5}$$

If we denote the jump times of the compound Poisson process X by

$$\sigma_n \triangleq \inf\{t > \sigma_{n-1} : X_t \neq X_{t-}\}, \quad n \geq 1 \qquad (\sigma_0 \equiv 0), \tag{2.6}$$

then $\mathbb{P}_i\{\sigma_1 \in dt_1, \ldots, \sigma_n \in dt_n, \sigma_{n+1} > t, Y_1 \in dy_1, \ldots, Y_n \in dy_n\}$ equals

$$\left[\lambda_i e^{-\lambda_i t_1} dt_1\right] \cdots \left[\lambda_i e^{-\lambda_i (t_n - t_{n-1})} dt_n\right] \left[e^{-\lambda_i (t-t_n)}\right] \prod_{k=1}^n p_i(y_k)\mu(dy_k) = e^{-\lambda_i t} \prod_{k=1}^n \lambda_i dt_k p_i(y_k)\mu(dy_k)$$

for every $n \geq 0$ and $0 < t_1 \leq \ldots \leq t_n \leq t$. Therefore, $\Pi_i(t)$ of (2.5) is the relative likelihood

$$\left.\frac{\pi_i \left[\lambda_i e^{-\lambda_i t_1} dt_1\right] \cdots \left[\lambda_i e^{-\lambda_i (t_n - t_{n-1})} dt_n\right] \left[e^{-\lambda_i (t-t_n)}\right] \prod_{k=1}^n p_i(y_k)\mu(dy_k)}{\sum_{j=1}^M \pi_j \left[\lambda_j e^{-\lambda_j t_1} dt_1\right] \cdots \left[\lambda_j e^{-\lambda_j (t_n - t_{n-1})} dt_n\right] \left[e^{-\lambda_j (t-t_n)}\right] \prod_{k=1}^n p_j(y_k)\mu(dy_k)}\right|_{\substack{n=N_t,\\ t_1=\sigma_1,\ldots,t_n=\sigma_n,\\ y_1=Y_1,\ldots,y_n=Y_n}}$$

of the path $X(s)$, $0 \leq s \leq t$ under hypothesis H_i as a result of Bayes rule. The explicit form in (2.5) also gives the recursive dynamics of the process Π as in

$$\left\{\begin{array}{l} \Pi(t) = x\left(t - \sigma_{n-1}, \Pi(\sigma_{n-1})\right), \quad t \in [\sigma_{n-1}, \sigma_n) \\ \Pi(\sigma_n) = \left(\dfrac{\lambda_1 p_1(Y_n)\Pi_1(\sigma_n-)}{\sum_{j=1}^M \lambda_j p_j(Y_n)\Pi_j(\sigma_n-)}, \; \ldots \;, \dfrac{\lambda_M p_M(Y_n)\Pi_M(\sigma_n-)}{\sum_{j=1}^M \lambda_j p_j(Y_n)\Pi_j(\sigma_n-)}\right) \end{array}\right\}, \quad n \geq 1 \tag{2.7}$$

in terms of the deterministic mapping $x : \mathbb{R} \times E \mapsto E$ defined by

$$x(t, \vec{\pi}) \equiv (x_1(t,\vec{\pi}), \ldots, x_M(t,\vec{\pi})) \triangleq \left(\frac{e^{-\lambda_1 t}\, \pi_1}{\sum_{j=1}^M e^{-\lambda_j t}\, \pi_j}, \; \ldots \;, \frac{e^{-\lambda_M t}\, \pi_M}{\sum_{j=1}^M e^{-\lambda_j t}\, \pi_j}\right). \tag{2.8}$$

This mapping satisfies the semi-group property $x(t+s, \vec{\pi}) = x(t, x(s, \vec{\pi}))$ for $s, t \geq 0$. Moreover, Π is a piecewise-deterministic process by (2.7). The process Π follows the deterministic curves $t \mapsto x(t, \vec{\pi})$, $\vec{\pi} \in E$ between arrival times of X and jumps from one curve to another at the arrival times $(\sigma_n)_{n \geq 1}$ of X; see Figure 1.

Since under every $\mathbb{P}_i$, $i \in I$ marks and interarrival times of the process X are i.i.d., and the latter are exponentially distributed, the process Π is a piecewise-deterministic $(\mathbb{P}_i, \mathbb{F})$-Markov

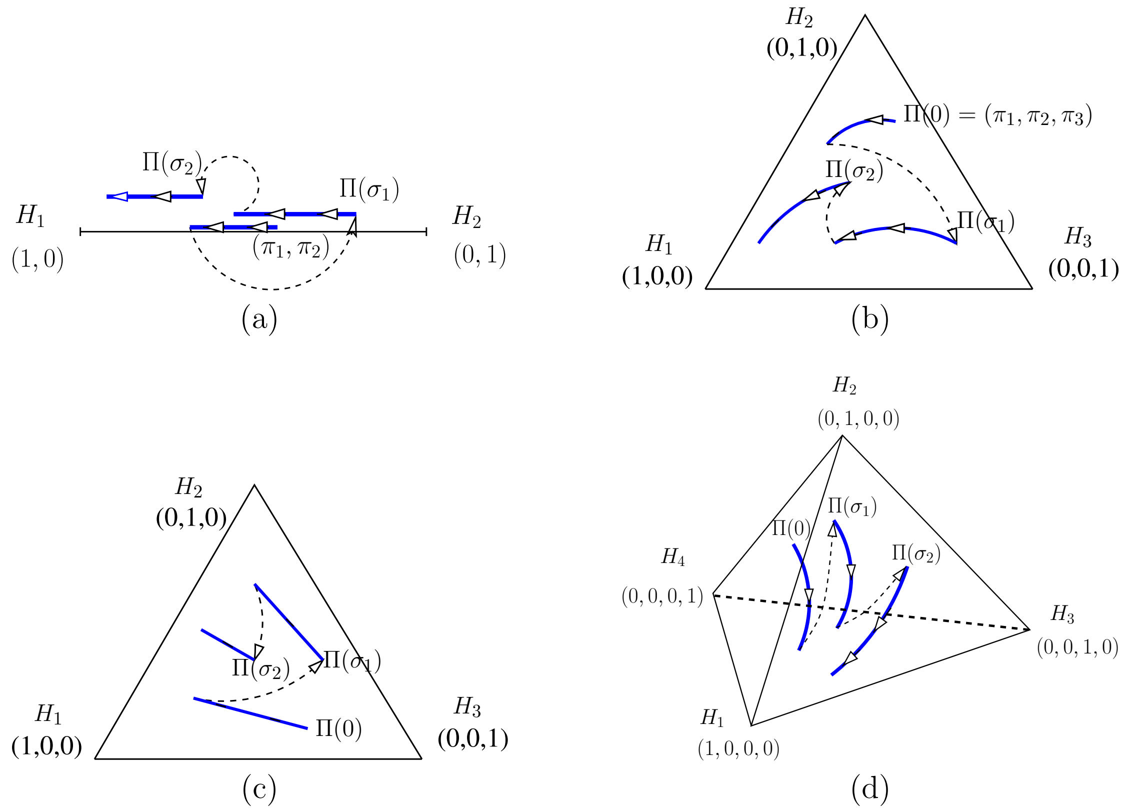


Figure 1: Sample paths $t \mapsto \Pi(t)$ when $M = 2$ in (a), $M = 3$ in (b) and (c), and $M = 4$ in (d). Between arrival times $\sigma_1, \sigma_2, \ldots$ of X, the process Π follows the deterministic curves $t \mapsto x(t, \vec{\pi})$, $\vec{\pi} \in E$. At the arrival times, it switches from one curve to another. In (a), (b), and (d), the arrival rates are different under hypotheses. In (c), $\lambda_1 = \lambda_2 < \lambda_3$, and (2.8) implies $\Pi_1(t)/\Pi_2(t) = \Pi_1(\sigma_n)/\Pi_2(\sigma_n)$ for every $t \in [\sigma_n, \sigma_{n+1})$, $n \geq 1$.

process for every $i \in I$. Therefore, as shown in the appendix,

$$\mathbb{E}^{\vec{\pi}}[g(\Pi(t+s))|\mathcal{F}_t] = \sum_{i=1}^{M} \Pi_i(t)\mathbb{E}_i[g(\Pi(t+s))|\mathcal{F}_t] = \sum_{i=1}^{M} \Pi_i(t)\mathbb{E}_i[g(\Pi(t+s))|\Pi(t)] \tag{2.9}$$

for every bounded Borel function $g : E \mapsto \mathbb{R}$, numbers $s, t \geq 0$, and $\vec{\pi} \in E$, and Π is also a *piecewise-deterministic* $(\mathbb{P}^{\vec{\pi}}, \mathbb{F})$*-Markov process* for every $\vec{\pi} \in E$. By (2.8), we have

$$\frac{\partial x_i(t, \vec{\pi})}{\partial t} = x_i(t, \vec{\pi}) \left(-\lambda_i + \sum_{j=1}^{M} \lambda_j x_j(t, \vec{\pi}) \right), \qquad i \in I, \tag{2.10}$$

and Jensen's inequality gives

$$\sum_{i=1}^{M} \lambda_i \frac{\partial x_i(t,\vec{\pi})}{\partial t} = -\sum_{i=1}^{M} \lambda_i^2 \, x_i(t,\vec{\pi}) + \left(\sum_{i=1}^{M} \lambda_i \, x_i(t,\vec{\pi}) \right)^2 \leq 0.$$

Therefore, the sum on the right hand side of (2.10) decreases in t, and next remark follows.

Remark 2.2. For every $i \in I$, there exists $t_i(\vec{\pi}) \in [0,\infty]$ such that the function $x_i(t,\vec{\pi})$ of (2.8) is increasing in t on $[0, t_i(\vec{\pi})]$ and decreasing on $[t_i(\vec{\pi}), \infty)$. Since $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$, (2.10) implies that $t_i(\cdot) = \infty$ for every $i \in I$ such that $\lambda_i = \lambda_1$. Similarly, $t_i(\cdot) = 0$ for every i such that $\lambda_i = \lambda_M$. In other words, as long as no arrivals are observed, the process Π give more and more weights on the hypotheses under which the arrival rate is the smallest.

On the other hand, if $\lambda_1 = \cdots = \lambda_M$, then $x(t,\vec{\pi}) = \vec{\pi}$ for every $t \geq 0$ and $\vec{\pi} \in E$ by (2.8). Hence by (2.7), Π does not change between arrivals, and the relative likelihood of hypotheses in (1.2) are updated only at arrival times according to the observed marks.

Finally, we conclude this section with the following lemma, which will be useful in describing a numerical algorithm in Section 4.2; see the appendix for its proof.

Lemma 2.3. Denote a "neighborhood of the corners" in E by

$$\widehat{E}_{(q_1,\ldots,q_M)} \triangleq \bigcup_{j=1}^{M} \{\vec{\pi} \in E \,:\, \pi_j \geq q_j\} \qquad \text{for every} \quad (q_1,\ldots,q_M) \in (0,1)^M. \tag{2.11}$$

If $\lambda_1 < \lambda_2 < \cdots < \lambda_M$ in (1.2), then for every $(q_1,\ldots,q_M) \in (0,1)^M$, the hitting time $\inf\{t \geq 0 \,:\, x(t,\vec{\pi}) \in \widehat{E}_{(q_1,\ldots,q_M)}\}$ of the path $t \mapsto x(t,\vec{\pi})$ to the set $\widehat{E}_{(q_1,\ldots,q_M)}$ is bounded uniformly in $\vec{\pi} \in E$ from above by some finite number $s\,(q_1,\ldots,q_M)$.

3. Successive approximations

By limiting the horizon of the problem in (2.3) to the nth arrival time σ_n of the process X, we obtain the family of optimal stopping problems

$$V_n(\pi_1,\ldots,\pi_M) \triangleq \inf_{\tau \in \mathbb{F}} \mathbb{E}^{\vec{\pi}} \left[\tau \wedge \sigma_n + h\left(\Pi(\tau \wedge \sigma_n)\right) \right], \qquad \vec{\pi} \in E, \quad n \geq 0, \tag{3.1}$$

where $h(\cdot)$ is the optimal terminal decision cost in (2.4). The functions $V_n(\cdot)$, $n \geq 1$ and $V(\cdot)$ are nonnegative and bounded since so is $h(\cdot)$. The sequence $\{V_n(\cdot)\}_{n \geq 1}$ is decreasing and has pointwise limit. Next proposition shows that it converges to $V(\cdot)$ uniformly on E.

Remark 3.1. In the proof of Proposition 3.2 and elsewhere, it will be important to remember that the infimum in (2.3) and (3.1) can be taken over stopping times whose $\mathbb{P}^{\vec{\pi}}$-expectation is bounded from above by $\sum_{i,j} a_{ij}$ for every $\vec{\pi} \in E$, since "immediate stopping" already gives the upper bound $h(\cdot) \le \sum_{i,j} a_{ij} < \infty$ on the value functions $V(\cdot)$ and $V_n(\cdot)$, $n \ge 0$.

Proposition 3.2. *The sequence $\{V_n(\cdot)\}_{n\ge 1}$ converges to $V(\cdot)$ uniformly on E. In fact,*

$$0 \le V_n(\vec{\pi}) - V(\vec{\pi}) \le \left(\sum_{i,j} a_{ij}\right)^{3/2} \sqrt{\frac{\lambda_M}{n-1}} \qquad \text{for every } \vec{\pi} \in E \text{ and } n \ge 1. \tag{3.2}$$

Proof. By definition, we have $V_n(\cdot) \ge V(\cdot)$. For every $\mathbb{F}$-stopping time τ, the expectation $\mathbb{E}^{\vec{\pi}}\left[\tau + \mathbf{1}_{\{\tau<\infty\}} h\left(\Pi(\tau)\right)\right]$ in (2.3) can be written as

$$\mathbb{E}^{\vec{\pi}}\left[\tau \wedge \sigma_n + h\left(\Pi(\tau\wedge\sigma_n)\right)\right] + \mathbb{E}^{\vec{\pi}}\left(\mathbf{1}_{\{\tau>\sigma_n\}}\left[\tau - \sigma_n + h\left(\Pi(\tau)\right) - h\left(\Pi(\sigma_n)\right)\right]\right)$$
$$\ge \mathbb{E}^{\vec{\pi}}\left[\tau \wedge \sigma_n + h\left(\Pi(\tau\wedge\sigma_n)\right)\right] - \sum_{i,j} a_{ij}\, \mathbb{E}^{\vec{\pi}}\, \mathbf{1}_{\{\tau>\sigma_n\}}.$$

We have $\mathbb{E}^{\vec{\pi}}\, \mathbf{1}_{\{\tau>\sigma_n\}} \le \mathbb{E}^{\vec{\pi}}\left[\mathbf{1}_{\{\tau>\sigma_n\}}\left(\tau/\sigma_n\right)^{1/2}\right] \le \sqrt{\mathbb{E}^{\vec{\pi}}\tau\, \mathbb{E}^{\vec{\pi}}\left(1/\sigma_n\right)}$ due to Cauchy-Schwartz inequality, and $\mathbb{E}^{\vec{\pi}}\left(1/\sigma_n\right) \le \lambda_M/(n-1)$ by (2.1). If $\mathbb{E}^{\vec{\pi}}\tau \le \sum_{i,j} a_{ij}$, then we obtain

$$\mathbb{E}^{\vec{\pi}}\left[\tau + \mathbf{1}_{\{\tau<\infty\}} h\left(\Pi(\tau)\right)\right] \ge \mathbb{E}^{\vec{\pi}}\left[\tau \wedge \sigma_n + h\left(\Pi(\tau\wedge\sigma_n)\right)\right] - \left(\sum_{i,j} a_{ij}\right)^{3/2} \sqrt{\frac{\lambda_M}{n-1}},$$

and taking the infimum of both sides over the $\mathbb{F}$-stopping times whose $\mathbb{P}^{\vec{\pi}}$-expectation is bounded by $\sum_{i,j} a_{ij}$ for every $\vec{\pi} \in E$ completes the proof by Remark 3.1. □

The dynamic programming principle suggests that the functions in (3.1) satisfy the relation $V_{n+1} = J_0 V_n$, $n \ge 0$, where J_0 is an operator acting on bounded functions $w : E \mapsto \mathbb{R}$ by

$$J_0 w(\vec{\pi}) \triangleq \inf_{\tau\in\mathbb{F}} \mathbb{E}^{\vec{\pi}}\left[\tau \wedge \sigma_1 + \mathbf{1}_{\{\tau<\sigma_1\}} h\left(\Pi(\tau)\right) + \mathbf{1}_{\{\tau\ge\sigma_1\}} w\left(\Pi(\sigma_1)\right)\right], \quad \vec{\pi} \in E. \tag{3.3}$$

We show that (3.3) is in fact a minimization problem over *deterministic* times, by using the next result about the characterization of stopping times of piecewise-deterministic Markov processes; see Brémaud (1981, Theorem T33, p. 308), Davis (1993, Lemma A2.3, p. 261)) for its proof.

Lemma 3.3. *For every $\mathbb{F}$-stopping time τ and every $n \ge 0$, there is an $\mathcal{F}_{\sigma_n}$-measurable random variable $R_n : \Omega \mapsto [0,\infty]$ such that $\tau \wedge \sigma_{n+1} = (\sigma_n + R_n) \wedge \sigma_{n+1}$ $\mathbb{P}^{\vec{\pi}}$-a.s. on $\{\tau \ge \sigma_n\}$.*

Particularly, if τ is an $\mathbb{F}$-stopping time, then there exists some constant $t = t(\vec{\pi}) \in [0,\infty]$ such that $\tau \wedge \sigma_1 = t \wedge \sigma_1$ holds $\mathbb{P}^{\vec{\pi}}$-a.s. Therefore, (3.3) becomes

$$J_0 w(\vec{\pi}) = \inf_{t\in[0,\infty]} \quad Jw(t,\vec{\pi}) \triangleq \mathbb{E}^{\vec{\pi}}\left[t\wedge\sigma_1 + \mathbf{1}_{\{t<\sigma_1\}} h\left(\Pi(t)\right) + \mathbf{1}_{\{t\geq\sigma_1\}} w\left(\Pi(\sigma_1)\right)\right]. \tag{3.4}$$

Since the first arrival time σ_1 of the process X is distributed under $\mathbb{P}^{\vec{\pi}}$ according to a mixture of exponential distributions with rates λ_i, $i \in I$ and weights π_i, $i \in I$, the explicit dynamics in (2.7) of the process Π gives

$$Jw(t,\vec{\pi}) = \int_0^t \sum_{i=1}^M \pi_i e^{-\lambda_i u}\left[1 + \lambda_i \cdot (S_i w)(x(u,\vec{\pi}))\right] du + \left(\sum_{i=1}^M \pi_i e^{-\lambda_i t}\right) \cdot h\left(x(t,\vec{\pi})\right), \tag{3.5}$$

where S_i is an operator acting on bounded functions $w : E \mapsto \mathbb{R}$ and is defined as

$$S_i w(\vec{\pi}) \triangleq \int_{\mathbb{R}^d} w\left(\frac{\lambda_1 p_1(y)\pi_1}{\sum_{j=1}^M \lambda_j p_j(y)\pi_j}, \ldots, \frac{\lambda_M p_M(y)\pi_M}{\sum_{j=1}^M \lambda_j p_j(y)\pi_j}\right) p_i(y)\mu(dy), \quad \vec{\pi}\in E,\ i \in I. \tag{3.6}$$

Hence, by (3.4) and (3.5) we conclude that the optimal stopping problem in (3.3) is essentially a deterministic minimization problem. Using the operator J_0, let us now define the sequence

$$v_0(\cdot) \triangleq h(\cdot) \qquad \text{and} \qquad v_{n+1}(\cdot) \triangleq J_0 v_n(\cdot), \quad n \geq 0. \tag{3.7}$$

One of the main results of this section is Proposition 3.7 below and shows that $V_n = v_n$, $n \geq 0$. Therefore, the solutions $(v_n)_{n\geq1}$ of deterministic minimization problems in (3.7) give the successive approximations $(V_n)_{n\geq1}$ in (3.1) of the function V in (2.3) by Proposition 3.2. That result hinges on certain properties of the operator J_0 and sequence $(v_n)_{n\geq1}$ summarized by the next proposition, whose proof is in the appendix.

Proposition 3.4. *For two bounded functions $w_1(\cdot) \leq w_2(\cdot)$ on E, we have $Jw_1(t,\cdot) \leq Jw_2(t,\cdot)$ for every $t \geq 0$, and $J_0 w_1(\cdot) \leq J_0 w_2(\cdot) \leq h(\cdot)$. If $w(\cdot)$ is concave, so are $Jw(t,\cdot)$, $t \geq 0$ and $J_0 w(\cdot)$. Finally, if $w : E \mapsto \mathbb{R}_+$ is bounded and continuous, so are $(t,\vec{\pi}) \mapsto J_0 w(t,\vec{\pi})$ and $\vec{\pi} \mapsto J_0 w(\vec{\pi})$.*

Corollary 3.5. *Each function $v_n(\cdot)$, $n \geq 0$ of (3.7) is continuous and concave on E. We have $h(\cdot) \equiv v_0(\cdot) \geq v_1(\cdot) \geq \ldots \geq 0$, and the pointwise limit $v(\cdot) \triangleq \lim_{n\to\infty} v_n(\cdot)$ exists and is concave on E.*

Proof. Since $v_0(\cdot) \equiv h(\cdot)$ and $v_n(\cdot) = J_0 v_{n-1}(\cdot)$, the claim on concavity and continuity follows directly from Proposition 3.4 by induction. To show the monotonicity, by construction we have $0 \leq v_0(\cdot) \equiv h(\cdot)$, therefore $0 \leq v_1(\cdot) \leq v_0(\cdot)$ again by Remark 3.4. Assume now that

$0 \le v_n(\cdot) \le v_{n-1}(\cdot)$. Applying the operator J_0 to both sides, we get $0 \le v_{n+1}(\cdot) = J_0 v_n(\cdot) \le J_0 v_{n-1}(\cdot) = v_n(\cdot)$. Hence the sequence is decreasing and the pointwise limit $\lim_{n\to\infty} v_n(\cdot)$ exists on E. Finally, since it is the lower envelop of concave functions v_n's, the function v is also concave. □

Remark 3.6. For $Jw(t,\cdot)$ given in (3.5) let us define

$$J_t w(\vec{\pi}) = \inf_{s\in[t,\infty]} Jw(s,\vec{\pi}), \tag{3.8}$$

which agrees with (3.4) for $t = 0$. For every bounded continuous function $w : E \mapsto \mathbb{R}_+$, the mapping $t \mapsto Jw(t,\vec{\pi})$ is again bounded and continuous on $[0,\infty]$, and the infimum in (3.8) attained for all $t \in [0,\infty]$.

Proposition 3.7. *For every $n \ge 0$, we have $v_n(\cdot) = V_n(\cdot)$ in (3.1) and (3.7). If for every $\varepsilon \ge 0$ we define*

$$r_n^\varepsilon(\vec{\pi}) \triangleq \inf\left\{s \in (0,\infty] : Jv_n\big(s,\vec{\pi}\big) \le J_0 v_n(\vec{\pi}) + \varepsilon\right\}, \quad n \ge 0,\ \vec{\pi} \in E,$$

$$S_1^\varepsilon \triangleq r_0^\varepsilon\big(\Pi(0)\big) \wedge \sigma_1, \quad \text{and} \quad S_{n+1}^\varepsilon \triangleq \left\{\begin{array}{ll} r_n^{\varepsilon/2}\big(\Pi(0)\big), & \text{if } \sigma_1 > r_n^{\varepsilon/2}\big(\Pi(0)\big) \\ \sigma_1 + S_n^{\varepsilon/2}\circ\theta_{\sigma_1}, & \text{if } \sigma_1 \le r_n^{\varepsilon/2}\big(\Pi(0)\big) \end{array}\right\}, \quad n \ge 1,$$

where θ_s is the shift-operator on Ω; i.e., $X_t \circ \theta_s = X_{s+t}$, then

$$\mathbb{E}^{\vec{\pi}}\left[S_n^\varepsilon + h(\Pi(S_n^\varepsilon))\right] \le v_n(\vec{\pi}) + \varepsilon, \qquad \forall\, \varepsilon \ge 0,\ n \ge 1,\ \vec{\pi} \in E. \tag{3.9}$$

Corollary 3.8. *Since $v_n(\cdot) = V_n(\cdot)$, $n \ge 0$, we have $v(\cdot) = V(\cdot)$ by Proposition 3.2. The function $V(\cdot)$ is continuous and concave on E by Proposition 3.2 and Corollary 3.5.*

Proposition 3.9. *The value function V of (2.3) satisfies $V = J_0 V$, and it is the largest bounded solution of $U = J_0 U$ smaller than or equal to h.*

Remark 3.10. If the arrival rates in (1.2) are identical, then $x(t,\vec{\pi}) = \vec{\pi}$ for every $t \ge 0$ and $\vec{\pi} \in E$; see Remark 2.2. If moreover $\lambda_1 = \cdots = \lambda_M = 1$, then (3.4)-(3.5) reduce to

$$J_0 w(\vec{\pi}) = \min\left\{h(\vec{\pi})\ ,\ 1 + \sum_{i=1}^M \pi_i \cdot (S_i w)(\vec{\pi})\right\}, \qquad \vec{\pi} \in E \tag{3.10}$$

for every bounded function $w : E \mapsto \mathbb{R}$. Blackwell and Girshick (1954, Chapter 9) show that $\lim_{n\to\infty} J_0^n h \equiv \lim_{n\to\infty} J_0(J_0^{n-1}h) \equiv \lim_{n\to\infty} J_0 v_n = v \equiv V$ gives the minimum Bayes risk of (1.6) in the *discrete-time* sequential Bayesian hypothesis testing problem, where (i) the expected sum of the number of samples $Y_1, Y_2, \ldots$ and terminal decision cost is minimized, and (ii) the unknown probability density function p of the Y_k's must be identified among the alternatives $p_1, \dots, p_M$.

Proposition 3.11. *For every bounded function $w : E \mapsto \mathbb{R}_+$, we have*

$$J_t w(\vec{\pi}) = Jw(t, \vec{\pi}) + \left(\sum_{i=1}^{M} \pi_i \, e^{-\lambda_i t} \right) \left[J_0 w\big(x(t, \vec{\pi})\big) - h\big(x(t, \vec{\pi})\big) \right], \tag{3.11}$$

where the operators J and J_t are defined in (3.5) and (3.8), respectively.

Proof. Let us fix a constant $s \geq t$ and $\vec{\pi} \in E$. Using the operator J in (3.5) and the semigroup property of $x(t, \vec{\pi})$ in (2.8), $Jw(s, \vec{\pi})$ can be written as

$$\begin{aligned} Jw(s, \vec{\pi}) = Jw(t, \vec{\pi}) &+ \sum_{i=1}^{M} \pi_i \, e^{-\lambda_i \, t} \int_t^s \sum_{i=1}^{M} x_i(t, \vec{\pi}) \, e^{-\lambda_i \, u} \left[1 + \lambda_i \cdot S_i w(x(t+u, x(t, \vec{\pi}))) \right] du \\ &- \left(\sum_{i=1}^{M} \pi_i \, e^{-\lambda_i \, t} \right) \cdot h\left(x(t, \vec{\pi}) \right) + \left(\sum_{i=1}^{M} \pi_i \, e^{-\lambda_i \, (u-t)} \, e^{-\lambda_i \, t} \right) \cdot h\left(x\big(s-t, x(t, \vec{\pi})\big) \right) \\ &= Jw(t, \vec{\pi}) + \left(\sum_{i=1}^{M} \pi_i \, e^{-\lambda_i \, t} \right) \left[Jw\big(s-t, x(t, \vec{\pi})\big) - h\big(x(t, \vec{\pi})\big) \right]. \end{aligned}$$

Taking the infimum above over $s \in [t, \infty]$ concludes the proof. ☐

Corollary 3.12. *If*

$$r_n(\vec{\pi}) \equiv r_n^{(0)}(\vec{\pi}) = \inf \left\{ s \in (0, \infty] : Jv_n\left(s, \vec{\pi}\right) = J_0 v_n(\vec{\pi}) \right\} \tag{3.12}$$

and $\inf \varnothing = \infty$ *as $r_n^{(\varepsilon)}(\vec{\pi})$ in Proposition 3.7 with $\varepsilon = 0$, then*

$$r_n(\vec{\pi}) = \inf \left\{ t > 0 : v_{n+1}\left(x(t, \vec{\pi})\right) = h\left(x(t, \vec{\pi})\right) \right\}. \tag{3.13}$$

Remark 3.13. Substituting $w = v_n$ in (3.11) gives the "dynamic programming equation" for the sequence $\{v_n(\cdot)\}_{n \geq 0}$

$$v_{n+1}(\vec{\pi}) = Jv_n(t, \vec{\pi}) + \left(\sum_{i=1}^{M} \pi_i \, e^{-\lambda_i t} \right) \Big[v_{n+1}\left(x(t, \vec{\pi})\right) - h\left(x(t, \vec{\pi})\right) \Big], \quad t \in [0, r_n(\vec{\pi})].$$

Moreover Corollary 3.8 and Proposition 3.11 give

$$J_t V(\vec{\pi}) = JV(t, \vec{\pi}) + \left(\sum_{i=1}^{M} \pi_i \, e^{-\lambda_i t} \right) \left[V\left(x(t, \vec{\pi})\right) - h\left(x(t, \vec{\pi})\right) \right], \quad t \in \mathbb{R}_+. \tag{3.14}$$

Similar to (3.12), let us define

$$r(\vec{\pi}) \triangleq \inf\{t > 0 : JV(t,\vec{\pi}) = J_0V(\vec{\pi})\}. \tag{3.15}$$

Then, similar arguments as in Corollary 3.12 lead to

$$r(\vec{\pi}) = \inf\{t > 0 : V(x(t,\vec{\pi})) = h\,(x(t,\vec{\pi}))\}, \tag{3.16}$$

and

$$V(\vec{\pi}) = JV(t,\vec{\pi}) + \left(\sum_{i=1}^{M} \pi_i\, e^{-\lambda_i t}\right) [V(x(t,\vec{\pi})) - h\,(x(t,\vec{\pi}))]\,, \quad t \in [0, r(\vec{\pi})]. \tag{3.17}$$

Remark 3.14. By Corollary 3.8, the function $V(\cdot)$ in (2.3) is continuous on E. Since $t \mapsto x(t,\vec{\pi})$ of (2.8) is continuous, the mapping $t \to V(x(t,\vec{\pi}))$ is continuous. Moreover, the paths $t \mapsto \Pi(t)$ follows the deterministic curves $t \mapsto x(t,\cdot)$ between two jumps. Hence the process $V\,(\Pi(t))$ has right-continuous paths with left limits.

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

$$U_\varepsilon \triangleq \inf\{t \geq 0 : V\,(\Pi(t)) - h\,(\Pi(t)) \geq -\varepsilon\}\,, \qquad \varepsilon \geq 0. \tag{3.18}$$

Then the regularity of the paths $t \mapsto \Pi_i(t)$ implies that

$$V\,(\Pi(U_\varepsilon)) - h\,(\Pi(U_\varepsilon)) \geq -\varepsilon \quad \text{on the event} \quad \{U_\varepsilon < \infty\}\,. \tag{3.19}$$

Proposition 3.15. *Let $M_t \triangleq t + V\,(\Pi(t))$, $t \geq 0$. For every $n \geq 0$ and $\varepsilon \geq 0$, we have $\mathbb{E}^{\vec{\pi}}[M_0] = \mathbb{E}^{\vec{\pi}}[M_{U_\varepsilon \wedge \sigma_n}]$; i.e.,*

$$V(\vec{\pi}) = \mathbb{E}^{\vec{\pi}}\left[U_\varepsilon \wedge \sigma_n + V\,(\Pi(U_\varepsilon \wedge \sigma_n))\right]. \tag{3.20}$$

Proposition 3.16. *The stopping time U_ε in (3.18) and N_{U_ε} have bounded $\mathbb{P}^{\vec{\pi}}$-expectations for every $\varepsilon \geq 0$ and $\vec{\pi} \in E$. Moreover, it is ε-optimal for the problem in (2.3); i.e.,*

$$\mathbb{E}^{\vec{\pi}}\left[U_\varepsilon + \mathbf{1}_{\{U_\varepsilon<\infty\}} h\,(\Pi(U_\varepsilon))\right] \leq V(\vec{\pi}) + \varepsilon \quad \textit{for every} \quad \vec{\pi} \in E.$$

Proof. To show the first claim, note that by Proposition 3.5 and Corollary 3.8 we have $\sum_{i,j} a_{ij} \geq h(\vec{\pi}) \geq V(\vec{\pi}) \geq 0$. Using Proposition 3.15 above we have

$$\sum_{i,j} a_{ij} \geq V(\vec{\pi}) = \mathbb{E}^{\vec{\pi}}\left[U_\varepsilon \wedge \sigma_n + V\,(\Pi(U_\varepsilon \wedge \sigma_n))\right] \geq \mathbb{E}^{\vec{\pi}}\left[U_\varepsilon \wedge \sigma_n\right],$$

and by monotone convergence theorem it follows that $\sum_{i,j} a_{ij} \geq \mathbb{E}^{\vec{\pi}}\left[U_\varepsilon\right]$.

Also note that the process $N_t - \sum_{i=1}^M \Pi_i(t)(\lambda_i t)$, $t \geq 0$ is a $\mathbb{P}^{\vec{\pi}}$-martingale. Then using the stopped martingale we obtain

$$\mathbb{E}^{\vec{\pi}}\left[N_{U_\varepsilon \wedge t}\right] = \mathbb{E}^{\vec{\pi}}\left[\sum_{i=1}^M \Pi_i(U_\varepsilon \wedge t) \cdot \lambda_i \cdot (U_\varepsilon \wedge t)\right] \leq \lambda_M \cdot \mathbb{E}^{\vec{\pi}} U_\varepsilon \leq \lambda_M \sum_{i,j} a_{ij}.$$

Letting $t \to \infty$ and applying monotone convergence theorem we get $\mathbb{E}^{\vec{\pi}} N_{U_\varepsilon} \leq \lambda_M \sum_{i,j} a_{ij}$.

Next, the almost-sure finiteness of U_ε implies

$$V(\vec{\pi}) = \lim_{n\to\infty} \mathbb{E}^{\vec{\pi}}\left[U_\varepsilon \wedge \sigma_n + V\left(\Pi(U_\varepsilon \wedge \sigma_n)\right)\right] = \mathbb{E}^{\vec{\pi}}\left[U_\varepsilon + V\left(\Pi(U_\varepsilon)\right)\right],$$

by the monotone and bounded convergence theorems, and by Proposition 3.15. Since $V\left(\Pi(U_\varepsilon)\right) - h\left(\Pi(U_\varepsilon)\right) \geq -\varepsilon$ by (3.19) we have

$$V(\vec{\pi}) \geq \mathbb{E}^{\vec{\pi}}\left[U_\varepsilon + V\left(\Pi(U_\varepsilon)\right) - h\left(\Pi(U_\varepsilon)\right) + h\left(\Pi(U_\varepsilon)\right)\right] \geq \mathbb{E}^{\vec{\pi}}\left[U_\varepsilon + h\left(\Pi(U_\varepsilon)\right)\right] - \varepsilon$$

and the proof is complete. □

Corollary 3.17. *Taking $\varepsilon = 0$ in Proposition 3.16 implies that U_0 is an optimal stopping time for the problem in (2.3).*

4. Solution

Let us introduce the *stopping region* $\Gamma_\infty \triangleq \{\vec{\pi} \in E : V(\phi) = h(\phi)\}$ and *continuation region* $C_\infty \triangleq E \setminus \Gamma_\infty$ for the problem in (2.3). Because $h(\vec{\pi}) = \min_{i\in I} h_i(\vec{\pi})$, we have

$$\Gamma_\infty = \cup_{i\in I}\Gamma_{\infty,i}, \qquad \text{where} \quad \Gamma_{\infty,i} \triangleq \{\vec{\pi} \in E : V(\vec{\pi}) = h_i(\vec{\pi})\},\ i \in I. \tag{4.1}$$

According to Proposition 2.1 and Corollary 3.17, an admissible strategy $(U_0, d(U_0))$ that attains the minimum Bayes risk in (1.6) is (i) to observe X until the process Π of (2.7)-(2.8) enters the stopping region Γ_∞, and then (ii) to stop the observation and select hypothesis H_i if $d(U_0) = i$, equivalently, if $\Pi(U_0)$ is in $\Gamma_{\infty,i}$ for some $i \in I$.

In order to implement this strategy one needs to calculate, at least approximately, the subsets $\Gamma_{\infty,i}$, $i \in I$ or the function $V(\cdot)$. In Sections 4.1 and 4.2 we describe numerical methods that give strategies whose Bayes risks are within any given positive margin of the minimum, and the following structural results and observations are needed along the way.

Because the functions $V(\cdot)$ and $h_i(\cdot)$, $i \in I$ are continuous by Corollary 3.8 and (2.4), the sets $\Gamma_{\infty,i}$, $i \in I$ are closed. They are also convex, since both $V(\cdot)$ is convex by the same corollary, and $h(\cdot)$ is affine. Indeed, if we fix $i \in I$, $\vec{\pi}_1, \vec{\pi}_2 \in \Gamma_{\infty,i}$, and $\alpha \in [0,1]$, then

$V(\alpha\vec{\pi}_1+(1-\alpha)\vec{\pi}_2) \geq \alpha V(\vec{\pi}_1)+(1-\alpha)V(\vec{\pi}_2) = \alpha h_i(\vec{\pi}_1)+ (1-\alpha)h_i(\vec{\pi}_2) = h_i(\alpha\vec{\pi}_1+(1-\alpha)\vec{\pi}_2) \geq V(\alpha\vec{\pi}_1+(1-\alpha)\vec{\pi}_2)$. Hence, $V(\alpha\vec{\pi}_1+(1-\alpha)\vec{\pi}_2) = h_i(\alpha\vec{\pi}_1+(1-\alpha)\vec{\pi}_2)$ and $\alpha\vec{\pi}_1+(1-\alpha)\vec{\pi}_2 \in \Gamma_{\infty,i}$.

Proposition 3.9 and next proposition, whose proof is in the appendix, show that the stopping region Γ_∞ always includes a nonempty open neighborhood of every corner of the $(M-1)$-simplex E. However, some of the sets $\Gamma_{\infty,i}$, $i \in I$ may be empty in general, unless $a_{ij} > 0$ for every $1 \leq i \neq j \leq M$, in which case for some $\overline{p}_i < 1$, $i \in I$ we have $\Gamma_{\infty,i} \supseteq \{\vec{\pi} \in E : \pi_i \geq \overline{p}_i\} = \varnothing$ for every $i \in I$.

Proposition 4.1. *Let $w : E \mapsto \mathbb{R}_+$ be a bounded function. For every $i \in I$, define*

$$\overline{\pi}_i \triangleq \left[1 - 1/(2\,\lambda_M \max_k a_{ik})\right]^+, \qquad \text{and whenever } \lambda_i > \lambda_1 \tag{4.2}$$

$$\pi_i^* \triangleq \inf\left\{\pi \in [\overline{\pi}_i, 1) : \frac{\pi}{\lambda_i(1-\pi)}\left[1 - \left(\frac{1-\overline{\pi}_i}{\overline{\pi}_i}\cdot\frac{\pi}{1-\pi}\right)^{-\lambda_i/(\lambda_i-\lambda_1)}\right] \geq \max_k a_{ik}\right\}. \tag{4.3}$$

Then $\overline{p} \triangleq (\max_{i:\,\lambda_i=\lambda_1} \overline{\pi}_i) \vee (\max_{i:\,\lambda_i>\lambda_1} \pi_i^*) < 1$, and

$$\{\vec{\pi} \in E : J_0 w(\vec{\pi}) = h(\vec{\pi})\} \supseteq \bigcup_{i=1}^M \{\vec{\pi} \in E : \pi_i \geq \overline{p}\}, \tag{4.4}$$

and if $a_{ij} > 0$ for every $1 \leq i \neq j \leq M$, then

$$\{\vec{\pi} \in E : J_0 w(\vec{\pi}) = h_i(\vec{\pi})\} \supseteq \left\{\vec{\pi} \in E : \pi_i \geq \overline{p} \vee \frac{\max_k a_{ik}}{(\min_{j\neq i} a_{ji}) + (\max_k a_{ik})}\right\},\ i \in I. \tag{4.5}$$

For the auxiliary problems in (3.1), let us define the stopping and continuation regions as $\Gamma_n \triangleq \{\vec{\pi} \in E : V_n(\phi) = h(\phi)\}$ and $C_n \triangleq E \setminus \Gamma_n$, respectively, for every $n \geq 0$. Similar arguments as above imply that

$$\Gamma_n = \bigcup_{i=1}^M \Gamma_{n,i}, \qquad \text{where} \quad \Gamma_{n,i} \triangleq \{\vec{\pi} \in E : V_n(\vec{\pi}) = h_i(\vec{\pi})\}, i \in I$$

are closed and convex subsets of E. Since $V(\cdot) \leq \ldots \leq V_1(\cdot) \leq V_0(\cdot) \equiv h(\cdot)$, we have

$$\begin{gathered}\bigcup_{i=1}^M \{\vec{\pi} \in E : \pi_i \geq \pi_i^*\} \subseteq \Gamma_\infty \subseteq \cdots \subseteq \Gamma_n \subseteq \Gamma_{n-1} \subseteq \ldots \subseteq \Gamma_1 \subseteq \Gamma_0 \equiv E, \\ \{\vec{\pi} \in E : \pi_1 < \pi_1^*, \ldots, \pi_M < \pi_M^*\} \supseteq C \supseteq \cdots \supseteq C_n \supseteq C_{n-1} \supseteq \cdots \supseteq C_1 \supseteq C_0 \equiv \varnothing.\end{gathered} \tag{4.6}$$

Remark 4.2. If $J_0 h(\cdot) = h(\cdot)$, then (3.7) and Corollary 3.8 imply $V(\cdot) = h(\cdot)$, and that it is optimal to stop immediately and select the hypothesis that gives the smallest terminal cost. A sufficient condition for this is that $1/\lambda_i \geq \max_{k\in I} a_{ki}$ for every $i \in I$; i.e., the expected cost of (or waiting-time for) an additional observation is higher than the maximum cost of a wrong decision. Note that $h(\cdot) \geq J_0 h(\cdot)$ by definition of the operator J_0 in (3.4), and since $h(\cdot) \geq 0$, we have

$$J_0 h(\vec{\pi}) \geq \inf_{t\in[0,\infty]} \left[\int_0^t \sum_{i=1}^M \pi_i e^{-\lambda_i s} ds + \left(\sum_{i=1}^M \pi_i e^{-\lambda_i s} \right) h\left(x(t,\vec{\pi})\right) \right]$$
$$= \inf_{k,t} \left[\int_0^t \sum_{i=1}^M \pi_i e^{-\lambda_i s} ds + \left(\sum_{i=1}^M \pi_i e^{-\lambda_i s} \right) h_k\left(x(t,\vec{\pi})\right) \right] = \inf_{k,t} \sum_{i=1}^M \pi_i \left[\frac{1}{\lambda_i} + \left(a_{k,i} - \frac{1}{\lambda_i} \right) e^{-\lambda_i t} \right].$$

If $1/\lambda_i \geq \max_k a_{ki}$ for every $i \in I$, then last double minimization is attained when $t = 0$ and becomes $h(\vec{\pi})$; therefore, we obtain $h(\cdot) \geq J_0 h(\cdot) \geq h(\cdot)$; i.e., immediate stopping is optimal.

4.1. **Nearly-optimal strategies.** In Section 3, the value function $V(\cdot)$ of (1.6) is approximated by the sequence $\{V_n(\cdot)\}_{n\geq 1}$, whose elements can be obtained by applying the operator J_0 successively to the function $h(\cdot)$ in (2.4); see (3.7) and Proposition 3.7. According to Proposition 3.2, $V(\cdot)$ can be approximated this way within any desired positive level of precision after finite number of iterations. More precisely, (3.2) implies that

$$N \geq 1 + \frac{\lambda_M}{\varepsilon^2} \left(\sum_{i,j} a_{ij} \right)^3 \quad \Longrightarrow \quad ||V_N - V|| \triangleq \sup_{\vec{\pi}\in E} |V_N(\vec{\pi}) - V(\vec{\pi})| \leq \varepsilon, \ \forall\, \varepsilon > 0. \tag{4.7}$$

In Section 4.2, we give a numerical algorithm to calculate the functions $V_1, V_2, \ldots$ successively. By using those functions, we describe here two ε-optimal strategies.

Recall from Proposition 3.7 the ε-optimal stopping times S_n^ε for the truncated problems $V_n(\cdot)$, $n \geq 1$ in (3.1). For a given $\varepsilon > 0$, if we fix N by (4.7) such that $||V_N - V|| \leq \varepsilon/2$, then $S_N^{\varepsilon/2}$ is ε-optimal for $V(\cdot)$ in the sense that

$$\mathbb{E}^{\vec{\pi}} \left[S_N^{\varepsilon/2} + h\left(\Pi(S_N^{\varepsilon/2}) \right) \right] \leq V_N(\vec{\pi}) + \varepsilon/2 \leq V(\vec{\pi}) + \varepsilon \qquad \text{for every } \vec{\pi} \in E,$$

and $\big(S_N^{\varepsilon/2}, d(S_N^{\varepsilon/2})\big)$ is an ε-optimal Bayes strategy for the sequential hypothesis testing problem of (1.6); recall the definition of $d(\cdot)$ from Proposition 2.1.

As described by Proposition 3.7, the stopping rule $S_N^{\varepsilon/2}$ requires that we wait until the least of $r_{N-1}^{\varepsilon/4}(\vec{\pi})$ and first jump time σ_1 of X. If $r_{N-1}^{\varepsilon/4}(\vec{\pi})$ comes first, then we stop and select hypothesis H_i, $i \in I$ that gives the smallest terminal cost $h_i\big(x\big(r_{N-1}^{\varepsilon/4}(\vec{\pi}), \vec{\pi}\big)\big)$. Otherwise, we update the posterior probabilities at the first jump time to $\Pi(\sigma_1)$ and wait until the least of $r_{N-2}^{\varepsilon/4}(\Pi(\sigma_1))$ and next jump time $\sigma_1 \circ \theta_{\sigma_1}$ of the process X. If $r_{N-2}^{\varepsilon/4}(\vec{\pi})$ comes first, then we stop and select a hypothesis, or we wait as before. We stop at the Nth jump time of the process X and select the best hypothesis if we have not stopped earlier.

Let N again be an integer as in (4.7) with $\varepsilon/2$ instead of ε. Then $\big(U^{(N)}_{\varepsilon/2}, d(U^{(N)}_{\varepsilon/2})\big)$ is another ε-optimal strategy, if we define

$$U^{(N)}_{\varepsilon/2} \triangleq \inf\{t \geq 0; \quad h(\Pi(t)) \leq V_N(\Pi(t)) + \varepsilon/2\}. \tag{4.8}$$

Indeed, $\mathbb{E}^{\vec{\pi}}\left[U^{(N)}_{\varepsilon/2} + h\left(\Pi(U^{(N)}_{\varepsilon/2})\right)\right] \leq V_N(\vec{\pi}) + \varepsilon/2 \leq V(\vec{\pi}) + \varepsilon$ by the arguments similar in the proof of Proposition 3.16. After we calculate $V_N(\cdot)$ as in Section 4.2, this strategy requires that we observe X until the process Π enters the region $\{\vec{\pi} \in E : V_N(\vec{\pi}) - h(\vec{\pi}) \geq -\varepsilon/2\}$ and then select the hypothesis with the smallest terminal cost as before.

4.2. **Computing the successive approximations.** For the implementation of the nearly-optimal strategies described above, we need to compute $V_1(\cdot), V_2(\cdot), \ldots, V_N(\cdot)$ in (3.1) for any given integer $N \geq 1$.

If the arrival rates in (1.2) are distinct; i.e., $\lambda_1 < \lambda_2 < \ldots < \lambda_M$, then by Lemma 2.3 the entrance time $t_{\vec{\pi}}(\pi_1^*, \ldots, \pi_M^*)$ of the path $t \mapsto x(t, \vec{\pi})$ in (2.8) to the region $\cup_{i=1}^M \{\vec{\pi} \in E : \pi_i \geq \pi_i^*\}$ defined in Proposition 4.1 is bounded uniformly in $\vec{\pi} \in E$ from above by some finite number $s(\pi_1^*, \ldots, \pi_M^*)$. Therefore, the minimization in the problem $V_{n+1}(\vec{\pi}) = J_0 V_n(\vec{\pi}) = \inf_{t \in [0,\infty]} JV_n(t, \vec{\pi})$ can be restricted to the compact interval $[0, t_{\vec{\pi}}(\pi_1^*, \ldots, \pi_M^*)] \subset [0, s(\pi_1^*, \ldots, \pi_M^*)]$ for every $n \geq 1$ and $\vec{\pi} \in E$, thanks to Corollary 3.12. On the other hand, if the arrival rates $\lambda_1 = \ldots = \lambda_M$ are the same, then this minimization problem over $t \in [0, \infty]$ reduces to a simple comparison of two numbers as in (3.10) because of degenerate operator J_0; see Remark 3.10.

If some of the arrival rates are equal, then $t_{\vec{\pi}}(\pi_1^*, \ldots, \pi_M^*)$ is not uniformly bounded in $\vec{\pi} \in E$ anymore; in fact, it can be infinite for some $\vec{\pi} \in E$. Excluding the simple case of identical arrival rates, one can still restrict the minimization in $\inf_{t \in [0,\infty]} JV_n((t, \vec{\pi})$ to a compact interval independent of $\vec{\pi} \in E$ and still control the difference arising from this. Note that the operator J defined in (3.5) satisfies $\sup_{\vec{\pi} \in E} |Jw(t, \vec{\pi}) - Jw(\infty, \vec{\pi})| \leq (1/\lambda_1) e^{-\lambda_1 t}\left[1 + (\lambda_1 + \lambda_M) \sum_{i,j} a_{ij}\right]$ for every $t \geq 0$ and $w : E \mapsto \mathbb{R}_+$ such that $\sup_{\vec{\pi} \in E} w(\vec{\pi}) \leq \sum_{i,j} a_{ij}$; recall Remark 3.1. If we define

$$t(\delta) \triangleq -\frac{1}{\lambda_1} \ln\left(\frac{(\delta/2)\,\lambda_1}{1 + (\lambda_1 + \lambda_M)\sum_{i,j} a_{ij}}\right) \qquad \text{for every } \delta > 0, \tag{4.9}$$

$$J_{0,t} w(\vec{\pi}) \triangleq \inf_{s \in [0,t]} Jw(s, \vec{\pi}), \quad \text{for every bounded } w : E \mapsto \mathbb{R},\ t \geq 0,\ \vec{\pi} \in E, \tag{4.10}$$

then for every $\delta > 0$, $t_1, t_2 \geq t(\delta)$ and $\vec{\pi} \in E$, we have $|Jw(t_1, \vec{\pi}) - Jw(t_2, \vec{\pi})| \leq |Jw(t_1, \vec{\pi}) - Jw(\infty, \vec{\pi})| + |Jw(\infty, \vec{\pi}) - Jw(t_2, \vec{\pi})| \leq \delta/2 + \delta/2 \leq \delta$, and therefore, $\sup_{\vec{\pi} \in E} \big| J_{0,t(\delta)} w(\vec{\pi}) -$

$J_0 w(\vec{\pi})\big| < \delta$. Let us now define

$$V_{\delta,0}(\cdot) \triangleq h(\cdot) \quad \text{and} \quad V_{\delta,n+1}(\cdot) \triangleq J_{0,t(\delta)} V_{\delta,n}(\cdot), \qquad n \geq 1,\ \delta > 0. \tag{4.11}$$

Lemma 4.3. *For every $\delta > 0$ and $n \geq 0$, we have $V_n(\cdot) \leq V_{\delta,n}(\cdot) \leq n\delta + V_n(\cdot)$*

Proof. The inequalities holds for $n = 0$, since $V_{\delta,0}(\cdot) = V_0(\cdot) = h(\cdot)$ by definition. Suppose that they hold for some $n \geq 0$. Then $V_{n+1}(\cdot) = J_0 V_n(\cdot) \leq J_0 V_{\delta,n}(\cdot) \leq J_{0,t(\delta)} V_{\delta,n}(\cdot) = V_{\delta,n+1}(\cdot)$, which proves the first inequality for $n+1$. To prove the second inequality, note that $V_{\delta,n}(\cdot)$ is bounded from above by $h(\cdot) \leq \sum_{i,j} a_{ij}$. Then by (4.10) we have for every $\vec{\pi} \in E$ that

$$V_{\delta,n+1}(\vec{\pi}) = \inf_{t \in [0,t(\delta)]} JV_{\delta,n}(t,\vec{\pi}) \leq \inf_{t \in [0,\infty]} JV_{\delta,n}(t,\vec{\pi}) + \delta \leq \inf_{t \in [0,\infty]} JV_n(t,\vec{\pi})$$
$$+ \int_0^t \sum_{i=1}^M \pi_i\, e^{-\lambda_i u} \lambda_i\, n\delta\, du + \delta \leq V_{n+1}(t,\vec{\pi}) + \int_0^\infty \sum_{i=1}^M \pi_i\, e^{-\lambda_i u} \lambda_i\, n\delta\, du + \delta = V_{n+1}(\vec{\pi}) + (n+1)\delta,$$

and the proof is complete by induction on $n \geq 0$. □

If some (but not all) of the arrival rates are equal, then Lemma 4.3 lets us approximate the function $V(\cdot)$ of (2.3) by the functions $\{V_{\delta,n}(\cdot)\}_{\delta>0, n\geq 1}$ in (4.11). In this case, there is additional loss in precision due to truncation at δ, and this is compensated by increasing the number of iterations. For example, for a every $\varepsilon > 0$ the choices of N and $\delta > 0$ such that

$$\begin{gathered} N \geq 1 + \frac{1}{\varepsilon^2} \left[1 + \sqrt{\lambda_M} \left(\sum_{i,j} a_{ij} \right)^{3/2} \right]^2 \quad \text{and} \quad \delta \leq \frac{1}{N\sqrt{N-1}} \quad \text{imply that} \\ ||V_{\delta,N} - V|| \leq ||V_{\delta,N} - V_N|| + ||V_N - V|| \leq N\delta + \left(\sum_{i,j} a_{ij} \right)^{3/2} \sqrt{\frac{\lambda_M}{N-1}} \leq \varepsilon \end{gathered} \tag{4.12}$$

by (3.2) and Lemma 4.3. Hence, the function $V(\cdot)$ in (2.3) can be calculated at any desired precision level $\varepsilon > 0$ after some finite $N = N(\varepsilon)$ number of applications of the operator $J_{0,t(\delta)}$ in (4.10) for some $\delta = \delta(N) > 0$ to the function $h(\cdot)$ of (2.4), and one can choose

$$N(\varepsilon) \triangleq \text{smallest integer greater than or equal to } \left[1 + \sqrt{\lambda_M} \Big(\sum_{i,j} a_{ij} \Big)^{3/2} \right]^2, \tag{4.13}$$

$$\delta(N) \triangleq \frac{1}{N\sqrt{N-1}} \qquad \text{for every } \varepsilon > 0 \text{ and } N > 1. \tag{4.14}$$

If we set $N = N(\varepsilon/2)$ and $\delta = \delta(N)$, and define $U^{(\delta,N)}_{\varepsilon/2} \triangleq \inf\{t \geq 0;\ h(\Pi(t)) \leq V_{\delta,N}(\Pi(t)) + \varepsilon/2\}$, then for every $\vec{\pi} \in E$ we have $\mathbb{E}^{\vec{\pi}}\left[U^{(\delta,N)}_{\varepsilon/2} + h\left(\Pi(U^{(\delta,N)}_{\varepsilon/2}) \right) \right] \leq V_{\delta,N}(\vec{\pi}) + \varepsilon/2 \leq V(\vec{\pi}) + \varepsilon$

Initialization. To approximate V by V_N or $V_{\delta,N}$ such that $||V_N - V|| \le \varepsilon$ or $||V_{\delta,N} - V|| \le \varepsilon$ for some $\varepsilon > 0$, choose the number of iterations N and truncation parameter $\delta > 0$ as follows:

- If $\lambda_1 < \lambda_2 < \cdots < \lambda_M$ or $\lambda_1 = \cdots = \lambda_M$, then take $N \ge 1 + (\lambda_M/\varepsilon^2)(\sum_{i,j} a_{ij})^3$.
- Else take $N \ge 1 + (1/\varepsilon^2)[1 + \sqrt{\lambda_M}(\sum_{i,j} a_{ij})^{3/2}]^2$ and $0 < \delta \le 1/(N\sqrt{N-1})$.

By using bisection search, find π_i^* for $i \in I$ by using (4.3). Set $n = 0$, $V_0(\cdot) \equiv h(\cdot)$ or $V_{\delta,0}(\cdot) \equiv h(\cdot)$.

Repeat

- If $\lambda_1 < \lambda_2 < \cdots < \lambda_M$, then calculate $V_{n+1}(\vec{\pi}) = \min_{t \in [0, t_{\vec{\pi}}(\pi_1^*, \ldots, \pi_M^*)]} JV_n(t, \vec{\pi})$ for every $\vec{\pi} \in E$, where $t_{\vec{\pi}}(\pi_1^*, \ldots, \pi_M^*)$ is the hitting time of the path $t \mapsto x(t, \vec{\pi})$ in (2.8) to the region $\widehat{E}_{(\pi_1^*, \ldots, \pi_M^*)}$ in (2.11).
- Else if $\lambda_1 = \cdots = \lambda_M$, calculate $V_{n+1}(\vec{\pi}) = \min\{h(\vec{\pi}), (1/\lambda_1) + \sum_i \pi_i \cdot (S_i V_n)(\vec{\pi})\}$ as in (3.10) for every $\vec{\pi} \in E$.
- Else, calculate $V_{\delta,n+1}(\vec{\pi}) = \inf_{t \in [0, t(\delta)]} JV_{\delta,n}(t, \vec{\pi})$ for every $\vec{\pi} \in E$, where $t(\delta)$ is given by (4.9).
- Increase n by one.

Until $n \ge N$.

Figure 2: Numerical algorithm that approximates the value function $V(\cdot)$ in (2.3).

as in the proof of Proposition 3.16. As in the case of (4.8), observing the process X until the stopping time $U_{\varepsilon/2}^{(\delta,N)}$ and then selecting a hypothesis with the smallest expected terminal cost give an ε-optimal strategy for the $V(\cdot)$ in (1.6).

Figure 2 summarizes our discussion in the form of an algorithm that computes the elements of the sequence $\{V_n(\cdot)\}_{n \ge 1}$ or $\{V_{\delta,n}(\cdot)\}_{n \ge 1}$ and approximates the function $V(\cdot)$ in (2.3) numerically. This algorithm is used in the numerical examples of Section 5.

5. Examples

Here we solve several examples by using the numerical methods of previous section. We restrict ourselves to the problems with $M = 3$ and 4 alternative hypotheses since the drawings of the value functions and stopping regions are then possible and help checking numerically Section 4's theoretical conclusions on their structures; compare them also to Dayanik and Sezer's (2005) examples with $M = 2$ alternative hypotheses. In each example, an approximate value function is calculated by using the numerical algorithm in Figure 2 with a negligible ε-value.

Figure 3 displays the results of two examples on testing unknown arrival rate of a simple Poisson process. In both examples, the terminal decision costs $a_{ij} = 2$, $i \ne j$ are the same for wrong terminal decisions.

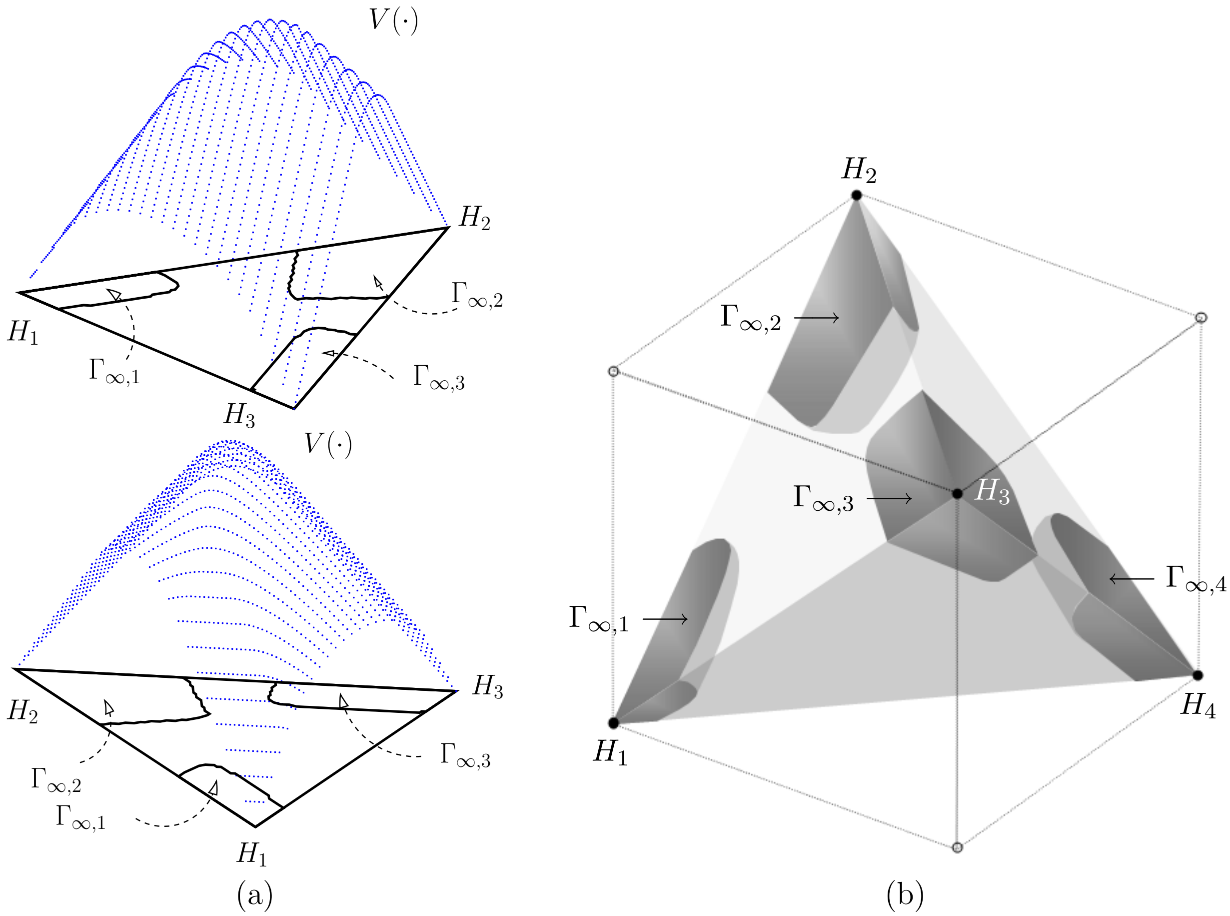


Figure 3: Testing hypotheses $H_1 : \lambda = 5$, $H_2 : \lambda = 10$, $H_3 : \lambda = 15$ in (a) and $H_1 : \lambda = 5$, $H_2 : \lambda = 10$, $H_3 : \lambda = 15$, $H_4 : 20$ in (b) about the unknown arrival rate of a simple Poisson process. Wrong terminal decision costs are $a_{ij} = 2$ for $i \neq j$ in both examples. In (a), the approximations of the function $V(\cdot)$ in (1.6)-(2.3) and stopping regions $\Gamma_{\infty,1}, \ldots, \Gamma_{\infty,3}$ in (4.1) are displayed from two perspectives. In (b), the stopping regions $\Gamma_{\infty,1}, \ldots, \Gamma_{\infty,4}$ are displayed by the darker regions in 3-simplex E. Observe that in (b) π_4 equals zero on the upper left $H_1H_2H_3$-face of the tetrahedron, and the stopping regions on that face coincide with those of the example in (a) as expected.

In Figure 3 (a), three alternative hypotheses $H_1 : \lambda = 5$, $H_2 : \lambda = 10$, $H_3 : \lambda = 15$ are tested sequentially. The value function and boundaries between stopping and continuation regions are drawn from two perspectives. As noted in Section 4, the stopping region consists of convex and closed neighborhoods of the corners of 2-simplex E at the bottom of the figures on the left. The value function above E looks concave and continuous in agreement with Corollary 3.8. In Figure 3 (b), four hypotheses $H_1 : \lambda = 5$, $H_2 : \lambda = 10$, $H_3 : \lambda = 15$, and $H_4 : \lambda = 20$ are tested sequentially and stopping regions $\Gamma_{\infty,1}, \ldots, \Gamma_{\infty,4}$ are displayed on

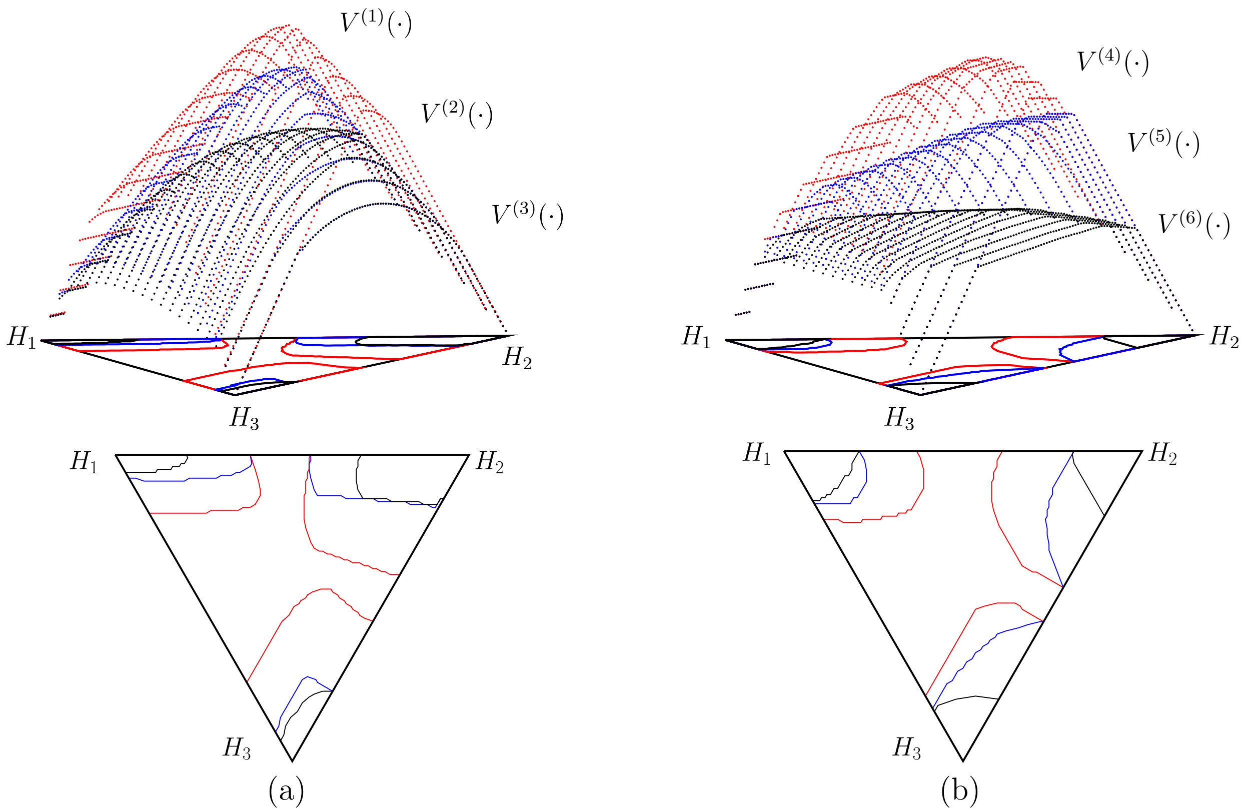


Figure 4: Impact on the minimum Bayes risk in (1.6) and on its stopping regions of different alternative hypotheses in (1.2) about unknown characteristics of a compound Poisson process. On the left, $V^{(1)}$, $V^{(2)}$, $V^{(3)}$ are the minimum Bayes risks when (i) alternative arrival rates $(\lambda_1, \lambda_2, \lambda_3)$ are $(1,3,5)$, $(1,3,10)$, $(0.1,3,10)$, respectively, and (ii) the alternative mark distributions on $\{1,2,3,4,5\}$ are given by (5.1). On the right, $V^{(4)}$, $V^{(5)}$, $V^{(6)}$ are the minimum Bayes risks when (i) alternative arrival rates $(\lambda_1, \lambda_2, \lambda_3) \equiv (3,3,3)$ are the same, and (ii) the marks are Gamma distributed random variables with common scale parameter 2 and unknown shape parameter alternatives $(3,6,9)$, $(1,6,9)$, $(1,6,15)$, respectively.

the 3-simplex E. Note that the first three hypotheses are the same as those of the previous example. Therefore, this problem reduces to the previous one if $\pi_4 = 0$, and the intersection of its stopping regions with the $H_1H_2H_3$-face of the simplex E coincides with those stopping regions of the previous problem.

Next we suppose that the marks of a compound Poisson process X take values in $\{1,2,3,4,5\}$ and their unknown common distribution $\nu(\cdot)$ is one of the alternatives

$$\nu_1 = \left\{\frac{1}{15}, \frac{2}{15}, \frac{3}{15}, \frac{4}{15}, \frac{5}{15}\right\}, \; \nu_2 = \left\{\frac{3}{15}, \frac{3}{15}, \frac{3}{15}, \frac{3}{15}, \frac{3}{15}\right\}, \; \nu_3 = \left\{\frac{5}{15}, \frac{4}{15}, \frac{3}{15}, \frac{2}{15}, \frac{1}{15}\right\}. \tag{5.1}$$

For three different sets of alternative arrival rates $(\lambda_1, \lambda_2, \lambda_3)$ in (1.2); namely, $(1, 3, 5)$, $(1, 3, 10)$, $(0.1, 3, 10)$, we solve the problem (1.6) and calculate the corresponding minimum Bayes risk functions $V^{(1)}(\cdot), V^{(2)}(\cdot), V^{(3)}(\cdot)$, respectively. The terminal decision cost is the same $a_{ij} = 2$ for every $1 \le i \ne j \le 3$. In Figure 4(a), the functions $V^{(1)}, V^{(2)}, V^{(3)}$ and stopping region boundaries are plotted. It suggests that $V^{(1)}(\cdot) \ge V^{(2)}(\cdot) \ge V^{(3)}(\cdot)$ as expected, since the wider the alternative arrival rates are stretched from each other, the easier is to discriminate the correct rate.

Figure 4(b) displays the minimum Bayes risks $V^{(4)}$, $V^{(5)}$, $V^{(6)}$ and boundaries of stopping regions of three more sequential hypothesis-testing problems with $M = 3$ alternative hypotheses. For each of those three problems, the alternative arrival rates $(\lambda_1, \lambda_2, \lambda_3)$ are the same and equal $(3, 3, 3)$, and the common mark distribution $\nu(\cdot)$ is Gamma$(2, \beta)$ with scale parameter 2 and unknown shape parameter β. Alternative β-values under three hypotheses are $(3, 6, 9)$, $(1, 6, 9)$, $(1, 6, 15)$ for three problems $V^{(4)}$, $V^{(5)}$, $V^{(6)}$, respectively. Figure 4(b) suggests that $V^{(4)}(\cdot) \ge V^{(5)}(\cdot) \ge V^{(6)}(\cdot)$ as expected, since alternative Gamma distributions are easier to distinguish as the shape parameters stretched out from each other.

Furthermore, the kinks of the functions $V^{(4)}(\cdot)$, $V^{(5)}(\cdot)$, $V^{(6)}(\cdot)$ suggest that these are not smooth; compare with $V^{(1)}(\cdot)$, $V^{(2)}(\cdot)$, $V^{(3)}(\cdot)$ in Figure 4 (a). Dayanik and Sezer (2005) show for $M = 2$ that the function $V(\cdot)$ in (1.6) may not be differentiable, even in the interior of the continuation, if the alternative arrival rates are equal. This result applies to the above example with $M = 3$, at least on the regions where one of the components of $\vec{\pi}$ equals zero and the example reduces to a two-hypothesis testing problem.

6. Alternative Bayes risks

The key facts behind the analysis of the problem (1.6) were that (i) the Bayes risk R in (1.3) is the expectation of a special functional of the piecewise-deterministic Markov process Π in (2.2), (2.7)-(2.10), (ii) the stopping times of such processes are "predictable between jump times" in the sense of Lemma 3.3, and (iii) the minimum Bayes risk $V(\cdot)$ in (1.6) and (2.3) is approximated *uniformly* by the truncations $V_n(\cdot)$, $n \ge 1$ in (3.1) of the same problem at arrival times of observation process X. Fortunately, all of these facts remain valid for the other two problems in (1.6),

$$V'(\vec{\pi}) \triangleq \inf_{(\tau,d)\in\mathcal{A}} R'(\vec{\pi}, \tau, d) \qquad \text{and} \qquad V''(\vec{\pi}) \triangleq \inf_{(\tau,d)\in\mathcal{A}} R''(\vec{\pi}, \tau, d), \tag{6.1}$$

of the alternative Bayes risks R' and R'' in (1.4) and (1.5), respectively. Minor changes to the proof of Proposition 2.1 give that

$$(6.2)\qquad \begin{aligned} V'(\vec\pi) &= \inf_{\tau\in\mathbb{F}} \mathbb{E}^{\vec\pi}\left[N_\tau + \mathbf{1}_{\{\tau<\infty\}} h(\Pi(\tau))\right], \quad \vec\pi\in E,\\ V''(\vec\pi) &= \inf_{\tau\in\mathbb{F}} \mathbb{E}^{\vec\pi}\left[\sum_{k=1}^{N_\tau} e^{-\rho\sigma_k} f(Y_k) + e^{-\rho\tau} h\left(\Pi(\tau)\right)\right], \quad \vec\pi\in E \end{aligned}$$

in terms of the minimum terminal decision cost function $h(\cdot)$ in (2.4). Moreover, for every a.s.-finite optimal $\mathbb{F}$-stopping times τ' and τ'' of the problems in (6.2) the admissible strategies $(\tau', d(\tau'))$ and $(\tau'', d(\tau''))$ attain the minimum Bayes risks $V'(\cdot)$ and $V''(\cdot)$ in (6.1), respectively, where the random variable $d(t)\in \operatorname{argmin}_{i\in I}\sum_{j\in I} a_{ij}\Pi_i(t)$, $t\ge 0$ is the same as in Proposition 2.1. The sequence of functions

$$(6.3)\qquad \begin{aligned} V'_n(\vec\pi) &\triangleq \inf_{\tau\in\mathbb{F}} \mathbb{E}^{\vec\pi}\left[n\wedge N_\tau + \mathbf{1}_{\{\tau<\infty\}} h(\Pi(\tau\wedge\sigma_n))\right], \; n\ge 0 \quad \text{and}\\ V''_n(\vec\pi) &\triangleq \inf_{\tau\in\mathbb{F}} \mathbb{E}^{\vec\pi}\left[\sum_{k=1}^{n\wedge N_\tau} e^{-\rho\sigma_k} f(Y_k) + e^{-\rho(\tau\wedge\sigma_n)} h\left(\Pi(\tau\wedge\sigma_n)\right)\right], \; n\ge 0, \end{aligned}$$

obtained from the problems in (6.2) by truncation at the nth arrival time σ_n ($\sigma_0\equiv 0$) of the observation process X, approximate respectively the functions $V'(\vec\pi)$ and $V''(\vec\pi)$ uniformly in $\vec\pi\in E$, and the error bounds are similar to (3.2); i.e.,

$$0\le ||V'_n - V'||\,, ||V''_n - V''|| \le \left(\sum_{i,j} a_{ij}\right)^{3/2}\sqrt{\frac{\lambda_M}{n-1}} \qquad \text{for every } n\ge 1.$$

Moreover, the sequences $(V'_n(\cdot))_{n\ge1}$ and $(V''_n(\cdot))_{n\ge1}$ can be calculated successively as in

$$V'_{n+1}(\vec\pi) = \inf_{t\in[0,\infty]} J'V'_n(t,\vec\pi),\; n\ge 0 \qquad \text{and} \qquad V''_{n+1}(\vec\pi) = \inf_{t\in[0,\infty]} J''V''_n(t,\vec\pi),\; n\ge 0$$

by using the operators J' and J'' acting on bounded functions $w: E\mapsto\mathbb{R}$ according to

$$(6.4)\qquad \begin{aligned} J'w(t,\vec\pi) &\triangleq \int_0^t \sum_{i=1}^M \pi_i\, e^{-\lambda_i u}\lambda_i\left[1 + (S_i w)(x(u,\vec\pi))\right] du + \sum_{i=1}^M \pi_i\, e^{-\lambda_i t} h(x(t,\vec\pi)),\\ J''w(t,\vec\pi) &\triangleq \int_0^t \sum_{i=1}^M \pi_i\, e^{-(\rho+\lambda_i)u}\lambda_i\left[\mu_i + (S_i w)(x(u,\vec\pi))\right] du + \sum_{i=1}^M \pi_i\, e^{-(\rho+\lambda_i)t} h(x(t,\vec\pi)) \end{aligned}$$

for every $t\ge 0$; here, S_i, $i\in I$ is the same operator in (3.6), and

$$\mu_i \triangleq \mathbb{E}^\bullet[f(Y_1)\mid \Theta = i] = \int_{\mathbb{R}^d} f(y)\nu_i(dy) \equiv \int_{\mathbb{R}^d} f(y)p_i(y)\mu(dy), \qquad i\in I$$

is the conditional expectation of the random variable $f(Y_1)$ given that the correct hypothesis is $\Theta = i$. Recall from Section 1 that the negative part $f^-(\cdot)$ of the Borel function $f : \mathbb{R}^d \mapsto \mathbb{R}$ is ν_i-integrable for every $i \in I$; therefore, the expectation μ_i, $i \in I$ exists.

As in Proposition 3.16 and Corollary 3.17 the $\mathbb{F}$-stopping times

$$U_0' \triangleq \inf\Big\{t \geq 0;\, V'\big(\Pi(t)\big) = h\big(\Pi(t)\big)\Big\} \quad \text{and} \quad U_0'' \triangleq \inf\Big\{t \geq 0;\, V''\big(\Pi(t)\big) = h\big(\Pi(t)\big)\Big\}$$

are optimal for the problems $V'(\cdot)$ and $V''(\cdot)$ in (6.2), respectively. The functions $V'(\cdot)$ and $V''(\cdot)$ are concave and continuous, and the *nonempty* stopping regions

$$\Gamma_\infty' \triangleq \{\vec{\pi} \in E \,:\, V'(\vec{\pi}) = h(\vec{\pi})\} \qquad \text{and} \qquad \Gamma_\infty'' \triangleq \{\vec{\pi} \in E \,:\, V''(\vec{\pi}) = h(\vec{\pi})\}$$

are the union of closed and convex neighborhoods

$$\Gamma_{\infty,i}' \triangleq \Gamma_\infty' \cap \{\vec{\pi} \in E : h(\vec{\pi}) = h_i(\vec{\pi})\} \quad \text{and} \quad \Gamma_{\infty,i}'' \triangleq \Gamma_\infty'' \cap \{\vec{\pi} \in E : h(\vec{\pi}) = h_i(\vec{\pi})\},\ i \in I$$

of M corners of the $(M-1)$-simplex E. It is optimal to stop at time U_0' (respectively, U_0'') and select the hypothesis H_i, $i \in I$ if $\Pi(U_0') \in \Gamma_{\infty,i}'$ (respectively, $\Pi(U_0'') \in \Gamma_{\infty,i}''$).

The numerical algorithm in Figure 2 approximates $V'(\cdot)$ in (6.2) with its outcome $V_N'(\cdot)$ or $V_{\delta,N}'(\cdot)$ if J, $\overline{\pi}_i$, π_i^* are replaced with J' in (6.4),

$$\overline{\pi}_i' \triangleq \left\{1 - \frac{\lambda_1}{2\,\lambda_M\,(\max_k a_{ik})}\right\} \vee \frac{\max_k a_{ik}}{(\min_{j\neq i} a_{ji}) + (\max_k a_{ik})},$$

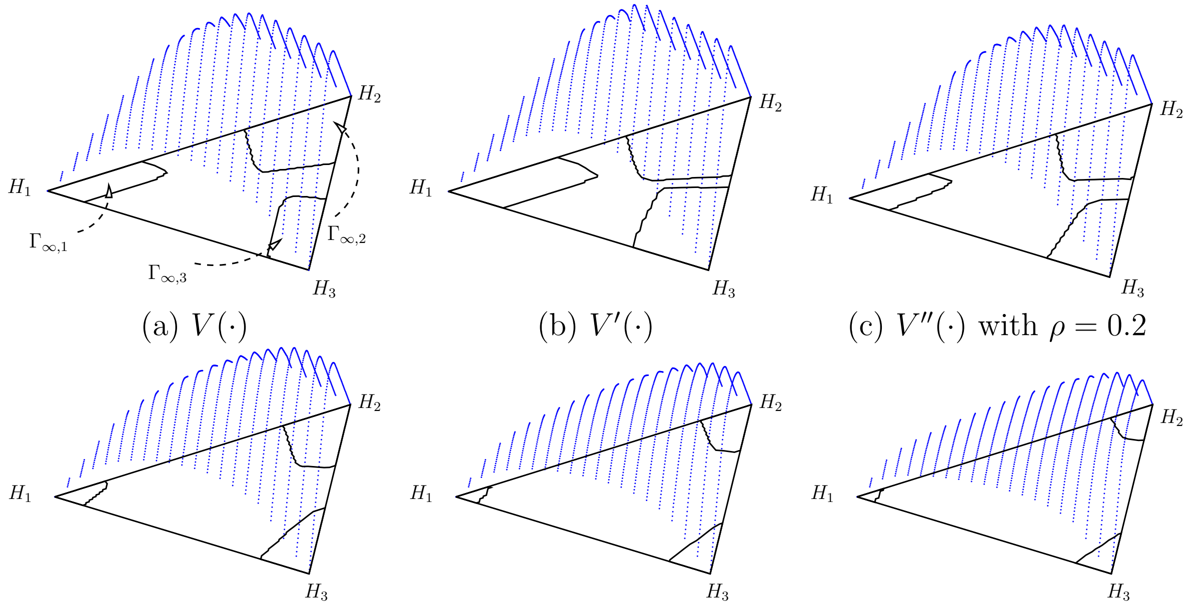
$$\pi_i' \triangleq \inf\left\{\pi_i \in [\overline{\pi}_i', 1] \,:\, \frac{\pi_i}{(1-\pi_i)}\left[1 - \left(\frac{1-\overline{\pi}_i}{\overline{\pi}_i}\cdot\frac{\pi_i}{1-\pi_i}\right)^{-\lambda_i/(\lambda_i-\lambda_1)}\right] \geq \max_k a_{ik}\right\},$$

respectively, for every $i \in I$. Similarly, its outcome $V_N''(\cdot)$ or $V_{\delta,N}''(\cdot)$ approximates the function $V''(\cdot)$ in (6.2) if in Figure 2 J, $\overline{\pi}_i$, π_i^* are replaced with J'' in (6.4),

$$\overline{\pi}_i'' \triangleq \left\{1 - \frac{\lambda_1}{2\,(\lambda_M+\rho)\,(\max_k a_{ik})}\right\} \vee \frac{\max_k a_{ik}}{(\min_{j\neq i} a_{ji}) + (\max_k a_{ik})},$$

$$\pi_i'' \triangleq \inf\left\{\pi_i \in [\overline{\pi}_i'', 1] \,:\, \frac{\pi_i\lambda_i}{(\lambda_i+\rho)(1-\pi_i)}\left[1 - \left(\frac{1-\overline{\pi}_i}{\overline{\pi}_i}\cdot\frac{\pi_i}{1-\pi_i}\right)^{-(\lambda_i+\rho)/(\lambda_i-\lambda_1)}\right] \geq \max_k a_{ik}\right\},$$

respectively, for every $i \in I$. As in Section 4.1, these numerical solutions can be used to obtain nearly-optimal strategies. For every $\varepsilon > 0$, if we fix N such that $||V_N' - V'|| \leq \varepsilon/2$ and define $U_{\varepsilon/2,N}' \triangleq \inf\{t \geq 0;\ h(\Pi(t)) \leq V_N'(\Pi(t)) + \varepsilon/2\}$, then $\big(U_{\varepsilon/2,N}', d(U_{\varepsilon/2,N}')\big)$ is an ε-optimal admissible strategy for the minimum Bayes risk $V'(\cdot)$ in (6.1). Similarly, if we fix N such that $||V_N'' - V''|| \leq \varepsilon/2$ and define $U_{\varepsilon/2,N}'' \triangleq \inf\{t \geq 0;\ h(\Pi(t)) \leq V_N''(\Pi(t)) + \varepsilon/2\}$, then $\big(U_{\varepsilon/2,N}'', d(U_{\varepsilon/2,N}'')\big)$ is an ε-optimal admissible strategy for the minimum Bayes risk $V''(\cdot)$ in



(a) $V(\cdot)$ (b) $V'(\cdot)$ (c) $V''(\cdot)$ with $\rho = 0.2$

(d) $V''(\cdot)$ with $\rho = 0.4$ (e) $V''(\cdot)$ with $\rho = 0.6$ (f) $V''(\cdot)$ with $\rho = 0.8$

Figure 5: The minimum Bayes risk functions $V(\cdot)$, $V'(\cdot)$, and $V''(\cdot)$ corresponding to Bayes risks R in (1.3), R' in (1.4), and R'' in (1.5) with discount rates $\rho = 0.2, 0.4, 0.6, 0.8$, respectively, when three alternative hypotheses, $H_1 : \lambda = 1$, $H_2 : \lambda = 2$, and $H_3 : \lambda = 3$, about the arrival rate λ of a simple Poisson process are tested. In each figure, the stopping regions $\Gamma_{\infty,1}$, $\Gamma_{\infty,2}$, and $\Gamma_{\infty,3}$ in the neighborhoods of the corners at H_1, H_2, and H_3 are also displayed. The terminal decision costs are $a_{ij} = 10$ for every $1 \leq i \neq j \leq 3$.

(6.1). The proofs are omitted since they are similar to those discussed in detail for $V(\cdot)$ of (1.6). However, beware that the Bayes risk R' in (1.4) does not immediately reduce to R of (1.3), since the $(\mathbb{P}^{\vec{\pi}}, \mathbb{F})$-compensator of the process N_t, $t \geq 0$ is $\sum_{i=1}^{M} \Pi_i(t)\lambda_i$, $t \geq 0$.

Figure 5 compares the minimum Bayes risk functions and optimal stopping regions for a three-hypothesis testing problem. Comparison of (a) and (b) confirms that the Bayes risks R and R' are different. Since the stopping regions in (b) are larger than those in (a), optimal decision strategies arrive at a conclusion earlier under R' than under R. As the discount rate ρ decreases to zero, the minimum risk $V''(\cdot)$ and stopping regions become larger and converge, apparently to $V'(\cdot)$ and its stopping regions, respectively.

Acknowledgment

The research of Savas Dayanik was supported by the Air Force Office of Scientific Research, under grant AFOSR-FA9550-06-1-0496. The research of H. Vincent Poor and Semih O. Sezer was supported by the U.S. Army Pantheon Project.

Appendix: proofs of selected results

Proof of (2.5). The right hand side of (2.5) is $\mathcal{F}_t$-measurable, and it is sufficient to show

$$\mathbb{E}^{\vec{\pi}}\left[\mathbf{1}_A\cdot\mathbb{P}^{\vec{\pi}}(\Theta=i\,|\,\mathcal{F}_t)\right]=\mathbb{E}^{\vec{\pi}}\left[\mathbf{1}_A\cdot\frac{\pi_i e^{-\lambda_i t}\prod_{k=1}^{N_t}\lambda_i\cdot p_i(Y_k)}{\sum_{j=1}^M\pi_j e^{-\lambda_j t}\prod_{k=1}^{N_t}\lambda_j\cdot p_j(Y_k)}\right]$$

for sets of the form $A\triangleq\{N_{t_1}=m_1,\ldots,N_{t_k}=m_k,(Y_1,\ldots,Y_{m_k})\in B_{m_k}\}$ for every $k\geq 1$, $0\equiv t_0\leq t_1\leq\ldots\leq t_k=t$, $0\equiv m_0\leq m_1\leq\ldots\leq m_k$, and Borel subset B_{m_k} of $\mathbb{R}^{m_k\times d}$. If A is as above, then $\mathbb{E}^{\vec{\pi}}\left[\mathbf{1}_A\,\mathbb{P}^{\vec{\pi}}(\Theta=i\,|\,\mathcal{F}_t)\right]=\mathbb{E}^{\vec{\pi}}\left[\mathbb{E}^{\vec{\pi}}(\,\mathbf{1}_{A\cap\{\Theta=i\}}\,|\,\mathcal{F}_t)\,\right]=\mathbb{E}^{\vec{\pi}}\left[\mathbf{1}_{A\cap\{\Theta=i\}}\right]=\pi_i\,\mathbb{P}_i(A)$ becomes

$$\pi_i\,e^{-\lambda_i t_k}\prod_{n=1}^k\frac{(t_n-t_{n-1})^{m_n-m_{n-1}}}{(m_n-m_{n-1})!}\int_{B_{m_k}}\prod_{\ell=1}^{m_k}p_i(y_\ell)\mu(dy_\ell)$$
$$=\prod_{n=1}^k\frac{(t_n-t_{n-1})^{m_n-m_{n-1}}}{(m_n-m_{n-1})!}\int_{B_{m_k}}L_i(t_k,m_k,y_1,\ldots,y_{m_k})\prod_{\ell=1}^{m_k}\mu(dy_\ell)$$

in terms of $L_i(t,m,y_1,\ldots,y_m)\triangleq\pi_i e^{-\lambda_i t}\prod_{\ell=1}^m\lambda_i p_i(y_\ell)$, $i\in I$. If $L(t,m,y_1,\ldots,y_m)\triangleq\sum_{i=1}^M L_i(t,m,y_1,\ldots,y_m)$, then $\mathbb{E}^{\vec{\pi}}\left[\mathbf{1}_A\,\mathbb{P}^{\vec{\pi}}(\Theta=i\mid\mathcal{F}_t)\right]$ equals

$$\prod_{n=1}^k\frac{(t_n-t_{n-1})^{m_n-m_{n-1}}}{(m_n-m_{n-1})!}\int_{B_{m_k}}L(t_k,m_k,y_1,\ldots,y_{m_k})\cdot\frac{L_i(t_k,m_k,y_1,\ldots,y_{m_k})}{L(t_k,m_k,y_1,\ldots,y_{m_k})}\prod_{\ell=1}^{m_k}\mu(dy_\ell)$$
$$=\sum_{j=1}^M\pi_j e^{-\lambda_j t_k}\lambda_j^{m_k}\prod_{n=1}^k\frac{(t_n-t_{n-1})^{m_n-m_{n-1}}}{(m_n-m_{n-1})!}\int_{B_{m_k}}\frac{L_i(t_k,m_k,y_1,\ldots,y_{m_k})}{L(t_k,m_k,y_1,\ldots,y_{m_k})}\prod_{\ell=1}^{m_k}p_j(y_\ell)\mu(dy_\ell)$$
$$=\sum_{j=1}^M\pi_j\mathbb{E}^{\vec{\pi}}\left[\mathbf{1}_A\frac{L_i(t_k,N_{t_k},Y_1,\ldots,Y_{N_{t_k}})}{L(t_k,N_{t_k},Y_1,\ldots,Y_{N_{t_k}})}\middle|\,\Theta=j\right]=\mathbb{E}^{\vec{\pi}}\left[\mathbf{1}_A\frac{\pi_i e^{-\lambda_i t_k}\prod_{\ell=1}^{N_{t_k}}\lambda_i p_i(Y_\ell)}{\sum_{j=1}^M\pi_j e^{-\lambda_j t_k}\prod_{\ell=1}^{N_{t_k}}\lambda_j p_j(Y_\ell)}\right].\quad\square$$

Proof of (2.9). For every bounded $\mathcal{F}_t$-measurable r.v. Z, we have

$$\mathbb{E}^{\vec{\pi}}[Z\,g(\Pi(t+s))]=\sum_{i=1}^M\pi_i\,\mathbb{E}_i[Z\,g(\Pi(t+s))]=\sum_{i=1}^M\pi_i\,\mathbb{E}_i[Z\,\mathbb{E}_i[g(\Pi(t+s))|\mathcal{F}_t]]$$
$$=\sum_{i=1}^M\mathbb{P}^{\vec{\pi}}\{\Theta=i\}\,\mathbb{E}^{\vec{\pi}}[Z\,\mathbb{E}_i[g(\Pi(t+s))|\mathcal{F}_t]\mid\Theta=i]=\sum_{i=1}^M\mathbb{E}^{\vec{\pi}}\left[Z\,\mathbf{1}_{\{\Theta=i\}}\mathbb{E}_i[g(\Pi(t+s))|\mathcal{F}_t]\right]$$
$$=\sum_{i=1}^M\mathbb{E}^{\vec{\pi}}\left[Z\,\Pi_i(t)\,\mathbb{E}_i[g(\Pi(t+s))|\mathcal{F}_t]\right]=\mathbb{E}^{\vec{\pi}}\left[Z\left(\sum_{i=1}^M\Pi_i(t)\,\mathbb{E}_i[g(\Pi(t+s))|\mathcal{F}_t]\right)\right].\quad\square$$

Proof of Proposition 2.1. Let τ be an $\mathbb{F}$- stopping time taking countably many real values $(t_n)_{n\geq 1}$. Then by monotone convergence theorem

$$\text{(A.1)}\quad \mathbb{E}^{\vec{\pi}}\left[\tau+\sum_{i=1}^{M}\sum_{j=1}^{M}a_{ij}\cdot\mathbf{1}_{\{d=i,\,\Theta=j\}}\right]=\sum_n \mathbb{E}^{\vec{\pi}}\mathbf{1}_{\{\tau=t_n\}}\left[t_n+\sum_{i=1}^{M}\sum_{j=1}^{M}a_{ij}\cdot\mathbf{1}_{\{d=i,\,\Theta=j\}}\right]$$
$$=\sum_n \mathbb{E}^{\vec{\pi}}\mathbf{1}_{\{\tau=t_n\}}\left[t_n+\sum_{i=1}^{M}\sum_{j=1}^{M}a_{ij}\cdot\mathbf{1}_{\{d=i\}}\mathbb{P}^{\vec{\pi}}\{\Theta=j\,|\mathcal{F}_{t_n}\}\right]$$
$$=\sum_n \mathbb{E}^{\vec{\pi}}\mathbf{1}_{\{\tau=t_n\}}\left[t_n+\sum_{i=1}^{M}\sum_{j=1}^{M}a_{ij}\cdot\mathbf{1}_{\{d=i\}}\Pi_j(t_n)\right]=\mathbb{E}^{\vec{\pi}}\left[\tau+\sum_{i=1}^{M}\sum_{j=1}^{M}a_{ij}\cdot\mathbf{1}_{\{d=i,\}}\Pi_j(\tau)\right].$$

Next take an arbitrary $\mathbb{F}$-stopping time τ with finite expectation $\mathbb{E}^{\vec{\pi}}\tau$. There is a decreasing sequence $(\tau_k)_{k\geq 1}$ of $\mathbb{F}$-stopping times converging to τ such that (i) every τ_k, $k\geq 1$ takes countably many values, and (ii) $\mathbb{E}^{\vec{\pi}}\tau_1$ is finite. (For example, if $\varphi_k(t)\triangleq(j-1)k2^{-k}$ whenever $t\in[(j-1)k2^{-k},jk2^{-k})$ for some $j\geq 1$, then $\tau_k\triangleq\varphi_k(\tau)$, $k\geq 1$ are $\mathbb{F}$-stopping times taking countably many real values and $\tau_k\leq\tau<\tau_k+k2^k$ for every $k\geq 1$.) Therefore, $d\in\mathcal{F}_\tau\subset\mathcal{F}_{\tau_k}$ for every $k\geq 1$. Then the right-continuity of the sample-paths of the process Π_i in (2.2), dominated convergence theorem, and (A.1) lead to

$$\mathbb{E}^{\vec{\pi}}\left[\tau+\sum_{i=1}^{M}\sum_{j=1}^{M}a_{ij}\cdot\mathbf{1}_{\{d=i,\,\Theta=j\}}\right]=\lim_{k\to\infty}\mathbb{E}^{\vec{\pi}}\left[\tau_k+\sum_{i=1}^{M}\sum_{j=1}^{M}a_{ij}\cdot\mathbf{1}_{\{d=i,\,\Theta=j\}}\right]$$
$$=\lim_{k\to\infty}\mathbb{E}^{\vec{\pi}}\left[\tau_k+\sum_{i=1}^{M}\sum_{j=1}^{M}a_{ij}\cdot\mathbf{1}_{\{d=i\}}\Pi_j(\tau_k)\right]=\mathbb{E}^{\vec{\pi}}\left[\tau+\sum_{i=1}^{M}\sum_{j=1}^{M}a_{ij}\cdot\mathbf{1}_{\{d=i\}}\Pi_j(\tau)\right].$$

The last expectation is minimized if we take $d(\tau)\in\arg\min_{i\in I}\sum_{j=1}^{M}a_{ij}\cdot\Pi_j(\tau)$. Hence,

$$\inf_{(\tau,d)\in\mathcal{A}}R(\tau,d)\geq\inf_{\tau}\ \mathbb{E}^{\vec{\pi}\in\mathbb{F}}\left[\tau+\sum_{i=1}^{M}\sum_{j=1}^{M}a_{ij}\cdot\mathbf{1}_{\{d=i\}}\Pi_j(\tau)\right]$$

by Remark 3.1. Since $d(\tau)$ is $\mathcal{F}_\tau$-measurable, we also have the reverse inequality. □

The following result will be useful in establishing Lemma 2.3.

Lemma A.1. *Suppose $\lambda_1<\lambda_2<\ldots<\lambda_M$. Let $\widehat{E}_{(q_1,\ldots,q_M)}$ be as in (2.11) for some $0<q_1,\ldots,q_M<1$. For every fixed $k\in\{1,\ldots,M-1\}$ and arbitrary constant $0<\delta_k<1$, there exists some $\varepsilon_k\in(0,\delta_k]$ such that starting at time $t=0$ in the region*

$$\{\vec{\pi}\in E\setminus\widehat{E}_{(q_1,\ldots,q_M)}\,:\,\pi_j<\varepsilon_k\ \ \forall j<k\ \text{and}\ \pi_k\geq\delta_k\},$$

the entrance time of the path $t \mapsto x(t,\vec{\pi})$, $t \geq 0$ into $\widehat{E}_{(q_1,\ldots,q_M)}$ is bounded from above. More precisely, there exists some finite $t_k(\delta_k, q_k)$ such that

$$x_k\left(t_k(\delta_k,q_k),\vec{\pi}\right) \geq q_k \quad \textit{for every } \vec{\pi} \textit{ in the region above.}$$

Proof. Let $k=1$ and $\delta_1 \in (0,1)$ be fixed. The mapping $t \mapsto x_1(t,\vec{\pi})$ is increasing, and $\lim_{t\to\infty} x_1(t,\vec{\pi}) = 1$ on $\{\vec{\pi} \in E : \pi_1 \geq \delta_1\}$. Hence the given level $0 < q_1 < 1$ will eventually be exceeded. Then the explicit form of $x(t,\vec{\pi})$ in (2.8) implies that we can set

$$t_1(\delta_1, q_1) = \frac{1}{\lambda_2 - \lambda_1} \ln\left(\frac{1-\delta_1}{\delta_1} \cdot \frac{q_1}{1-q_1}\right).$$

For $1 < k < M-1$, let $\delta_k \in (0,1)$ be a given constant. By (2.10), the mapping $t \mapsto x_k(t,\vec{\pi})$ is increasing on the region $\{\vec{\pi} \in E : \pi_j \leq \hat{\varepsilon}_k, \text{ for all } j < k\}$, where

$$\hat{\varepsilon}_k \triangleq \min\left\{\frac{(1-q_k)(\lambda_{k+1}-\lambda_k)}{(k-1)(\lambda_{k+1}-\lambda_{k-1})}, \frac{1-\delta_k}{k-1}\right\}.$$

Let us fix

$$\varepsilon_k \triangleq \sup\Bigg\{0 \leq \pi_k \leq \delta_k \wedge \hat{\varepsilon}_k \wedge \frac{\lambda_M - \lambda_k}{\lambda_M - \lambda_1} : \delta_k\left(\frac{1-q_k}{q_k}\right) \geq$$
$$\pi_k(k-1)\left(\frac{\hat{\varepsilon}_k}{1-\hat{\varepsilon}_k} \cdot \frac{1-\pi_k}{\pi_k}\right)^{(\lambda_k-\lambda_1)/(\lambda_M-\lambda_1)} + \left(\frac{\hat{\varepsilon}_k}{1-\hat{\varepsilon}_k} \cdot \frac{1-\pi_k}{\pi_k}\right)^{(\lambda_k-\lambda_{k+1})(\lambda_M-\lambda_1)}\Bigg\}.$$

The right hand side of the inequality above is 0 at $\pi_k = 0$, and it is increasing on $\pi_k \in [0, (\lambda_M - \lambda_k)/(\lambda_M - \lambda_1)]$. Therefore, ε_k is well defined. Moreover, since $\varepsilon_k \leq \hat{\varepsilon}_k$, for a point $\vec{\pi}$ on the subset $\{\vec{\pi} \in E : \pi_j < \varepsilon_k \text{ for all } j < k\}$ we have by (2.8) for every $j < k$ that

$$x_j\left(t,\vec{\pi}\right) \leq \hat{\varepsilon}_k \qquad \text{for every} \quad t \leq \frac{1}{\lambda_M - \lambda_1} \ln\left(\frac{\hat{\varepsilon}_k}{1-\hat{\varepsilon}_k} \cdot \frac{1-\varepsilon_k}{\varepsilon_k}\right).$$

In other words, $t \mapsto x_k(t,\vec{\pi})$ is increasing on $t \in \left[0, \frac{1}{\lambda_M-\lambda_1} \ln\left(\frac{\hat{\varepsilon}_k}{1-\hat{\varepsilon}_k} \cdot \frac{1-\varepsilon_k}{\varepsilon_k}\right)\right]$ for every $\vec{\pi}$ in the smaller set $\{\vec{\pi} \in E : \pi_j < \varepsilon_k \text{ for all } j < k\}$. Moreover,

$$\delta_k\left(\frac{1-q_k}{q_k}\right) \geq \sum_{j<k} \varepsilon_k \left(\frac{\hat{\varepsilon}_k}{1-\hat{\varepsilon}_k} \cdot \frac{1-\varepsilon_k}{\varepsilon_k}\right)^{\frac{\lambda_k-\lambda_j}{\lambda_M-\lambda_1}} + \sum_{j>k} \pi_j \left(\frac{\hat{\varepsilon}_k}{1-\hat{\varepsilon}_k} \cdot \frac{1-\varepsilon_k}{\varepsilon_k}\right)^{\frac{\lambda_k-\lambda_j}{\lambda_M-\lambda_1}}.$$

For every $\vec{\pi} \in E$ such that $\pi_k \geq \delta_k$ and $\pi_j < \varepsilon_k$ for all $j < k$, rearranging this inequality and using (2.8) give

$$x_k\left(\frac{1}{\lambda_M - \lambda_1} \ln\left(\frac{\hat{\varepsilon}_k}{1-\hat{\varepsilon}_k} \cdot \frac{1-\varepsilon_k}{\varepsilon_k}\right), \vec{\pi}\right) \geq q_k,$$

which completes the proof for $1 < k < M-1$. □

Proof of Lemma 2.3. The claim is immediate if $\vec{\pi} \in \widehat{E}_{(q_1,\ldots,q_M)}$. To prove it on $E \setminus \widehat{E}_{(q_1,\ldots,q_M)}$, with the notation in Lemma A.1, fix

$$\delta_{M-1} = \frac{1-q_M}{M-1}, \qquad \text{and} \qquad \delta_i = \varepsilon_{i+1} \;\; \text{for} \;\; i = M-2,\ldots,1.$$

Then by Lemma A.1 we have $\delta_1 \le \ldots \le \delta_{M-1}$, and our choice of δ_{M-1} implies that $\{\vec{\pi} \in E \setminus E^* \,:\, \pi_i < \delta_i,\; i < M-1\} = \varnothing$. By the same lemma, for every starting point $\vec{\pi}$ at time $t=0$ in the set

$$\begin{aligned}\{\vec{\pi} \in E \setminus \widehat{E}_{(q_1,\ldots,q_M)} \,:\, \pi_i < \delta_i,\; \forall i < M-2,\; \text{and}\;\; \pi_{M-1} \ge \delta_{M-1}\} \\ \subseteq \{\vec{\pi} \in E \setminus \widehat{E}_{(q_1,\ldots,q_M)} \,:\, \pi_i < \delta_{M-2},\; \forall i < M-2,\; \text{and}\;\; \pi_{M-1} \ge \delta_{M-1}\}\end{aligned}$$

there exists $t_{M-1}(\delta_{M-1}, q_{M-1})$ such that the exit time of the path $t \mapsto x(t,\vec{\pi})$ from $E \setminus \widehat{E}_{(q_1,\ldots,q_M)}$ is bounded from above by $t_{M-1}(\delta_{M-1}, q_{M-1})$. Then the last two statements imply that the same bound holds on $\{\vec{\pi} \in E \setminus \widehat{E}_{(q_1,\ldots,q_M)} \,:\, \pi_i < \delta_i,\; \text{for}\; i < M-2\}$.

Now assume that for some $1 \le n < M-1$ and for every $\vec{\pi}$ in the set

$$\{\vec{\pi} \in E \setminus \widehat{E}_{(q_1,\ldots,q_M)} \,:\, \pi_i < \delta_i,\; \text{for}\; i < n\},$$

we have $\inf\{t \ge 0 \,:\, x(t,\vec{\pi}) \in \widehat{E}_{(q_1,\ldots,q_M)}\} \le \max_{k>n} t_k(\delta_k, q_k)$. Again by Lemma A.1 there exists $t_n(\delta_n, q_n) < \infty$ as the upper bound on the hitting time to $\widehat{E}_{(q_1,\ldots,q_M)}$ for all the points in $\{\vec{\pi} \in E \setminus \widehat{E}_{(q_1,\ldots,q_M)} \,:\, \pi_i < \delta_i,\; \text{for}\; i < n-1,\; \text{and}\;\; \pi_n \ge \delta_n\}$. Hence on $\{\vec{\pi} \in E \setminus \widehat{E}_{(q_1,\ldots,q_M)} \,:\, \pi_i < \delta_i,\; \text{for}\; i < n-1\}$ the new upper bound is $\max_{k>n-1} t_k(\delta_k, q_k)$. By induction, it follows that $s\,(q_1,\ldots,q_M)$ can be taken as $\max_{k\ge 1} t_k(\delta_k, q_k)$. □

Proof of Proposition 3.4. Monotonicity of the operators Jw and $J_0 w$, and boundedness of $J_0 w(\cdot)$ from above by $h(\cdot)$ follow from (3.4) and (3.5). The check the concavity, let $w(\cdot)$ be a concave function on E. Then $w(\vec{\pi}) = \inf_{k\in K}\; \beta_0^k + \sum_{j=1}^M \beta_j^k \pi_j$ for some index set K and some constants β_j^k, $k \in K$, $j \in I \cup \{0\}$. Then $\sum_{i=1}^M \pi_i e^{-\lambda_i u}\lambda_i \cdot (S_i w)(x(u,\vec{\pi}))$ equals

$$\begin{aligned}\sum_{i=1}^M \lambda_i \pi_i e^{-\lambda_i u} \int_{\mathbb{R}^d} \inf_{k\in K} \left[\beta_0^k + \frac{\sum_{j=1}^M \beta_j^k\, \lambda_j\, x_j(u,\vec{\pi})\, p_j(y)}{\sum_{j=1}^M \lambda_j\, x_j(u,\vec{\pi})\, p_j(y)}\right] p_i(y)\mu(dy) \\ = \int_{\mathbb{R}^d} \inf_{k\in K} \left[\beta_0^k \sum_{i=1}^M \lambda_i \pi_i e^{-\lambda_i u} p_i(y) + \sum_{i=1}^M \beta_i^k \lambda_i \pi_i e^{-\lambda_i u} p_i(y)\right] \mu(dy),\end{aligned}$$

where the last equality follows from the explicit form of $x(u,\vec{\pi})$ given in (2.8). Hence, the integrand and the integral in (3.5) are concave functions of $\vec{\pi}$. Similarly, using (2.8) and (2.4) we have $\left(\sum_{i=1}^M \pi_j e^{-\lambda_i t}\right) h\left(x(t,\vec{\pi})\right) = \inf_{k\in I}\ \sum_{j=1}^M a_{k,j}\,\pi_j\,e^{-\lambda_j t}$, and the second term in (3.5) is concave. Being the sum of two concave functions, $\vec{\pi}\mapsto Jw(t,\vec{\pi})$ is also concave. Finally, since $\vec{\pi}\mapsto J_0 w(\vec{\pi})$ is the infimum of concave functions, it is concave.

Next, we take $w : E \mapsto \mathbb{R}_+$ bounded and continuous, and we verify the continuity of the mappings $(t,\vec{\pi})\mapsto Jw(t,\vec{\pi})$ and $\vec{\pi}\mapsto J_0(\vec{\pi})$. To show that $(t,\vec{\pi})\mapsto Jw(t,\vec{\pi})$ of (3.5) is continuous, we first note that the mappings

$$(t,\vec{\pi})\mapsto \int_0^t \sum_{i=1}^M \pi_i e^{-\lambda_i u}\,du \qquad \text{and} \qquad (t,\vec{\pi})\mapsto \left(\sum_{i=1}^M \pi_i e^{-\lambda_i t}\right)\cdot h\left(x(t,\vec{\pi})\right)$$

are jointly continuous. For a bounded and continuous function $w(\cdot)$ on E, the mapping $\vec{\pi}\mapsto (S_i w)(\vec{\pi})$ is also bounded and continuous by bounded convergence theorem. Let $(t^{(k)},\vec{\pi}^{(k)})_{k\geq 1}$ be a sequence in $\mathbb{R}_+\times E$ converging to $(t,\vec{\pi})$. Then we have

$$\begin{aligned}
&\left|\int_0^t \sum_{i=1}^M e^{-\lambda_i u}\,\lambda_i\pi_i\cdot S_i w(x(u,\vec{\pi}))du - \int_0^{t^{(k)}} \sum_{i=1}^M e^{-\lambda_i u}\,\lambda_i\pi_i^{(k)}\cdot S_i w(x(u,\vec{\pi}^{(k)}))du\right| \\
&\leq \left|\int_0^t \sum_{i=1}^M e^{-\lambda_i u}\,\lambda_i\pi_i\cdot S_i w(x(u,\vec{\pi}))du - \int_0^{t^{(k)}} \sum_{i=1}^M e^{-\lambda_i u}\,\lambda_i\pi_i\cdot S_i w(x(u,\vec{\pi}))du\right| \\
&\qquad + \left|\int_0^{t^{(k)}} \sum_{i=1}^M e^{-\lambda_i u}\,\lambda_i\pi_i\cdot S_i w(x(u,\vec{\pi}))du - \int_0^{t^{(k)}} \sum_{i=1}^M e^{-\lambda_i u}\,\lambda_i\pi_i^{(k)}\cdot S_i w(x(u,\vec{\pi}^{(k)}))du\right| \\
&\leq \lambda_M\,||w||\cdot\left|t-t^{(k)}\right| + \int_0^{t^{(k)}} \sum_{i=1}^M e^{-\lambda_i u}\,\lambda_i\left|\pi_i\cdot S_i w(x(u,\vec{\pi})) - \pi_i^{(k)}\cdot S_i w(x(u,\vec{\pi}^{(k)}))\right|du \\
&\leq \lambda_M\,||w||\cdot\left|t-t^{(k)}\right| + \lambda_M\int_0^{\infty} e^{-\lambda_1 u}\sum_{i=1}^M \left|\pi_i\cdot S_i w(x(u,\vec{\pi})) - \pi_i^{(k)}\cdot S_i w(x(u,\vec{\pi}^{(k)}))\right|du
\end{aligned}$$

where $||w|| \triangleq \sup_{\vec{\pi}\in E}\,|w(\vec{\pi})| < \infty$. Letting $k\to\infty$, the first term of the last line above goes to 0. The second term also goes to 0 by the continuity of $\vec{\pi}\mapsto S_i w(x(u,\vec{\pi}))$ and dominated convergence theorem. Therefore, $Jw(t,\vec{\pi})$ is jointly continuous in $(t,\vec{\pi})$.

To see the continuity of $\vec{\pi}\mapsto J_0 w(\vec{\pi})$, let us define the "truncated" operators

$$J_0^k w(\vec{\pi}) \triangleq \inf_{t\in[0,k]} Jw(t,\vec{\pi}), \qquad \vec{\pi}\in E,\ k\geq 1$$

on bounded functions $w : E \mapsto \mathbb{R}$. Since $(t, \vec{\pi}) \mapsto Jw(t, \vec{\pi})$ is uniformly continuous on $[0, k] \times E$, the mapping $\vec{\pi} \mapsto J_0^k w(\vec{\pi})$ is also continuous on E. Also note that

$$\begin{aligned}
Jw(t,\vec{\pi}) &= Jw(t\wedge k,\vec{\pi}) + \int_{t\wedge k}^{t} \sum_{i=1}^{M} e^{-\lambda_i u}\, \pi_i \left[1+\lambda_i \cdot S_i w(x(u,\vec{\pi}))\right] du \\
&\qquad + \sum_{i=1}^{M} e^{-\lambda_i t}\, \pi_i \cdot h\left(x(t,\vec{\pi})\right) - \sum_{i=1}^{M} e^{-\lambda_i (t\wedge k)}\, \pi_i \cdot h\left(x(t\wedge k,\vec{\pi})\right) \\
&\geq Jw(t\wedge k,\vec{\pi}) - \mathbf{1}_{\{t>k\}} \sum_{i=1}^{M} e^{-\lambda_i k}\, \pi_i \cdot h\left(x(k,\vec{\pi})\right) \\
&\geq Jw(t\wedge k,\vec{\pi}) - \left(\sum_{i,j}^{M} a_{ij}\right) \sum_{i=1}^{M} e^{-\lambda_i k}\, \pi_i \geq Jw(t\wedge k,\vec{\pi}) - \left(\sum_{i,j}^{M} a_{ij}\right) e^{-\lambda_1 k},
\end{aligned}$$

since $0 \leq w(\cdot)$ and $0 \leq h(\cdot) \leq \sum_{i,j}^{M} a_{ij}$. Taking the infimum over $t \geq 0$ of both sides gives

$$J_0^k w(\vec{\pi}) \geq J_0 w(\vec{\pi}) \geq J_0^k w(\vec{\pi}) - \left(\sum_{i,j}^{M} a_{ij}\right) e^{-\lambda_1 k}.$$

Hence, $J_0^k w(\vec{\pi}) \to J_0 w(\vec{\pi})$ as $k \to \infty$ uniformly in $\vec{\pi} \in E$, and $\vec{\pi} \mapsto J_0 w(\vec{\pi})$ is continuous on the compact set E. □

Proof of Proposition 3.7. Here, the dependence of $r_n^{\varepsilon/2}$ on $\vec{\pi}$ is omitted for typographical reasons. Let us first establish the inequality

$$\mathbb{E}\left[\tau\wedge\sigma_n + h\big(\Pi(\tau\wedge\sigma_n)\big)\right] \geq v_n(\vec{\pi}), \qquad \tau \in \mathbb{F},\ \vec{\pi} \in E, \tag{A.2}$$

by proving inductively on $k = 1, \ldots, n+1$ that

$$\begin{aligned}
&\mathbb{E}^{\vec{\pi}}\left[\tau\wedge\sigma_n + h\left(\Pi(\tau\wedge\sigma_n)\right)\right] \\
&\quad \geq \mathbb{E}^{\vec{\pi}}\left[\tau\wedge\sigma_{n-k+1} + \mathbf{1}_{\{\tau\geq\sigma_{n-k+1}\}} v_{k-1}\left(\Pi(\sigma_{n-k+1})\right) + \mathbf{1}_{\{\tau<\sigma_{n-k+1}\}} h\left(\Pi(\tau)\right)\right] =: RHS_{k-1}.
\end{aligned} \tag{A.3}$$

Note that (A.2) will follow from (A.3) when we set $k = n+1$.

For $k = 1$, (A.3) is immediate since $v_0(\cdot) \equiv h(\cdot)$. Assume that it holds for some $1 \leq k < n+1$ and prove it for $k+1$. Note that RHS_{k-1} defined in (A.3) can be decomposed as

$$RHS_{k-1} = RHS_{k-1}^{(1)} + RHS_{k-1}^{(2)} \tag{A.4}$$

where $RHS_{k-1}^{(1)} \triangleq \mathbb{E}^{\vec{\pi}}\left[\tau\wedge\sigma_{n-k} + \mathbf{1}_{\{\tau<\sigma_{n-k}\}} h(\Pi(\tau))\right]$, and $RHS_{k-1}^{(2)}$ is defined by

$$\mathbb{E}^{\vec{\pi}}\left[\mathbf{1}_{\{\tau\geq\sigma_{n-k}\}}\left\{\tau\wedge\sigma_{n-k+1} - \sigma_{n-k} + \mathbf{1}_{\{\tau\geq\sigma_{n-k+1}\}} v_{k-1}(\Pi(\sigma_{n-k+1})) + \mathbf{1}_{\{\tau<\sigma_{n-k+1}\}} h(\Pi(\tau))\right\}\right].$$

By Lemma 3.3, there is an $\mathcal{F}_{\sigma_{n-k}}$-measurable random variable R_{n-k} such that $\tau \wedge \sigma_{n-k+1} = (\sigma_{n-k} + R_{n-k}) \wedge \sigma_{n-k+1}$ on $\{\tau \geq \sigma_{n-k}\}$. Therefore, $RHS^{(2)}_{k-1}$ becomes

$$\begin{aligned}
&\mathbb{E}^{\vec{\pi}} \Big[\mathbf{1}_{\{\tau\geq\sigma_{n-k}\}} \big\{(\sigma_{n-k} + R_{n-k}) \wedge \sigma_{n-k+1} - \sigma_{n-k} + \mathbf{1}_{\{\sigma_{n-k}+R_{n-k}\geq\sigma_{n-k+1}\}} v_{k-1}(\Pi(\sigma_{n-k+1})) \\
&\qquad\qquad + \mathbf{1}_{\{\sigma_{n-k}+R_{n-k}<\sigma_{n-k+1}\}} h(\Pi(\sigma_{n-k} + R_{n-k}))\big\}\Big] \\
&\quad = \mathbb{E}^{\vec{\pi}} \Big[\mathbf{1}_{\{\tau\geq\sigma_{n-k}\}} \mathbb{E}^{\vec{\pi}} \Big\{ (r \wedge \sigma_1) \circ \theta_{\sigma_{n-k}}\Big|_{r=R_{n-k}} + \mathbf{1}_{\{R_{n-k}\geq\sigma_{n-k+1}-\sigma_{n-k}\}} v_{k-1}(\Pi(\sigma_{n-k+1})) \\
&\qquad\qquad + \mathbf{1}_{\{R_{n-k}<\sigma_{n-k+1}-\sigma_{n-k}\}}\, h(\Pi(\sigma_{n-k} + R_{n-k}))\big|\, \mathcal{F}_{\sigma_{n-k}} \Big\}\Big] .
\end{aligned}$$

Since Π has the strong Markov property and the same jump times σ_k, $k \geq 1$ as the process X, the last expression for $RHS^{(2)}_{k-1}$ can be rewritten as

$$RHS^{(2)}_{k-1} = \mathbb{E}^{\vec{\pi}} \left[\mathbf{1}_{\{\tau\geq\sigma_{n-k}\}}\, f_{n-k}(R_{n-k}, \Pi(\sigma_{n-k}))\right] ,$$

where $f_{n-k}(r, \vec{\pi}) \triangleq \mathbb{E}^{\vec{\pi}} \left[r \wedge \sigma_1 + \mathbf{1}_{\{r>\sigma_1\}} v_{k-1}(\Pi(\sigma_1)) + \mathbf{1}_{\{r\leq\sigma_1\}} h(\Pi(r))\right] \equiv Jv_{k-1}(r, \vec{\pi})$. Thus, $f_{n-k}(r, \vec{\pi}) \geq J_0 v_{k-1}(\vec{\pi}) = v_k(\vec{\pi})$, and $RHS^{(2)}_{k-1} \geq \mathbb{E}\left[\mathbf{1}_{\{\tau>\sigma_{n-k}\}}\, v_k(\Pi(\sigma_{n-k}))\right]$. Using this inequality with (A.3) and (A.4) we get

$$\begin{aligned}
&\mathbb{E}^{\vec{\pi}} \left[\tau \wedge \sigma_n + h\left(\Pi(\tau \wedge \sigma_n)\right)\right] \geq RHS_{k-1} \\
&\qquad\qquad \geq \mathbb{E}^{\vec{\pi}} \left[\tau \wedge \sigma_{n-k} + \mathbf{1}_{\{\tau<\sigma_{n-k}\}} h\left(\Pi(\tau)\right) + \mathbf{1}_{\{\tau\geq\sigma_{n-k}\}} v_k\left(\Pi(\tau)\right)\right] = RHS_k.
\end{aligned}$$

This proves (A.3) for $k+1$ instead of k, and (A.2) follows after induction for $k = n+1$. Taking the infimum on the left-hand side of (A.2) over all $\mathbb{F}$-stopping times τ gives $V_n(\cdot) \geq v_n(\cdot)$.

To prove the reverse inequality $V_n(\cdot) \geq v_n(\cdot)$, it is enough to show (3.9) since by construction S^{ε}_n is an $\mathbb{F}$-stopping time. We will prove (3.9) also inductively. For $n = 1$, the left hand side of (3.9) reduces to

$$\mathbb{E}^{\vec{\pi}} \left[r^{\varepsilon}_0 \wedge \sigma_1 + h\left(\Pi(r^{\varepsilon}_0 \wedge \sigma_1)\right)\right] = Jv_0(r^{\varepsilon}_0, \vec{\pi}) \leq J_0 v_0(\vec{\pi}) + \varepsilon,$$

where the inequality follows from Remark 3.6, and the inequality in (3.9) holds with $n = 1$.

Suppose now that (3.9) holds for some $n \geq 1$. By definition, $S^{\varepsilon}_{n+1} \wedge \sigma_1 = r^{\varepsilon/2}_n \wedge \sigma_1$, and

$$\mathbb{E}^{\vec{\pi}}\left[S^{\varepsilon}_{n+1} + h\left(\Pi(S^{\varepsilon}_{n+1})\right)\right] = \mathbb{E}^{\vec{\pi}}\left[S^{\varepsilon}_{n+1} \wedge \sigma_1 + (S^{\varepsilon}_{n+1} - \sigma_1)\,\mathbf{1}_{\{S^{\varepsilon}_{n+1} \geq \sigma_1\}} + h\left(\Pi(S^{\varepsilon}_{n+1})\right)\right]$$

$$= \mathbb{E}^{\vec{\pi}}\left[r^{\varepsilon/2}_n \wedge \sigma_1 + \left(S^{\varepsilon/2}_n \circ \theta_{\sigma_1} + h\left(\Pi(\sigma_1 + S^{\varepsilon/2}_n \circ \theta_{\sigma_1})\right)\right)\mathbf{1}_{\{r^{\varepsilon/2}_n \geq \sigma_1\}} + \mathbf{1}_{\{r^{\varepsilon/2}_n < \sigma_1\}} h\left(\Pi(r^{\varepsilon/2}_n)\right)\right]$$

$$= \mathbb{E}^{\vec{\pi}}\left[r^{\varepsilon/2}_n \wedge \sigma_1 + \mathbf{1}_{\{r^{\varepsilon/2}_n < \sigma_1\}} h\left(\Pi(r^{\varepsilon/2}_n)\right) + \mathbf{1}_{\{r^{\varepsilon/2}_n \geq \sigma_1\}} \mathbb{E}^{\vec{\pi}}\left[\left(S^{\varepsilon/2}_n + h(\Pi(S^{\varepsilon/2}_n))\right) \circ \theta_{\sigma_1} \middle| \mathcal{F}_{\sigma_1}\right]\right]$$

$$\leq \mathbb{E}^{\vec{\pi}}\left[r^{\varepsilon/2}_n \wedge \sigma_1 + \mathbf{1}_{\{r^{\varepsilon/2}_n < \sigma_1\}} h\left(\Pi(r^{\varepsilon/2}_n)\right) + \mathbf{1}_{\{r^{\varepsilon/2}_n \geq \sigma_1\}} v_n\left(\Pi_{\sigma_1}\right)\right] + \frac{\varepsilon}{2} = Jv_n(r^{\varepsilon/2}_n, \vec{\pi}) + \frac{\varepsilon}{2},$$

where the inequality follows from the induction hypothesis and strong Markov property. Hence, $\mathbb{E}^{\vec{\pi}}\left[S^{\varepsilon}_{n+1} + h\left(\Pi(S^{\varepsilon}_{n+1})\right)\right] \leq v_n(\vec{\pi}) + \varepsilon$ by Remark 3.6, and (3.9) holds for $n+1$. □

Proof of Proposition 3.9. Note that the sequence $\{v_n\}_{n\geq 1}$ is decreasing. Therefore,

$$V(\vec{\pi}) = v(\vec{\pi}) = \inf_{n\geq 1} v_n(\vec{\pi}) = \inf_{n\geq 1} J_0 v_{n-1}(\vec{\pi}) = \inf_{n\geq 1}\inf_{t\geq 0} Jv_{n-1}(t, \vec{\pi}) = \inf_{t\geq 0}\inf_{n\geq 1} Jv_{n-1}(t, \vec{\pi})$$

$$= \inf_{t\geq 0}\inf_{n\geq 1}\left[\int_0^t \sum_{i=1}^M \pi_i e^{-\lambda_i u}\left[1 + \lambda_i S_i v_{n-1}(x(u, \vec{\pi}))\right] du + \sum_{i=1}^M e^{-\lambda_i t}\,\pi_i h\left(x(t, \vec{\pi})\right)\right]$$

$$= \inf_{t\geq 0}\left[\int_0^t \sum_{i=1}^M \pi_i e^{-\lambda_i u}\left[1 + \lambda_i S_i v(x(u, \vec{\pi}))\right] du + \sum_{i=1}^M e^{-\lambda_i t}\,\pi_i h\left(x(t, \vec{\pi})\right)\right] = J_0 v(\vec{\pi}) = J_0 V(\vec{\pi}),$$

where the seventh equality follows from the bounded convergence theorem since $0 \leq v(\cdot) \leq v_n(\cdot) \leq h(\cdot) \leq \sum_{i,j} a_{ij}$. Thus, V satisfies $V = J_0 V$.

Next, let $U(\cdot) \leq h(\cdot)$ be a bounded solution of $U = J_0 U$. Proposition 3.4 implies that $U = J_0 U \leq J_0 h = v_1$. Suppose that $U \leq v_n$ for some $n \geq 0$. Then $U = J_0 U \leq J_0 v_n = v_{n+1}$, and by induction we have $U \leq v_n$ for all $n \geq 1$. This implies $U \leq \lim_{n\to\infty} v_n = V$. □

Proof of Corollary 3.12. For typographical reasons, we will again omit the dependence of $r_n(\vec{\pi})$ on $\vec{\pi}$. By Remark 3.6, $Jv_n(r_n, \vec{\pi}) = J_0 v_n(\vec{\pi}) = J_{r_n} v_n(\vec{\pi})$. Let us first assume that $r_n < \infty$. Taking $t = r_n$ and $w = v_n$ in (3.11) gives

$$Jv_n(r_n, \vec{\pi}) = J_{r_n} v_n(\vec{\pi}) = Jv_n(r_n, \vec{\pi}) + \sum_{i=1}^M \pi_i\, e^{-\lambda_i r_n}\Big[v_{n+1}(x(r_n, \vec{\pi})) - h\big(x(r_n, \vec{\pi})\big)\Big].$$

Therefore, $v_{n+1}(x(r_n,\vec{\pi})) < h(x(r_n,\vec{\pi}))$. If $0 < t < r_n$, then $Jv_n(t,\vec{\pi}) > J_0 v_n(\vec{\pi}) = J_{r_n} v_n(\vec{\pi}) = J_t v_n(\vec{\pi})$. Using (3.11) once again with $w = v_n$ and $t \in (0, r_n)$ yields

$$J_0 v_n(\vec{\pi}) = J_t v_n(\vec{\pi}) = Jv_n(t,\vec{\pi}) + \sum_{i=1}^{M} \pi_i\, e^{-\lambda_i t}\Big[v_{n+1}(x(t,\vec{\pi})) - h\big(x(t,\vec{\pi})\big)\Big].$$

Hence $v_{n+1}(x(t,\vec{\pi})) < h(x(t,\vec{\pi}))$ for $t \in (0, r_n)$. If $r_n = \infty$, then the same argument as above implies that $v_{n+1}(x(t,\vec{\pi})) < h(x(t,\vec{\pi}))$ for $t \in [0,\infty]$, and (3.13) still holds since $\inf \varnothing \equiv \infty$. □

Proof of Proposition 3.15. We will prove (3.20) inductively. For $n = 1$, by Lemma 3.3 there exists a constant $u \in [0,\infty]$ such that $U_\varepsilon \wedge \sigma_1 = u \wedge \sigma_1$. Then

$$\begin{aligned}
&\text{(A.5)}\quad \mathbb{E}^{\vec{\pi}}[M_{U_\varepsilon\wedge\sigma_1}] = \mathbb{E}^{\vec{\pi}}\left[u\wedge\sigma_1 + V\left(\Pi(u\wedge\sigma_1)\right)\right]\\
&\quad = \mathbb{E}^{\vec{\pi}}\Big[u\wedge\sigma_1 + \mathbf{1}_{\{u\ge\sigma_1\}} V\left(\Pi(\sigma_1)\right) + \mathbf{1}_{\{u<\sigma_1\}} h\left(\Pi(u)\right)\Big] + \mathbb{E}^{\vec{\pi}}\Big[\mathbf{1}_{\{u<\sigma_1\}}\left[V\left(\Pi(u)\right) - h\left(\Pi(u)\right)\right]\Big]\\
&\quad = JV(u,\vec{\pi}) + \sum_{i=1}^{M}\pi_i\, e^{-\lambda_i u}\left[V\big(x(u,\vec{\pi})\big) - h\big(x(u,\vec{\pi})\big)\right] = J_u V(\vec{\pi})
\end{aligned}$$

because of the delay equation in (3.14). Fix any $t \in [0,u)$. Same equation implies that $JV(t,\vec{\pi}) = J_t V(\vec{\pi}) - \sum_{i=1}^{M}\pi_i\, e^{-\lambda_i t}\big[V\left(x(t,\vec{\pi})\right) - h\big(x(t,\vec{\pi})\big)\big]$ is less than or equal to

$$J_0 V(\vec{\pi}) - \sum_{i=1}^{M}\pi_i\, e^{-\lambda_i t}\big[V\left(x(t,\vec{\pi})\right) - h\big(x(t,\vec{\pi})\big)\big] = J_0 V(\vec{\pi}) - \mathbb{E}^{\vec{\pi}}\left[\mathbf{1}_{\{\sigma_1>t\}}\big[V(\Pi(t)) - h(\Pi(t))\big]\right].$$

On $\{\sigma_1 > t\}$, we have $U_\varepsilon > t$ (otherwise, $U_\varepsilon \le t < \sigma_1$ would imply $U_\varepsilon = u \le t$, which contradicts with our initial choice of $t < u$). Thus, $V(\Pi(t)) - h(\Pi(t)) < -\varepsilon$ on $\{\sigma_1 > t\}$, and

$$JV(t,\vec{\pi}) \ge J_0 V(\vec{\pi}) + \varepsilon\,\mathbb{E}^{\vec{\pi}}\left[\mathbf{1}_{\{\sigma_1>t\}}\right] \ge J_0 V(\vec{\pi}) + \varepsilon\sum_{i=1}^{M}\pi_i\, e^{-\lambda_i t} \quad \text{for every } t \in [0,u).$$

Therefore, $J_0 V(\vec{\pi}) = J_u V(\vec{\pi})$, and (A.5) implies $\mathbb{E}^{\vec{\pi}}[M_{U_\varepsilon\wedge\sigma_1}] = J_u V(\vec{\pi}) = J_0 V(\vec{\pi}) = V(\vec{\pi}) = \mathbb{E}^{\vec{\pi}}[M_0]$. This completes the proof of (3.20) for $n = 1$. Now assume that (3.20) holds for some $n \ge 1$. Note that $\mathbb{E}^{\vec{\pi}}[M_{U_\varepsilon\wedge\sigma_{n+1}}] = \mathbb{E}^{\vec{\pi}}[\mathbf{1}_{\{U_\varepsilon<\sigma_1\}} M_{U_\varepsilon}] + \mathbb{E}^{\vec{\pi}}[\mathbf{1}_{\{U_\varepsilon\ge\sigma_1\}} M_{U_\varepsilon\wedge\sigma_{n+1}}]$ equals

$$\mathbb{E}^{\vec{\pi}}[\mathbf{1}_{\{U_\varepsilon<\sigma_1\}} M_{U_\varepsilon}] + \mathbb{E}^{\vec{\pi}}\left[\mathbf{1}_{\{U_\varepsilon\ge\sigma_1\}}\left\{\sigma_1 + U_\varepsilon\wedge\sigma_{n+1} - \sigma_1 + V\left(\Pi(U_\varepsilon\wedge\sigma_{n+1})\right)\right\}\right].$$

Since $U_\varepsilon \wedge \sigma_{n+1} = \sigma_1 + [(U_\varepsilon\wedge\sigma_n)\circ\theta_{\sigma_1}]$ on the event $\{U_\varepsilon \ge \sigma_1\}$, the strong Markov property of Π implies that $\mathbb{E}^{\vec{\pi}}[M_{U_\varepsilon\wedge\sigma_{n+1}}]$ equals

$$\mathbb{E}^{\vec{\pi}}[\mathbf{1}_{\{U_\varepsilon<\sigma_1\}} M_{U_\varepsilon}] + \mathbb{E}^{\vec{\pi}}[\mathbf{1}_{\{U_\varepsilon\ge\sigma_1\}}\sigma_1] + \mathbb{E}^{\vec{\pi}}\left[\mathbf{1}_{\{U_\varepsilon\ge\sigma_1\}}\mathbb{E}^{\Pi(\sigma_1)}\left[U_\varepsilon\wedge\sigma_n + V\left(\Pi(U_\varepsilon\wedge\sigma_n)\right)\right]\right].$$

By the induction hypothesis the inner expectation equals $V(\Pi(\sigma_1))$, and

$$\mathbb{E}^{\vec{\pi}}[M_{U_\varepsilon\wedge\sigma_{n+1}}] = \mathbb{E}^{\vec{\pi}}[\mathbf{1}_{\{U_\varepsilon<\sigma_1\}}M_{U_\varepsilon}] + \mathbb{E}^{\vec{\pi}}\left[\mathbf{1}_{\{U_\varepsilon\geq\sigma_1\}}\left(\sigma_1 + V(\Pi(\sigma_1))\right)\right] = \mathbb{E}^{\vec{\pi}}[M_{U_\varepsilon\wedge\sigma_1}] = \mathbb{E}^{\vec{\pi}}[M_0],$$

where the last inequality follows from the above proof for $n = 1$. Hence, $\mathbb{E}^{\vec{\pi}}[M_{U_\varepsilon\wedge\sigma_{n+1}}] = \mathbb{E}^{\vec{\pi}}[M_0]$, and the proof is complete by induction. □

Proof of Proposition 4.1. By Proposition 3.4 we have $0 \leq J_0 w(\cdot) \leq h(\cdot)$ for every bounded and positive function $w(\cdot)$. To prove (4.4), it is therefore enough to show that $h(\cdot) \leq J_0 w(\cdot)$ on $\{\vec{\pi} \in E : \pi_i \geq \overline{p}\}$ for $\overline{p}$ defined in the lemma. Since $w(\cdot) \geq 0$, we have $J_0 w(\vec{\pi}) = \inf_{t\geq 0} Jw(t,\vec{\pi}) \geq \inf_{t\geq 0}\inf_{i\in I} K_i(t,\vec{\pi})$, where

$$K_i(t,\vec{\pi}) \triangleq \int_0^t \sum_{j=1}^M \pi_j e^{-\lambda_j u}\, du + \sum_{j=1}^M e^{-\lambda_j t} h_i(x(t,\vec{\pi})),$$

with the partial derivative

$$\text{(A.6)}\quad \frac{\partial K_i(t,\vec{\pi})}{\partial t} = \sum_{j=1}^M \pi_j e^{-\lambda_j t}\Bigg\{1 - \sum_{k=1}^M (\lambda_j + \lambda_k)\, a_{ik}\, x_k(t,\vec{\pi})$$
$$+\left(\sum_{k=1}^M a_{ik}\, x_k(t,\vec{\pi})\right)\left(\sum_{k=1}^M \lambda_k\, x_k(t,\vec{\pi})\right)\Bigg\} \geq \sum_{j=1}^M \pi_j e^{-\lambda_j t}\left\{1 - 2\,\lambda_M\,(\max_k a_{ik})(1 - x_i(t,\vec{\pi}))\right\}.$$

Note that if $\lambda_i = \lambda_1$, then $t \mapsto x_i(t,\vec{\pi})$ is non-decreasing. Hence $\partial K_i(t,\vec{\pi})/\partial t \geq 0$ for all $t \geq 0$ and $\pi_i \geq \overline{\pi}_i \triangleq (1 - 1/(2\,\lambda_M\, \max_k a_{i,k}))^+$. Hence on $\{\vec{\pi} \in E : \pi_i \geq \overline{\pi}_i\}$, we get $h(\vec{\pi}) \geq J_0 w(\vec{\pi}) \geq \inf_{t\geq 0}\inf_i K_i(t,\vec{\pi}) = \inf_i \inf_{t\geq 0} K_i(t,\vec{\pi}) = \inf_i K_i(0,\vec{\pi}) = \inf_i h_i(\vec{\pi}) = h(\vec{\pi})$.

Now assume $\lambda_i > \lambda_1$, and let $T_i(\vec{\pi},m) \triangleq \inf\{t \geq 0 : x_i(t,\vec{\pi}) \leq m\}$ be the the first time $t \mapsto x_i(t,\vec{\pi})$ reaches $[0,m]$. For $t \geq T_i(\vec{\pi},\overline{\pi}_i)$, the definition of $K_i(t,\vec{\pi})$ implies

$$K_i(t,\vec{\pi}) \geq \pi_i \int_0^{T_i(\vec{\pi},\overline{\pi}_i)} e^{-\lambda_i u} du \geq \frac{\pi_i}{\lambda_i}\left[1 - \left(\frac{1-\overline{\pi}_i}{\overline{\pi}_i}\cdot\frac{\pi_i}{1-\pi_i}\right)^{-\lambda_i/(\lambda_i-\lambda_1)}\right],$$

where the last inequality follows from the explicit form of $x(t,\vec{\pi})$ given in (2.8). The last expression above is 0 at $\pi_i = \overline{\pi}_i$, and it is increasing on $\pi_i \in [\overline{\pi}_i, 1)$ and increases to the limit $1/\lambda_i$ as $\pi_i \to 1$. For $\pi_i \geq \pi_i^*$ in (4.3), we have

$$\frac{\pi_i}{\lambda_i}\left[1 - \left(\frac{1-\overline{\pi}_i}{\overline{\pi}_i}\cdot\frac{\pi_i}{1-\pi_i}\right)^{-\lambda_i/(\lambda_i-\lambda_1)}\right] \geq (\max_k a_{ik})(1-\pi_i),$$

and the inequality holds with an equality at $\pi_i = \pi_i^*$. Hence, we have $K_i(t,\vec{\pi}) \geq h_i(\vec{\pi})$ for $t \geq T_i(\vec{\pi},\overline{\pi}_i)$, and for $\pi_i \geq \pi_i^*$. Since $t \mapsto K_i(t,\vec{\pi})$ is increasing on $[0, T_i(\vec{\pi},\overline{\pi}_i)]$, we get

$h(\vec{\pi}) \geq J_0 w(\vec{\pi}) \geq \inf_i \inf_{t\geq 0} K_i(t,\vec{\pi}) = \inf_i h_i(\vec{\pi}) = h(\vec{\pi})$ on $\{\vec{\pi} \in E : \pi_i \geq \pi_i^*\}$, and (4.4) follows. Finally, if $a_{ij} > 0$ for every $i \neq j$, then (4.5) follows from (4.4) since $h(\vec{\pi}) = h_i(\vec{\pi})$ on $\{\vec{\pi} \in E : \pi_i \geq \max_k a_{ik}/[(\min_{j\neq i} a_{ji}) + (\max_k a_{ik})]\}$ for every $i \in I$. □

References

Blackwell, D. and Girshick, M. (1954). *Theory of Games and Statistical Decisions*, New York, John Wiley & Sons.

Brémaud, P. (1981). *Point Processes and Queues*, Martingale dynamics, Springer Series in Statistics, Springer-Verlag, New York.

Davis, M. H. A. (1993). *Markov Models and Optimization*, Vol. 49 of *Monographs on Statistics and Applied Probability*, Chapman & Hall, London.

Dayanik, S. and Sezer, S. (2005). Sequential testing of simple hypotheses about compound poisson processes, *Stoch. Proces. and Appl., to appear. (www.princeton.edu/∼sdayanik/papers/hypothesis.pdf).*

Gapeev, P. (2002). Problems of the sequential discrimination of hypotheses for a compound poisson process with exponential jumps, *Uspekhi Mat. Nauk* **57**(6(348)): 171–172.

Peskir, G. and Shiryaev, A. (2000). Sequential testing problems for poisson processes, *Annals of Statistics* **28**(3): 837–859.

Shiryaev, A. (1978). *Optimal Stopping Rules*, Springer-Verlag, New York.

Wald, A. and Wolfowitz, J. (1950). Bayes solutions of sequential decision problems, *Ann. Math. Statist.* **21**: 82–99.

Zacks, S. (1971). *The Theory of Statistical Inference*, New York, John Wiley & Sons.

(S. Dayanik) Department of Operations Research and Financial Engineering, and the Bendheim Center for Finance, Princeton University, Princeton, NJ 08544

E-mail address: sdayanik@princeton.edu

(H. V. Poor) School of Engineering and Applied Science, Princeton University, Princeton, NJ 08544

E-mail address: poor@princeton.edu

(S. O. Sezer) Department of Operations Research and Financial Engineering, and the Bendheim Center for Finance, Princeton University, Princeton, NJ 08544

E-mail address: ssezer@princeton.edu