FINITE HORIZON DECISION TIMING WITH PARTIALLY OBSERVABLE POISSON PROCESSES

MICHAEL LUDKOVSKI AND SEMIH O. SEZER

ABSTRACT. We study decision timing problems on finite horizon with Poissonian information arrivals. In our model, a decision maker wishes to optimally time her action in order to maximize her expected reward. The reward depends on an unobservable Markovian environment, and information about the environment is collected through a (compound) Poisson observation process. Examples of such systems arise in investment timing, reliability theory, Bayesian regime detection and technology adoption models. We solve the problem by studying an optimal stopping problem for a piecewise-deterministic process, which gives the posterior likelihoods of the unobservable environment. Our method lends itself to simple numerical implementation and we present several illustrative numerical examples.

1. INTRODUCTION

Decision timing under uncertainty is one of the fundamental problems in Operations Research. In a typical setting, an economic agent (called the decision-maker or DM) has a set of possible actions $A$ where each action has a (random) reward associated with it. The objective of the DM is to select a single action and time it so as to maximize her expected reward. More precisely, the DM picks a stopping time $\tau$ and an action $k$ from the set $A$ at $\tau$. The reward $H$ that DM receives is a function of the pair $(\tau, k)$, as well as of some stochastic state variable $Y$. In classical examples (e.g. investment timing, American option pricing, natural resource management, etc.), $Y$ is an observable stochastic process (e.g. asset prices, market demand etc.), and the DM’s objective is a standard optimal stopping problem.

More complicated stopping problems involving unobserved system states have also been considered in the literature; see, for example, Bather [1973], Monahan [1980], Jensen [1982], McCardle [1985], Mazziotto [1986], Jensen and Hsu [1993], Stadje [1997], Schöttl [1998], Fuh [2003], Décamps et al. [2005], Dayanik and Goulding [2009]. Such models are especially natural when one wishes to capture the inherent conflict between gathering information (which makes waiting valuable) and the time-value of money (which makes waiting costly). Indeed, most realistic settings involve a DM who is only partially aware of the environment and must collect data before making a decision. In a multi-period setting, it is natural to capture this uncertainty in the environment through an unobservable stochastic process $M = \{M_t\}_{t \geq 0}$, where $M_t$ represents the state of the world at time $t$. The DM starts with an initial guess about $M$, collects information via relevant news, and updates her beliefs. At the time of decision she then receives a reward that depends on the present environment, $H = H(\tau, k, M_\tau)$.

In such problems, a common approach is to postulate that the process $M$ is a partially observable Markov (decision) process (POMDP), in which case we have a hidden Markov model (HMM). Such models have been studied extensively both in discrete- and continuous-time. The reader may refer
to Bertsekas [1976], Monahan [1982], Elliott et al. [1995] and Cappé et al. [2005] for a comprehensive treatment of discrete-time models and also for many other references on the subject. For continuous-time models and applications, Bensoussan [1992], Liptser and Shiryaev [2001], Mamon and Elliott [2007] (and Elliott et al. [1995]), and the references therein can be consulted.

In continuous-time models, if news (such as changes in asset prices) arrive in infinitesimal amounts, then it is intuitive to have a continuum of information, which is typically captured by the filtration of an observed diffusion process. However, in many instances, a more realistic representation is to use “discrete” information amounts. Corporate developments, engineering failures, insurance claims, and economic surveys are all discrete events and the corresponding news arrive in “chunks”. Note that discreteness of information is distinct from the discreteness of time. The model is still in continuous-time, since the events may take place at any instance. However, the event itself carries a strictly positive amount of information. Moreover, “no news” is still informative and affects the beliefs of the DM.

Mathematically, discrete information in continuous-time may be represented by the filtration of an observed marked point process (MPP). In such a model, the instantaneous arrival intensity and the distribution of the marks typically depend on the current state of the process \( M \). That is, the observable point process encodes information about the hidden environment \( M \) via its arrival times and/or marks. Filtering with continuous-time point process observations has been considered in Bremaud [1981], Arjas et al. [1992], Elliott and Malcolm [2005], and it is known that the dynamics of the conditional probabilities of \( M \) are of the piecewise deterministic process (PDP) type. In other words, the DM’s beliefs evolve deterministically between arrivals of new information, and experience random jumps at event times. From the control perspective, various aspects of optimal stopping of PDP’s have been studied by Lenhart and Liao [1985], Gugerli [1986] and Costa and Davis [1988].

In this paper, we study a class of finite-horizon decision-making problems within the PDP framework by considering a general partially-observable regime-switching model with Poisson information arrivals. More precisely, we consider a setting where the observations of DM come from a compound Poisson process \( X \) with arrival rate \( \lambda \), and mark/jump distribution \( \nu \). The local characteristics \((\lambda, \nu)\) of \( X \) are modulated by the current state of an unobservable finite-state Markov process \( M \).

In this setting, the DM can stop at any time \( \tau \) less than some horizon \( T < \infty \) and select an action from a set \( A \triangleq \{1, \ldots, a\} \). Action \( k \in A \) yields a terminal reward (cost) equal to \( \sum_{i \in E} \mu_{k,i} 1_{\{M_\tau = i\}} \), as a function of the unobservable state of \( M \). Here, \( \mu_{k,i} \) is a given finite number, which can also be interpreted as the expected value of an independent random variable \( \Phi_{k,i} \) representing the uncertain payoff of taking action \( k \) when \( M_t = i \).

The DM may alternatively delay her decision and continue to observe the process \( X \) in order to collect more information, or in order to stop later when \( M \) appears to be in a better state. Delaying the decision carries penalties (rewards) due to the cost of observation or lost opportunity (or operating revenues). We allow these terms to depend on \( M \) and we assume that an amount with present value \( \int_0^\tau e^{-\rho t} \left( \sum_{i \in E} c_i 1_{\{M_t = i\}} \right) dt \) is accumulated until the decision time \( \tau \). Here \( \rho \geq 0 \) is the discount factor, and \( c_i \) is the instantaneous cost or revenue of running the system when \( M \) is in state \( i \in E \). Also, we allow \( \rho \) to be zero. This makes the formulation suitable for non-financial applications where the quality of the decision is more important than its timing.

In this setup, the objective of the DM is to find an admissible pair \((\tau, d)\) that will maximize her total expected reward and resolve the trade-off between exploring (getting more observations) and exploiting (engaging in an action). Since the DM collects information by observing \( X \), \( \tau \) must be a
stopping time of the filtration $\mathcal{F}^X$ generated by $X$. Also, the decision $d$ should be measurable with respect to the information $\mathcal{F}^X_\tau$ revealed until $\tau$. Let $\bar{\pi} = (\pi_1, \ldots, \pi_n) \overset{\Delta}{=} (\mathbb{P}(M_0 = 1), \ldots, \mathbb{P}(M_0 = n))$ denote the prior beliefs of the DM about the initial state of $M$, and let $\mathbb{P}^{\bar{\pi}}$ be the corresponding probability measure. Then, the objective of the DM is to compute

\begin{equation}
U(T, \bar{\pi}) \overset{\Delta}{=} \sup_{\tau \leq T, \; d} \mathbb{E}^{\bar{\pi}} \left[ \int_0^\tau e^{-\rho t} \left( \sum_{i \in E} c_i 1\{M_t = i\} \right) \, dt + e^{-\rho \tau} \sum_{k \in A} 1\{d = k\} \left( \sum_{i \in E} \mu_{k,i} \cdot 1\{M_\tau = i\} \right) \right],
\end{equation}

and, if it exists, find a pair $((\tau, d))$ attaining this value.

In our paper, we solve the problem in (1.1) in its general form without any restrictive assumption. We give a full characterization of the value function with a direct proof of the dynamic programming principle. We also identify optimal and $\epsilon$-optimal policies for the DM. Moreover, we study the qualitative properties of the solution structure and provide a numerical approach that can be readily implemented; see Sections 4 and 5.

Special cases of this optimal stopping problem have been considered by Jensen and Hsu [1993] in connection with system reliability studies, Jensen [1997] and Schöttl [1998] in the context of insurance premium re-pricing and Peskir and Shiryaev [2000], Gapeev [2002], Bayraktar et al. [2006], Dayanik et al. [2008a] for classical Poisson disorder and regime detection problems. This line of work links together the MPP filtering literature with the PDP results of Lenhart and Liao [1985], Gugerli [1986], Costa and Davis [1988]. Here, we extend these works in three major directions. First, we consider a general continuous-time finite-state Markov chain for the environment $M$, and impose no restrictions on the arrival rate and mark distribution of the observed MPP $X$. Second, we consider a general discount/cost structure, that can be used to encode a variety of economic objectives. So far, all the aforementioned papers have dealt only with special cases by imposing additional assumptions on either $X$ or $M$ or $c_i, \rho, \mu_{k,i}$'s. Finally, we work in the context of finite horizon, where value functions are time-inhomogeneous. The introduction of time-to-maturity as a state variable adds analytical complexity and leads to the appearance of novel effects that are not possible with infinite horizon stationary models.

In Section 5, we illustrate the strength and wide-applicability of our approach on two key applications. In Section 5.1, we revisit the machine replacement problem of Jensen and Hsu [1993] in the finite horizon setting and without their assumptions. Next, in Section 5.2, we give the solution of the finite horizon formulation of the hypothesis-testing problem studied in Peskir and Shiryaev [2000], Gapeev [2002] and Dayanik et al. [2008a]. In the Bayesian sequential analysis literature, continuous-time change-detection and hypothesis-testing problems have attracted considerable interest, especially for Poisson and Wiener processes; see for example Dayanik et al. [2008a,b] and the references therein. Earlier works in this field study these two problems on the infinite horizon. One exception is Gapeev and Peskir [2004], which solves the finite horizon change-detection problem for the Wiener process long after its infinite horizon formulation was solved by Shiryaev [1978]. Our analysis in Sections 3 and 4 gives the solutions of these problems for the compound Poisson process as an immediate corollary, and this is another contribution of our paper.

This paper is organized as follows. Below in Section 2, we describe the formal setting of our model and then show that the problem in (1.1) is equivalent to an optimal stopping problem in terms of the conditional probability process, which is a piecewise deterministic process. In Section 3, we describe how the value function of this stopping problem can be computed via a sequential procedure. The results of Section 3 are used in Section 4 in order to identify an optimal strategy and study its properties. Following this, in Section 5 we give examples illustrating our results.
Finally, Appendices at the end include supplementary proofs and additional remarks. Appendix A extends our model to the case of discrete costs incurred at each event time, and Appendix B comments on the relationship between finite- and infinite-horizon problems and optimal controls.

2. Problem Statement

2.1. Model. Let \((\Omega, \mathcal{H}, \mathbb{P})\) be a probability space hosting a continuous-time Markov process \(M\) taking values on \(E \triangleq \{1, \ldots, n\}\), for \(n \in \mathbb{N}\), and with infinitesimal generator \(Q = (q_{ij})_{i,j \in E}\). Also, we have a collection of independent compound Poisson processes \(X^{(1)}, \ldots, X^{(n)}\) with local parameters \((\lambda_1, \nu_1), \ldots, (\lambda_n, \nu_n)\) respectively. In terms of these independent processes, we define the observation process

\[
X_t \triangleq X_0 + \int_{(0,t]} \sum_{i \in E} 1_{\{M_s = i\}} dX_s^{(i)}, \quad t \geq 0,
\]

which is a Markov-modulated Poisson process, also called a Cox process (see Cox and Isham [1980] and Grandell [1976]). In the remainder, we let \(\sigma_0, \sigma_1, \ldots\) denote the arrival times of the process \(X\):

\[
\sigma_m \triangleq \inf\{t > \sigma_{m-1} : X_t \neq X_{t-}\}, \quad m \geq 1, \quad \text{with} \ \sigma_0 \equiv 0,
\]

and the variables \(Y_1, Y_2, \ldots\) denote \(\mathbb{R}^d\)-valued marks observed at these arrival times: \(Y_m = X_{\sigma_m} - X_{\sigma_{m-1}}\), \(m \geq 1\). Finally, to compute relative likelihoods of different marks, we introduce the measure \(\nu\) defined as \(\nu \triangleq \nu_1 + \ldots + \nu_n\), and we let \(f_i(\cdot)\) denote the density of \(\nu_i\) with respect to \(\nu\).

2.2. Conditional probability process. For a point in \(D \triangleq \{ \vec{\pi} \in \mathbb{R}^n_+ : \pi_1 + \ldots + \pi_n = 1 \}\), let \(\mathbb{P}^{\vec{\pi}}\) denote the probability measure (with the expectation operator \(\mathbb{E}^{\vec{\pi}}\)) under which \(M\) has initial distribution \(\vec{\pi}\). Moreover, let \(\mathbb{F} \triangleq \{\mathcal{F}_t^X\}_{t \geq 0}\) be the filtration of the process \(X\) in (2.1). With this notation, we define the \(D\)-valued \emph{conditional probability process} \(\vec{\Pi}_t \triangleq (\Pi_t^{(1)}, \ldots, \Pi_t^{(n)})\) such that

\[
\Pi_t^{(i)} = \mathbb{P}^{\vec{\pi}}\{M_t = i|\mathcal{F}_t^X\}, \quad \text{for} \ i \in E, \ \text{and} \ t \geq 0.
\]

The process \(\vec{\Pi}\) is clearly adapted to \(\mathbb{F}\), and each component gives the conditional probability that the current state of \(M\) is \(\{i\}\) given the information generated by \(X\) until the current time \(t\). Moreover, using standard arguments as in Shiryaev [1978, pp. 166-167], and Dayanik et al. [2008a, Proof of Proposition 2.1], it can be shown that the problem in (1.1) is equivalent to a fully observed optimal stopping problem with the process \(\vec{\Pi}\) as the new hyperstate. More precisely, the value function \(U\) in (1.1) can be written as

\[
U(T, \vec{\pi}) = V(T, \vec{\pi}) \triangleq \sup_{\tau \leq T} \mathbb{E}^{\vec{\pi}}\left[ \int_0^\tau e^{-\rho t} C(\vec{\Pi}_t) dt + e^{-\rho T} H(\vec{\Pi}_T) \right],
\]

in terms of the functions

\[
C(\vec{\pi}) \triangleq \sum_{i \in E} c_i \pi_i \quad \text{and} \quad H(\vec{\pi}) \triangleq \max_{k \in A} H_k(\vec{\pi}), \quad \text{where} \quad H_k(\vec{\pi}) \triangleq \sum_{i \in E} \mu_{k,i} \pi_i.
\]

If there is a stopping time \(\tau^*\) attaining the supremum in (2.3), then the admissible strategy \((\tau^*, d(\tau^*))\) is an optimal rule for the problem in (1.1) if we define

\[
d(\tau) \in \arg\max_{k \in A} H_k(\vec{\Pi}_\tau).
\]
2.3. Sample paths of \( \bar{\Pi} \). Let us take a sample path of the observations process \( X \), in which \( m \)-many arrivals are observed on \([0, t]\). Let \((t_k)_{k \leq m}\) denote those arrival times. If we know that the process \( M \) stays at the state \( \{i\} \) without any transition, then the (conditional) likelihood of this path would be written as 

\[
\mathbb{P}^\pi \{ \sigma_k \in dt_k, Y_k \in dy_k ; k \leq m \mid M_s = i, s \leq t \} = \left[ \lambda_i e^{-\lambda_i t_1} dt_1 \right] \cdots \left[ \lambda_i e^{-\lambda_i (t_m-t_{m-1})} dt_{m-1} \right] \prod_{k=1}^{m} [f_i(y_k) \nu(dy_k)] = e^{-\lambda_i t} \prod_{k=1}^{m} \lambda_i dt_k \cdot f_i(y_k) \nu(dy_k).
\]

By construction, the observation process \( L \) also, let 

\[
L(t) \triangleq \int_0^t \sum_{i=1}^n \lambda_i 1_{\{M_s=i\}} ds \quad \text{and} \quad \ell(t, y) \triangleq \sum_{j \in E} 1_{\{M_i=j\}} \lambda_j \cdot f_j(y).
\]

Also, let \( L^\pi(t, m : (t_k, y_k), k \leq m) \triangleq \sum_{j \in E} L^\pi_j(t, m : (t_k, y_k), k \leq m) \). Then we have

\[
\Pi^\pi_i(t) = \left. \frac{L^\pi_i(t, N_i : (\sigma_k, Y_k), k \leq N_i)}{L^\pi(t, N_i : (\sigma_k, Y_k), k \leq N_i)} \right|_{m=N_i \wedge (t_k=\sigma_k \wedge y_k=Y_k) \leq m},
\]

\( \mathbb{P}^\pi \)-a.s., for all \( t \geq 0 \), and for \( i \in E \).

Lemma 2.1 indicates that the conditional probability of \( M_t \) being in state \( i \) is simply the relative likelihood of the observed path until \( t \) on the event \( \{M_t = i\} \). Using the explicit form in (2.9), we describe the behavior of the sample paths of \( \bar{\Pi} \) in Remark 2.1 below.

Remark 2.1. The process \( \bar{\Pi} \) has piecewise-deterministic sample paths: between two arrival times of \( X \), it moves deterministically, and at an arrival time, it jumps from one point to another depending on the observed mark size (see Figure 1). In precise terms, the sample paths have the characterization}

\[
\bar{\Pi}(t) = \bar{x}(t - \sigma_m, \bar{\Pi}(\sigma_m)), \quad \sigma_m \leq t < \sigma_{m+1}, \quad m \in \mathbb{N}
\]

\[
\bar{\Pi}(\sigma_m) = R(\bar{\Pi}(\sigma_m -), Y_m)
\]

where \( \bar{x}(t, \vec{\sigma}) \equiv (x_1(t, \vec{\sigma}), \ldots, x_n(t, \vec{\sigma})) \) is defined as

\[
x_i(t, \vec{\sigma}) \equiv \frac{\mathbb{P}^\pi \{1_{\{M_t=i\}} \cdot e^{-I(t)} \}}{\mathbb{E}^\pi [e^{-I(t)}]} = \frac{\mathbb{P}^\pi \{\sigma_1 > t, M_t = i\}}{\mathbb{P}^\pi \{\sigma_1 > t\}}, \quad \text{for } i \in E,
\]
Figure 1.
Sample paths of the process $\vec{\Pi}$ for four different examples. Solid lines represent actual sample paths. Dashed lines in panels (c) and (d) are the deterministic parts in (2.11). In panels (a) and (b), there are two hidden states, and in panels (c) and (d), there are three. The parameters of each example:

\[ Q_a = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad Q_b = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, \quad Q_c = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \quad Q_d = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \]

with $\vec{\lambda}_a = [1, 2]$, $\vec{\lambda}_b = [1, 4]$, $\vec{\lambda}_c = [1, 2, 3]$, $\vec{\lambda}_d = [1, 3, 5]$. In each example, jumps of the process $X$ are always of unit size.

and $R(\vec{\pi}, y)$ is defined by

\[ R(\vec{\pi}, y) \equiv R(\pi_1, \ldots, \pi_n, y) = \left( \frac{\lambda_1 \pi_1 f_1(y)}{\sum_{j \in E} \lambda_j \pi_j f_j(y)}, \ldots, \frac{\lambda_n \pi_n f_n(y)}{\sum_{j \in E} \lambda_j \pi_j f_j(y)} \right). \]

Note that the paths $t \mapsto \vec{x}(t, \vec{\pi})$ have the semigroup property $\vec{x}(t + u, \vec{\pi}) = \vec{x}(u, \vec{x}(t, \vec{\pi}))$, for $t, u \geq 0$.

The $i$'th component of the vector flow $x_i(\cdot, \cdot)$ indicates how likely it is to have a period of $[0, t]$ without any arrival on the event $\{M_t = i\}$. Moreover, from similar analysis in Dayanik et al. [2008a, Section 2], $\vec{\Pi}$ is a $(\mathbb{P}^{\vec{\pi}}, \mathbb{F})$-Markov process for every $\vec{\pi} \in D$.

**Corollary 2.1.** Using infinitesimal last step analysis, it can be shown (see, for example, Darroch and Morris [1968, page 416], and Karlin and Taylor [1998, Chapter 6.7]) that the vector

\[ \vec{m}(t, \vec{\pi}) \equiv (m_1(t, \vec{\pi}), \ldots, m_n(t, \vec{\pi})) \triangleq \left( \mathbb{E}^{\vec{\pi}} \left[ 1_{\{M_t = 1\}} \cdot e^{-I(u)} \right], \ldots, \mathbb{E}^{\vec{\pi}} \left[ 1_{\{M_t = n\}} \cdot e^{-I(u)} \right] \right) \]
has the form \( \hat{m}(t, \bar{\pi}) = \bar{\pi} \cdot e^{(Q - \Lambda)t} \) in terms of \( n \times n \) diagonal matrix \( \Lambda \) with \( \Lambda_{i,i} = \lambda_i \), and the components of \( \hat{m}(t, \bar{\pi}) \) satisfy \( dt_i \hat{m}_i(t, \bar{\pi})/dt = -\lambda_i m_i(t, \bar{\pi}) + \sum_{j \in E} m_j(t, \bar{\pi}) \cdot q_{j,i} \). Then thanks to the chain rule and (2.11) we have

\[
(2.14) \quad \frac{dx_i(t, \bar{\pi})}{dt} = \left( \sum_{j} q_{j,i}x_j(t, \bar{\pi}) - \lambda_i x_i(t, \bar{\pi}) + x_i(t, \bar{\pi}) \sum_{j} \lambda_j x_j(t, \bar{\pi}) \right).
\]

Hence, the process \( \Pi \) in (2.10) has the dynamics

\[
(2.15) \quad d\Pi^{(i)}_t = \left( \sum_{j} q_{j,i}\Pi^{(j)}_t - \lambda_i \Pi^{(i)}_t + \Pi^{(i)}_t \sum_{j} \lambda_j \Pi^{(j)}_t \right) dt + \int_{\mathbb{R}^d} \left[ \frac{\lambda_i f(y)\Pi^{(i)}_t}{\sum_{j \in E} \lambda_j f_j(y)\Pi^{(j)}_t} - 1 \right] p(dt, dy), \quad i \in E,
\]

where \( p(\cdot, \cdot) \) is the point process given by

\[
p((0, t] \times B) = \sum_{i \in \mathbb{N}} 1_{(0, t] \times B}(\sigma_i, Y_i), \quad \text{for every Borel set } B \in \mathcal{B}(\mathbb{R}^d) \text{ and } t \geq 0.
\]

### 3. Constructing the Value Function

The characterization of the sample paths in (2.15) and general theory of optimal stopping (see, for example, Bensoussan [1992], Lenhart and Liao [1985]) imply that the free-boundary problem associated with the optimal stopping problem in (2.3) has the form

\[
(3.1) \quad \max \left\{ (-\rho + \mathcal{L})f(s, \bar{\pi}) + C(\bar{\pi}) ; H(\bar{\pi}) - f(s, \bar{\pi}) \right\} = 0,
\]

in terms of the infinitesimal generator

\[
\mathcal{L} f(s, \bar{\pi}) = \frac{\partial f(s, \bar{\pi})}{\partial s} + \sum_{i \in E} \left( \sum_{j \in E} q_{j,i} \bar{\pi}_j - \lambda_i \pi_i + \pi_i \lambda_j \bar{\pi}_j \right) \frac{\partial f(s, \bar{\pi})}{\partial \pi_i} + \int_{y \in \mathbb{R}^d} [f(s, R(\bar{\pi}, y)) - f(s, \bar{\pi})] \sum_{i \in E} \pi_i \lambda_i \nu_i(dy),
\]

acting on (smooth) functions \( f(\cdot, \cdot) \) on \([0, T] \times D\). Studying the equation \( (-\rho + \mathcal{L})f(s, \bar{\pi}) + C(\bar{\pi}) = 0 \) and determining the stopping regions is not easy even when \( n = 2 \); see, for example, Peskir and Shiryaev [2000], which solves a free-boundary problem similar to (3.1) for an infinite horizon problem with \( n = 2 \). Moreover, it is known that the value function of such a stopping problem may not be differentiable at every point on its domain as illustrated in Dayanik and Sezer [2005], in which case the equation (3.1) should be considered in viscosity sense.

Instead of studying the problem in (3.1), we will employ a sequential approximation technique to compute the value function following Gugerli [1986] and Davis [1993, Chapter 5]. Similar approach is also taken in Bayraktar et al. [2006] and Dayanik et al. [2008a] for the disorder-detection and hypothesis-testing problems respectively on infinite horizon. Below, we tailor this method to fit it into the finite-horizon setting. We focus on the non-trivial modifications that arise due to time-dependent operators and the more general form of \( M \), and otherwise refer to the results of Dayanik et al. [2008a]. All the proofs are delegated to the Appendix.
3.1. A sequential approximation. Let us first define the sequence of functions

\[
V(s, \bar{\pi}) \triangleq \sup_{\tau \leq s} \mathbb{E}^D \left[ \int_0^\tau e^{-\rho t} C(\bar{\Pi}_t) dt + e^{-\rho \tau} H \left( \bar{\Pi}_\tau \right) \right], \quad \text{and}
\]

\[
V_m(s, \bar{\pi}) \triangleq \sup_{\tau \leq s} \mathbb{E}^D \left[ \int_0^{\tau \wedge \sigma_m} e^{-\rho t} C(\bar{\Pi}_t) dt + e^{-\rho \tau \wedge \sigma_m} H \left( \bar{\Pi}_{t \wedge \sigma_m} \right) \right],
\]

for \( m \in \mathbb{N} \), on the domain \([0, T] \times D\), where the first argument ‘s’ should be considered as the remaining time to maturity.

Proposition 3.1 below shows that \( V_m \)’s converge to \( V \) uniformly; see also the proof of Davis [1993, Theorem (53.40)] and Dayanik et al. [2008a, Proposition 3.1] for related results. Proposition 3.1 is a generalization of these results in the finite horizon case.

**Proposition 3.1.** The sequence \( \{V_m\}_{m \geq 1} \) converges to \( V \) uniformly on \([0, T] \times D\). More precisely, we have

\[
V_m(s, \bar{\pi}) \leq V(s, \bar{\pi}) \leq V_m(s, \bar{\pi}) + (T\|C\| + 2\|H\|) \left( \frac{\sqrt{T}}{m - 1} \right)^{1/2} \left( \frac{\lambda}{2\rho + \lambda} \right)^{m/2},
\]

for all \((s, \bar{\pi}) \in [0, T] \times D\) and \( m \in \mathbb{N} \), where \( \|C\| \triangleq \max_{\bar{\pi} \in D} |C(\bar{\pi})| \), \( \|H\| \triangleq \max_{\bar{\pi} \in D} |H(\bar{\pi})| \) and \( \lambda \triangleq \max_{i \in E} \lambda_i \).

Let us consider the second problem in (3.2) for fixed \( m \in \mathbb{N} \), and let \( \tau \leq s \) be an \( \mathcal{F} \)-stopping time. Note that the first arrival time \( \sigma_1 \) is a regeneration time of Markov process \( \bar{\Pi} \); therefore, on the event \( \{\tau \geq \sigma_1\} \), the maximal expected reward that the DM can achieve after \( \sigma_1 \) should be \( V_{m-1}(s - \sigma_1, \bar{\Pi}_{\sigma_1}) \). Define the operator

\[
Jw(\tau, s, \bar{\pi}) \triangleq \mathbb{E}^D \left[ \int_0^{\tau \wedge \sigma_1} e^{-\rho t} C(\bar{\Pi}_t) dt + 1_{\{\tau < \sigma_1\}} e^{-\rho \tau} H \left( \bar{\Pi}_\tau \right) + 1_{\{\sigma_1 \leq \tau\}} e^{-\rho \sigma_1} w \left( s - \sigma_1, \bar{\Pi}(\sigma_1) \right) \right].
\]

Then, the dynamic programming intuition suggests that \( V(\cdot) \) should solve the equation \( V_m(s, \bar{\pi}) = J_0 V_{m-1}(s, \bar{\pi}) \), where the operator \( J_0 \) is defined as

\[
J_0 w(s, \bar{\pi}) \triangleq \sup_{\tau \leq s} Jw(\tau, s, \bar{\pi}) = \sup_{t \in [0, s]} Jw(t, s, \bar{\pi})
\]

for a bounded function \( w : [0, T] \times D \mapsto \mathbb{R} \). The second equality in (3.5) is due to the characterization of \( \mathcal{F} \)-stopping times (Davis [1993, Lemma A2.3, p. 261]) whereby for every \( m \in \mathbb{N} \), there exists a \( \mathcal{F}_{\sigma_m} \)-measurable \( R_m \) such that \( \tau \wedge \sigma_{m+1} = (\sigma_m + R_m) \wedge \sigma_{m+1} \), \( \mathbb{P} \)-a.s. on \( \{\tau \geq \sigma_m\} \).

Note that, with the notation in (2.13), we have \( \mathbb{P}^\mathcal{F} [\sigma_1 > u] = \mathbb{E}^\mathcal{F} [e^{-I(u)}] \) and

\[
\mathbb{P}^\mathcal{F} [\sigma_1 \in du, M_u = i] = \mathbb{E}^\mathcal{F} \left[ \lambda_i 1_{\{M_u = i\}} e^{-I(u)} \right] du = \lambda_i m_i(u, \bar{\pi}) du,
\]

and using the characterization of the paths in (2.10) and (2.14) the operator \( J \) in (3.4) can be rewritten as

\[
Jw(t, s, \bar{\pi}) = \left( \sum_{i \in E} m_i(t, \bar{\pi}) \right) \cdot e^{-\rho t} \cdot H \left( \bar{x}(t, \bar{\pi}) \right)
\]

\[+ \int_0^t e^{-\rho u} \sum_{i \in E} m_i(u, \bar{\pi}) \cdot \left( C(\bar{x}(u, \bar{\pi})) + \lambda_i \cdot S_i w(s - u, \bar{x}(u, \bar{\pi})) \right) du,
\]
in terms of the operators (see (2.12))

\[ S_i w(t, \vec{\pi}) \triangleq \int_{\mathbb{R}^d} w(t, \mathcal{R}(\vec{\pi}, y)) f_i(y) \nu(dy), \quad \text{for } i \in E. \]

The following lemma provides basic properties of the operator \( J_0 \).

**Lemma 3.1.** If \( w(\cdot, \cdot) \) is a bounded continuous function on \([0, T] \times D\), then so is \( J_0 w(\cdot, \cdot) \). Also, if \( w_1(\cdot, \cdot) \leq w_2(\cdot, \cdot) \), then \( J_0 w_1(\cdot, \cdot) \leq J_0 w_2(\cdot, \cdot) \). Moreover, if the mapping \( \vec{\pi} \mapsto w(s, \vec{\pi}) \) is convex for each \( s \in [0, T] \), so is \( \vec{\pi} \mapsto J_0 w(s, \vec{\pi}) \) for each \( s \in [0, T] \).

Let us now define the sequence

\[ v_0(s, \vec{\pi}) \triangleq H(\vec{\pi}), \quad \text{and} \quad v_{m+1}(s, \vec{\pi}) \triangleq J_0 v_m(s, \vec{\pi}), \quad \text{for } m \geq 0, \quad \text{on } [0, T] \times D. \]

Thanks to Lemma 3.1 we immediately see that the sequence \( \{v_m(\cdot, \cdot)\}_{m \in \mathbb{N}} \) is non-decreasing, hence the pointwise limit \( v(\cdot, \cdot) \triangleq \sup_{m \in \mathbb{N}} v_m(\cdot, \cdot) \) is well defined on \([0, T] \times D\). Moreover, again by Lemma 3.1, we have that each \( v_m(\cdot, \cdot) \) is bounded and continuous on \([0, T] \times D\), and the mapping \( \vec{\pi} \mapsto v_m(s, \vec{\pi}) \) is convex for each \( s \in [0, T] \).

**Proposition 3.2.** The sequences defined in (3.2) and (3.8) coincide. That is, we have \( v_m(\cdot, \cdot) = V_m(\cdot, \cdot) \) for every \( m \in \mathbb{N} \).

**Corollary 3.1.** Each \( V_m \) is continuous and hence their uniform limit (see Proposition 3.1) \( V(\cdot, \cdot) \) is also continuous on \([0, T] \times D\). As the upper envelope of convex mappings \( \vec{\pi} \mapsto v_m(s, \vec{\pi}) = V_m(s, \vec{\pi}) \), the mapping \( \vec{\pi} \mapsto V(s, \vec{\pi}) \) is again convex for each \( s \in [0, T] \).

The Proposition 3.3 below is the dynamic programming equation for \( V(\cdot, \cdot) \), characterizing the value function as the fixed point of the operator \( J_0 \) defined in (3.5-3.6).

**Proposition 3.3.** The value function satisfies \( V(s, \vec{\pi}) = J_0 V(s, \vec{\pi}) \), and it is the smallest bounded solution of this equation greater than \( H(\cdot) \).

4. **An Optimal Strategy**

Recall that the process \( \bar{\Pi} \) has right-continuous paths (with left limits), and the functions \( V(\cdot, \cdot) \) and \( H(\cdot) \) are continuous due to Corollary 3.1. Hence the paths of the process \( V(t, \bar{\Pi}_t) - H(\bar{\Pi}_t) \) are also right-continuous and have left limits. Therefore, for \( \varepsilon \geq 0 \) the random time

\[ U_{\varepsilon}(s, \vec{\pi}) \triangleq \inf \left\{ t \in [0, s] : V(s - t, \bar{\Pi}_t) - \varepsilon \leq H(\bar{\Pi}_t) \right\} \]

is a well-defined \( \mathbb{F} \)-stopping time. We also have \( U_{\varepsilon}(s, \vec{\pi}) \wedge \sigma_1 = r_{\varepsilon}(s, \vec{\pi}) \wedge \sigma_1 \), where

\[ r_{\varepsilon}(s, \vec{\pi}) \triangleq \inf \left\{ t \in [0, s] : V(s - t, \bar{x}(t, \vec{\pi})) - \varepsilon \leq H(\bar{x}(t, \vec{\pi})) \right\}, \]

which can be considered as the deterministic counterpart of (4.1).

**Proposition 4.1.** The stopping time \( U_{\varepsilon}(s, \vec{\pi}) \) defined in (4.1) is an \( \varepsilon \)-optimal stopping time for the problem in (2.3), i.e.,

\[ \mathbb{E}^\mathbb{F} \left[ \int_0^{U_{\varepsilon}(s, \vec{\pi})} e^{-\rho t} C(\bar{\Pi}_t) \, dt + e^{-\rho U_{\varepsilon}(s, \vec{\pi})} H(\bar{\Pi}(U_{\varepsilon}(s, \vec{\pi}))) \right] \geq V(s, \vec{\pi}) - \varepsilon, \]

for all \( \varepsilon \geq 0 \) and \((s, \vec{\pi}) \in [0, T] \times D\).

Before proceeding with the proof of Proposition 4.1, we first state an immediate consequence of this result.
Corollary 4.1. The pair \((U_0(T, \bar{\pi}), d(U_0(T, \bar{\pi})))\) is an optimal admissible strategy for the problem in (1.1).

Proof of Proposition 4.1. Let us define

\[
Z_t \triangleq \int_0^t e^{-\rho u} C(\bar{\Pi}_u) \, du + e^{-\rho t} V(s - t, \bar{\Pi}_t), \quad t \in [0, s],
\]

which is a bounded process on \(t \in [0, s] \subseteq [0, T]\). The \(\varepsilon\)-optimality of \(U_\varepsilon(s, \bar{\pi})\) follows easily once we establish

\[
\mathbb{E}^{\bar{\pi}}[Z_{U_\varepsilon(s, \bar{\pi})}] = Z_0
\]

since this equality would imply \(V(s, \bar{\pi}) = \mathbb{E}^{\bar{\pi}}[Z_{U_\varepsilon(s, \bar{\pi})}]\) due to regularity of the paths \(t \mapsto V(t, \bar{\Pi}_t) - H(\bar{\Pi}_t)\). In the remainder of the proof we show (4.5) by establishing \(\mathbb{E}^{\bar{\pi}}[Z_{U_\varepsilon(s, \bar{\pi})}] \leq Z_0\) for \(m = 1, 2, \ldots,\) inductively. After taking the limit as \(m \to \infty\) in the equality above, we obtain (4.5) due to bounded convergence theorem.

For typographical convenience we write \(r_\varepsilon = r_\varepsilon(s, \bar{\pi})\) and \(U_\varepsilon = U_\varepsilon(s, \bar{\pi})\). First, we consider \(m = 1\). Recall that \(U_\varepsilon(s, \bar{\pi}) \land \sigma_1 = r_\varepsilon \land \sigma_1\). Then

\[
\mathbb{E}^{\bar{\pi}}[Z_{U_\varepsilon \land \sigma_1}] = \mathbb{E}^{\bar{\pi}}[Z_{r_\varepsilon \land \sigma_1}] =
\]

\[
\mathbb{E}^{\bar{\pi}} \left[ \int_0^{r_\varepsilon \land \sigma_1} e^{-\rho t} C(\bar{\Pi}_t) \, dt + e^{-\rho r_\varepsilon} H(\bar{\Pi}_{r_\varepsilon}) + e^{-\rho \sigma_1} V(s - \sigma_1, \bar{\Pi}_{\sigma_1}) + \mathbb{E}^{\bar{\pi}} \{ \sigma_1 > r_\varepsilon \} e^{-\rho r_\varepsilon} \left( V(s - r_\varepsilon, \bar{\Pi}_{r_\varepsilon}) - H(\bar{\Pi}_{r_\varepsilon}) \right) \right]
\]

\[
= JV(r_\varepsilon, s, \bar{\pi}) + e^{-\rho r_\varepsilon} \cdot \mathbb{P}^{\bar{\pi}} \{ \sigma_1 > r_\varepsilon \} \left( V(s - r_\varepsilon, \bar{x}(r_\varepsilon, \bar{\pi})) - H(\bar{x}(r_\varepsilon, \bar{\pi})) \right)
\]

where we used Proposition 3.3. Analogously to Dayanik et al. [2008a, Lemma 3.8], we have that for deterministic times \(u \leq t \leq s\), and for a bounded function \(w(\cdot, \cdot)\)

\[
Jw(t, s, \bar{\pi}) = Jw(u, s, \bar{\pi}) + \mathbb{P}^{\bar{\pi}} \{ \sigma_1 > u \} \cdot e^{-\rho u} \cdot \left( Jw(t - u, s - u, x(u, \bar{\pi})) - H(x(u, \bar{\pi})) \right).
\]

For \(t < r_\varepsilon(s, \bar{\pi})\), we have \(V(s - t, \bar{x}(t, \bar{\pi})) - H(\bar{x}(t, \bar{\pi})) > \varepsilon\). Then, (4.8) yields \(JV(t, s, \bar{\pi}) \leq \sup_{u \in [t, s]} JV(u, s, \bar{\pi}) - \varepsilon \mathbb{P}^{\bar{\pi}} \{ \sigma_1 > t \} e^{-\rho t} \). Combining (4.8) with (4.7), we get

\[
\mathbb{E}^{\bar{\pi}}[Z_{U_\varepsilon \land \sigma_1}] = \sup_{u \in [r_\varepsilon, s]} JV(u, s, \bar{\pi}) = J_0 V(s, \bar{\pi}) = V(s, \bar{\pi}) = Z_0.
\]
Now suppose by induction that $E^\pi[Z_{U_\varepsilon(s,\bar{\pi})}\wedge \sigma_m] = Z_0$ for $m \geq 1$ and consider the equality
\begin{equation}
E^\pi[Z_{U_\varepsilon\wedge \sigma_{m+1}}] = E^\pi\left[1\{U_\varepsilon < \sigma_1\} Z_{U_\varepsilon} + 1\{U_\varepsilon \geq \sigma_1\} Z_{U_\varepsilon\wedge \sigma_{m+1}}\right]
\end{equation}
\begin{equation}
= E^\pi\left[1\{U_\varepsilon < \sigma_1\} \left(\int_0^{U_\varepsilon} e^{-\rho t} C(\bar{\Pi}_t) \, dt + e^{-\rho U_\varepsilon} V(s - U_\varepsilon, \bar{\Pi}_{U_\varepsilon})\right) + 1\{U_\varepsilon \geq \sigma_1\} \left(\int_0^{U_\varepsilon\wedge \sigma_{m+1}} e^{-\rho t} C(\bar{\Pi}_t) \, dt + e^{-\rho U_\varepsilon\wedge \sigma_{m+1}} V\left(s - U_\varepsilon \wedge \sigma_{m+1}, \bar{\Pi}_{U_\varepsilon\wedge \sigma_{m+1}}\right)\right)\right].
\end{equation}

On the event $\{U_\varepsilon \geq \sigma_1\}$, we have $U_\varepsilon \wedge \sigma_{m+1} = \sigma_1 + [U_\varepsilon \wedge \sigma_m] \circ \theta_{\sigma_1}$, where $\theta$ is the time-shift operator on $\Omega$; i.e., $X_t \circ \theta_s = X_{t+s}$. Using the strong Markov property of $\bar{\Pi}$, equation (4.9) becomes
\begin{equation}
E^\pi[Z_{U_\varepsilon\wedge \sigma_{m+1}}] = E^\pi\left[1\{U_\varepsilon < \sigma_1\} \left(\int_0^{U_\varepsilon} e^{-\rho t} C(\bar{\Pi}_t) \, dt + e^{-\rho U_\varepsilon} V(s - U_\varepsilon, \bar{\Pi}_{U_\varepsilon})\right) + \int_0^{\sigma_1} e^{-\rho t} C(\bar{\Pi}_t) \, dt + 1\{U_\varepsilon \geq \sigma_1\} e^{-\rho \sigma_1} \eta(s - \sigma_1, \bar{\Pi}_{\sigma_1})\right],
\end{equation}
where
\begin{equation}
\eta(u, \bar{\pi}) \triangleq E^\pi\left[\int_0^{U_\varepsilon \wedge \sigma_m} e^{-\rho t} C(\bar{\Pi}_t) \, dt + e^{-\rho U_\varepsilon \wedge \sigma_m} V(u - U_\varepsilon \wedge \sigma_m, \bar{\Pi}_{U_\varepsilon \wedge \sigma_m})\right] = V(u, \bar{\pi}),
\end{equation}
thanks to the induction hypothesis for $m$. Combining (4.10) and (4.11) and the definition of $Z$ in (4.4) we get $E^\pi[Z_{U_\varepsilon\wedge \sigma_{m+1}}] = E^\pi\left[1\{U_\varepsilon < \sigma_1\} Z_{U_\varepsilon} + 1\{U_\varepsilon \geq \sigma_1\} Z_{\sigma_1}\right] = E^\pi[Z_{U_\varepsilon\wedge \sigma_1}] = Z_0$, where the last equality follows from our result for $m = 1$. Hence we have $E^\pi[Z_{U_\varepsilon\wedge \sigma_{m+1}}] = Z_0$ and this completes the induction step. \qed

4.1. A nearly-optimal strategy. On a practical level, one cannot compute $V$ directly, but instead computes the approximate value functions $V_m$’s defined in (3.2) and employs the corresponding nearly-optimal strategies (see 4.12). It is therefore important to know the error associated with this approximation.

For a given error level $\varepsilon > 0$, let us fix
\begin{equation}
m = \inf\left\{k \in \mathbb{N} : \left(T \|C\| + 2 \|H\|\right) \left(\frac{\bar{X} T}{k - 1}\right)^{1/2} \left(\frac{\bar{X}}{2\rho + \bar{X}}\right)^{k/2} \leq \varepsilon/2\right\},
\end{equation}
such that $\|V_m - V\| \leq \varepsilon/2$ on $[0, T] \times D$ via (3.3). Next, let us define the stopping times
\begin{equation}
U^{(m)}_{\varepsilon/2}(s, \bar{\pi}) \triangleq \inf\{t \in [0, s] : V_m(s, \bar{\Pi}_t) - \varepsilon/2 \leq H(\bar{\Pi}_t)\}.
\end{equation}
The regularity of the paths $t \mapsto \bar{\Pi}_t$ implies that $V\left(U^{(m)}_{\varepsilon/2}(s, \bar{\pi}), \bar{\Pi}_{U^{(m)}_{\varepsilon/2}(s, \bar{\pi})}\right) - H\left(\bar{\Pi}_{U^{(m)}_{\varepsilon/2}(s, \bar{\pi})}\right) \leq \varepsilon$. Then the arguments in the proof of Proposition 4.1 (see (4.4), (4.5), and (4.6)) can easily be modified to show that
\begin{equation}
V(s, \bar{\pi}) = E^\pi\left[\int_0^{U^{(m)}_{\varepsilon/2}(s, \bar{\pi})} e^{-\rho t} C(\bar{\Pi}_t) \, dt + e^{-\rho U^{(m)}_{\varepsilon/2}(s, \bar{\pi})} V\left(s - U^{(m)}_{\varepsilon/2}(s, \bar{\pi}), \bar{\Pi}_{U^{(m)}_{\varepsilon/2}(s, \bar{\pi})}\right)\right]
\end{equation}
\begin{equation}
\leq E^\pi\left[\int_0^{U^{(m)}_{\varepsilon/2}(s, \bar{\pi})} e^{-\rho t} C(\bar{\Pi}_t) \, dt + e^{-\rho U^{(m)}_{\varepsilon/2}(s, \bar{\pi})} H\left(\bar{\Pi}_{U^{(m)}_{\varepsilon/2}(s, \bar{\pi})}\right)\right] + \varepsilon.
\end{equation}
Hence, if we apply the admissible strategy $\left(U^{(m)}_{\varepsilon}(T, \bar{\pi}), d(U^{(m)}_{\varepsilon}(T, \bar{\pi}))\right)$, which requires computing (3.2) only up to $m$ defined above, the resulting error is no more than $\varepsilon$. 
4.2. Stopping and continuation regions. Let
\[
C_T \triangleq \{(s, \bar{\pi}) \in [0, T] \times D : V(s, \bar{\pi}) > H(\bar{\pi})\},
\]
\[
\Gamma_T \triangleq \{(s, \bar{\pi}) \in [0, T] \times D : V(s, \bar{\pi}) = H(\bar{\pi})\}
\]
denote the continuation and stopping regions respectively. The stopping region can further be decomposed as the union $\bigcup_{k \in A} \Gamma_{T,k}$ of the regions
\[
\Gamma_{T,k} \triangleq \{(s, \bar{\pi}) \in [0, T] \times D : V(s, \bar{\pi}) = H_k(\bar{\pi})\}, \quad k \in A,
\]
where $H_k$ is defined in (2.4). Corollary 4.1 states that in the optimal solution $(U_0(T, \bar{\pi}), d(U_0(T, \bar{\pi})))$, one observes the process $\bar{\Pi}$ until $U_0(T, \bar{\pi})$, whence it enters the region $\Gamma_T$. At this time, if $\bar{\Pi}$ is in the set $\Gamma_{T,k}$ we take $d(U_0(T, \bar{\pi})) = k$.

**Remark 4.1.** The definition of the value function $V$ in (2.3) implies that the mapping $s \mapsto V(s, \bar{\pi})$ is non-decreasing. Therefore if $(s, \bar{\pi}) \in \Gamma_{T,k}$ for some $(s, \bar{\pi}) \in [0, T] \times D$, then we have $(t, \bar{\pi}) \in \Gamma_{T,k}$ for all $t \leq s$.

**Remark 4.2.** For fixed $s \leq T$, let $(s, \bar{\pi}_1)$ and $(s, \bar{\pi}_2)$ be two points in the region $\Gamma_{T,k}$, and let $\alpha \in (0, 1)$. As the upper envelope of convex mappings $\bar{\pi} \mapsto v_m(s, \bar{\pi})$ (see Corollary 3.1), the mapping $\bar{\pi} \mapsto V(s, \bar{\pi})$ is convex for each $s \in [0, T]$. Using this property we obtain
\[
H_k(\alpha \cdot \bar{\pi}_1 + (1 - \alpha) \cdot \bar{\pi}_2)) \leq V(s, \alpha \cdot \bar{\pi}_1 + (1 - \alpha) \cdot \bar{\pi}_2) \leq \alpha \cdot V(s, \bar{\pi}_1) + (1 - \alpha) \cdot V(s, \bar{\pi}_2)
\]
\[
\quad = \alpha \cdot H_k(\bar{\pi}_1) + (1 - \alpha) \cdot H_k(\bar{\pi}_2) = H_k(\alpha \cdot \bar{\pi}_1 + (1 - \alpha) \cdot \bar{\pi}_2),
\]
which implies that $(s, \alpha \cdot \bar{\pi}_1 + (1 - \alpha) \cdot \bar{\pi}_2) \in \Gamma_{T,k}$, and the region $\Gamma_{T,k} \cap \{(s) \times D\}$ is convex for each fixed $s \leq T$ and $k \in A$.

**Remark 4.3.** Note that $\Gamma_T \supseteq \{(0, \bar{\pi}); \bar{\pi} \in D\} \neq \emptyset$. The region $\{(s, \bar{\pi}) \in \Gamma_T : s > 0\}$ may however be empty. In an example where $\min_{i \in E} c_i > 0$ and $\mu_{k,i}$'s are all the same it is never optimal to stop prior to terminal time $T$. Moreover, the region $\{(s, \bar{\pi}) \in \Gamma_T : s > 0\}$ may be non-empty but have an empty interior. For example, in the hypothesis testing problem discussed in Section 5.2 all the states of the unobservable Markov process are absorbing, and each component $\Pi_t^{(i)}$ is a martingale. Since the terminal cost function of the corresponding minimization problem (see (2.4)) $H(\cdot) = \min_{k \in E} H_k(\cdot)$ is concave, the process $H(\Pi_t)$ is a super-martingale on $[0, T]$. If we select $\rho = 0$ and $c_i = 0$ for all $i \in E$, it is therefore never optimal to stop early on the interior of $\{(s, \bar{\pi}) \in \Gamma_T : s > 0\}$. In this case, there is no penalty associated with a delay in the decision, so $\tau = T$ unless $\bar{\pi}$ is at a corner of the simplex $D$.

**Lemma 4.1.** For $i \in E$, let $A^*(i) \triangleq \{k \in A : \mu_{k,i} = \max_{j \in A} \mu_{j,i}\}$. If the inequality $c_i - \rho \mu_{k,i} + \sum_{j \neq i} (\mu_{k,j} - \mu_{k,i})q_{i,j} > 0$ holds for all $k \in A^*(i)$, then there exists $\pi^*_i < 1$, independent of $T$, such that $\{(s, \bar{\pi}) \in (0, T] \times D : \pi_i \geq \pi^*_i\} \subseteq C_T$.

If the hidden process $M$ is known to be in state $i \in E$, then the expression $-\rho \mu_{k,i}$ is the instantaneous decay of the payoff from selecting action $k \in A$ immediately, and $c_i$ is the instantaneous cost of waiting. Moreover, under action $k \in A$, the term $\sum_{j \neq i}^{} (\mu_{k,j} - \mu_{k,i})q_{i,j}$ is the marginal rate of return from waiting for the hidden process $M$ to jump to another state. Therefore the sum in Lemma 4.1 is the instantaneous net return enjoyed by the DM under action $k \in A$. Lemma 4.1 indicates that if there is strong posteriori evidence that $M$ is in state $i$, and if the instantaneous net return is positive under all favorable actions around the $i$'th corner of $D$, the decision maker should not stop (unless $T = 0$).
4.3. Stopping regions for reward maximization with running cost. Here, we consider the problem in (2.3) with the assumption \( c_i \leq 0 \) (running costs) for \( i \in E \), and \( \bar{\pi} \triangleq \max_{k,i} \mu_{k,i} > 0 \) (terminal rewards). The second condition is not restrictive if \( \rho = 0 \) since we can always add (and subtract) the same constant to (and from) the terminal reward function. Let us define
\[
I^* \triangleq \{ i \in E : \max_{k \in A} \mu_{k,i} = \bar{\pi} \},
\]
which is the set of the states of \( M \), at which the DM can get the highest terminal reward. Since \( c_i \leq 0 \) for all \( i \in E \), we obviously have \( \bigcup_{i \in I^*} \{ (s, \bar{\pi}) : s \in [0, T], \pi_i = 1 \} \subset \Gamma_T \).

In general, if there is a penalty associated with waiting, we expect that it is optimal to stop at the points \( (s, \bar{\pi}) \) for which the “best” component \( \pi_i, i \in I^* \), is sufficiently high, for any \( s > 0 \). Lemma 4.2 provides a sufficient condition for this to be true.

Lemma 4.2. Let \( i \in I^* \). If \( \rho > 0 \), or \( c_i < 0 \), then there exists a constant \( \pi_i^* < 1 \), independent of \( T \), such that
\[
\Gamma_T \supseteq \{ (s, \bar{\pi}) \in [0, T] \times D : \pi_i \geq \pi_i^* \}.
\]

Remark 4.4. If \( H(\cdot) \geq 0 \), the statement of the stopping problem in (2.3) implies that the value function \( V \) is non-increasing as a function of the discount factor \( \rho \). If we denote the dependence of the stopping region on \( \rho \) with \( \Gamma_T(\rho) \), then we have \( \Gamma_T(\rho_1) \subseteq \Gamma_T(\rho_2) \) whenever \( \rho_1 \leq \rho_2 \). Moreover, the dynamics of the process \( \bar{\Pi} \) are independent of \( \rho \) and \( U_0(s, \bar{\pi}) \) is the hitting time of \( \bar{\Pi} \) to \( \Gamma_T \). Therefore, the time that the DM can afford for observing the process \( X \) in the presence of a lower discount factor is no less than that spent under heavier discounting.

A similar claim also holds for dependence of \( U_0(s, \bar{\pi}) \) and \( \Gamma_T \) on the running costs \( c_i \). Namely, an observer with lower (in absolute value) running costs stops no-sooner than another one with heavier running costs.

5. Examples

Below, we re-visit the well-known Bayesian regime detection problem and the machine replacement problem of Jensen and Hsu [1993] in our finite horizon setting. For both problems, we also provide numerical solutions, which are obtained by discretizing the domain \([0, T] \times D \) of \( V(\cdot, \cdot) \) and solving the fixed point equation \( V(\cdot, \cdot) = J_0 V(t, \bar{\pi}) \) recursively. We set the number of iterations \( m \in \mathbb{N} \) such that the error \( \| V_m(\cdot) - V(\cdot) \| \) is negligible (see (3.3)). Our model is applicable in many other settings that have been considered elsewhere, including launch of insurance products Schöttl [1998], technology adoption Ulu and Smith [2007] and various disorder detection problems; see Ludkovski and Sezer [2007] for further details and examples.

5.1. Optimal replacement of a system. Here, we consider the reliability problem in Jensen and Hsu [1993] where the aim is to find the best time to replace a machine in order to maximize its lifetime net earnings. The objective is to compute
\[
\sup_{\tau} \mathbb{E}_\pi \left[ \int_0^\tau \sum_{i \in E} c_i 1_{\{ M_t = i \}} dt + \sum_{i \in E} \mu_i 1_{\{ M_t = i \}} \right].
\]
In this setting, the observations come from a simple Poisson process representing the number of defective items produced by the machine, and the process \( M \) represents the current productivity level. The \( n \)th state (defective state) is absorbing, while all others are transient. Related models have appeared in Makis and Jiang [2003], and Stadje [1994] and go all the way to classical POMDP work by Smallwood and Sondik [1973].
Figure 2. Value function $V(T, \bar{\pi})$ of the reliability example of Section 5.1. The shaded regions represent the stopping regions $\{\bar{\pi} \in D : V(T, \bar{\pi}) = H(\bar{\pi})\}$. Left and right panels are for the values $T = 1.5$ and $T = 0.2$ respectively. The shaded regions are the same in both panels. Note however the different $z$-scales. The panels also show the line $3.5\pi_1 + 1.5\pi_2 - \pi_3 = 0$, which is the stopping boundary of the ILA rule.

Assumption 1. In Jensen and Hsu [1993], it is assumed that (i) $q_i \neq 0$ for $i = 1, \ldots, n - 1$, with $q_n = 0$ (ii) $r_1 \geq r_2 \geq \ldots \geq r_n = c_n$, with $c_n < 0$ (iii) $0 < \lambda_1 \leq \ldots \leq \lambda_n$, (iv) $q_{in} > \lambda_n - \lambda_i$ for $i = 1, \ldots, n - 1$.

These assumptions ensure that the infinitesimal look-ahead rule $\tau^{ILA} := \inf\{t \geq 0: \sum_i r_i \Pi_t^{(i)} < 0\}$ is optimal where $r_i \triangleq c_i + \sum_{j \neq i} (\mu_j - \mu_i) q_{i,j}$ (cf. Lemma 4.1). It follows as a corollary to Jensen [1989, Theorem 3.1] that $\tau^{ILA} \wedge T$ is an optimal stopping rule for the finite horizon problem and the region $\{\bar{\pi} \in D : V(T, \bar{\pi}) = H(\bar{\pi})\}$ does not depend on $T$. This occurs because the instantaneous revenue rates $r_i$’s completely summarize the relative worth of different machine states, and the sum $\sum_{i \in E} r_i \Pi_t^{(i)}$ is monotonically non-increasing over time $\mathbb{P}^\bar{\pi}$-a.s. for all $\bar{\pi} \in D$ (see Jensen and Hsu [1993, Theorem 2]). Thus, $T$ only plays a role insofar as allowing the DM to collect profits before the machine deteriorates.

We illustrate this degeneracy in Figure 2. In this example, we select the parameters to fit the framework of Jensen and Hsu [1993]. We have a machine that moves through three regimes $E = \{1, 2, 3\}$ with transition matrix

$$ Q = \begin{pmatrix} -4 & 1.5 & 2.5 \\ 0 & -1.5 & 1.5 \\ 0 & 0 & 0 \end{pmatrix}. $$

At different states, the running profit from operating the machine is $\bar{c} = [1, 0, -1]$, and shutting down the machine involves a cost of $\bar{\mu} = [-1, -1, 0]$. In each state, the breakdowns occur according to independent Poisson processes with intensities $\bar{\lambda} = [2, 3, 4]$. In this setting, we have $\bar{r} = \{3.5, 1.5, -1\}$ so that $\tau^{ILA} = \inf\{t \geq 0: 3.5\Pi_t^{(1)} + 1.5\Pi_t^{(2)} - \Pi_t^{(3)} < 0\}$. The left and right panels of Figure 2 show the functions $V(T, \bar{\pi})$ and the regions $\{\bar{\pi} \in D : (T, \bar{\pi}) \in \Gamma_T\}$ for $T = 1.5$ and $T = 0.2$ respectively. We see that $V(0.2, \bar{\pi}) < V(1.5, \bar{\pi})$ but the regions $\{\bar{\pi} \in D : V(T, \bar{\pi}) = H(\bar{\pi})\}$ for $T = 0.2$ and $T = 1.5$ coincide with the region $\{\bar{\pi} \in D : 3.5\pi_1 + 1.5\pi_2 - \pi_3 \leq 0\}$, at least modulo the $D$-discretization necessary for numerical implementation.

This degenerate structure would disappear if one removes some of the assumptions in Jensen and Hsu [1993]. Nevertheless, the sequential construction of Section 3 can still safely be employed.
Figure 3. The second example for the reliability problem of Section 5.1 with the new parameters in (5.2). In the left panel $T = 2$, in the middle $T = 0.5$, and in the right panel $T = 0.1$. In each picture, the function $V(T, \pi)$ is plotted on $D$. Shaded regions are the sets $\{\pi \in D : V(T, \pi) = H(\pi)\}$.

with the optimal stopping rule given in Corollary 4.1. We give an example in Figure 3 where

\begin{equation}
Q = \begin{pmatrix}
-1 & 0.5 & 0.5 \\
0 & -0.5 & 0.5 \\
0 & 0 & 0
\end{pmatrix}
\quad \text{and} \quad \bar{\lambda} = [\lambda_1, \lambda_2, \lambda_3] = [1, 4, 7].
\end{equation}

We keep other parameters the same as in the previous example. Now, the instantaneous net gain $\sum_{i \in E} r_i \Pi_t^{(i)} = 1.5 \Pi_t^{(1)} + 0.5 \Pi_t^{(2)} - \Pi_t^{(3)}$ is not monotonically non-increasing $\mathbb{P}^{\pi}$-a.s. for all $\pi \in D$ anymore. Figure 3 shows that the stopping region is now time dependent and expands as time to maturity decreases. For this choice of $\bar{c}$ and $\bar{\mu}$, one can modify the proof of Lemma 4.2 to show that there always exists a stopping region around the absorbing state (for all $T \geq 0$). Furthermore, Lemma 4.1 implies that it is never optimal to stop around the corners of the simplex $D$ corresponding to non-absorbing states. Also note that the transition rates of $M$ are now lower. Therefore, the DM can obtain positive net gain when $M$ starts from the state $\{1\}$ and there is enough time to operate the system. Indeed, the first panel in Figure 3 shows that for $T = 2$ the value function is positive around the corner $\{1\}$.

5.2. Sequential hypothesis-testing. In this problem, a compound Poisson process $X = \{X_t\}_{t \geq 0}$ is observed starting from $t = 0$. The arrival rate $\lambda$ and mark distribution $\nu$ of $X$ are not known precisely. Rather they depend on the static regime of a Markov process $M$ with $n$ absorbing states (i.e., $M_t = M_0$ for all $t \geq 0$). Each state corresponds to the realization of one of the $n$ simple hypotheses

\begin{equation}
A_1 : (\lambda, \nu) = (\lambda_1, \nu_1), \ldots, A_n : (\lambda, \nu) = (\lambda_n, \nu_n),
\end{equation}

with given prior likelihoods $\pi_i$, for $i = 1, \ldots, n$. The objective of the DM is to identify the current regime as quickly as possible, with minimal probability of wrong decision.

In earlier work on this problem, the trade-off between observing and stopping is generally modeled via the Bayes risk

\begin{equation}
\mathbb{E}^{\pi} \left[ \tau + \sum_{k,i=1}^n \mu_{k,i} 1\{d=k, M_0=i\} \right],
\end{equation}
where \( \tau \) is the decision time, \( d \in \{1, \ldots, n\} \) represents the hypothesis selected and \( \mu_{k,i} \geq 0 \) is the cost of selecting the wrong hypothesis \( A_k \) when the correct one is \( A_i \). The DM then needs to minimize (5.4) and find a pair \((\tau, d)\), if one exists, that attains this infimum.

The infinite horizon version of (5.4) was solved for the first time by Peskir and Shiryaev [2000] for a simple Poisson process with \( n = 2 \). Later, Gapeev [2002] provided the solution (again with \( n = 2 \)), where the jump size is exponentially distributed under each hypothesis, and the mean of the exponential distribution is the same as the proposed arrival rate. The solution for any jump distribution and for \( n \in \mathbb{N} \) was recently provided by Dayanik et al. [2008a]. Below we treat the finite horizon version of that problem, where a decision must be made before horizon \( T < \infty \).

**Remark 5.1.** Let \( V(\infty, \bar{\pi}) \) denote the value function of this minimization problem on infinite-horizon, and for \( 1 \leq k \leq n \), let \( \Gamma_{\infty,k} \triangleq \{ \bar{\pi} \in D : \ V(\infty, \bar{\pi}) = H_k(\bar{\pi}) \} \) in terms of the functions \( H_k(\bar{\pi}) = \sum_{i \in E} \mu_{k,i} \pi_i \). Dayanik et al. [2008a] showed that each region \( \Gamma_{\infty,k} \) is closed and convex with a non-empty interior around the \( k \)th corner of the simplex \( D \). This structure also extends to the finite-horizon problem. Since \( V(\infty, \bar{\pi}) \leq V(T, \bar{\pi}) \), we have \( \Gamma_{\infty,k} \subseteq \Gamma_{T,k} \), for \( k \in E \) and \( T < \infty \). Then, Remarks 4.1 and 4.2 and Corollary 4.1 imply that there are time-dependent closed convex sets (with non-empty interiors) around the corners of \( D \) such that it is optimal to stop the first time the process \( \Pi \) enters one of these sets. At this time, if the conditional likelihoods process \( \bar{\Pi} \) is around the \( k \)th corner, we select hypothesis \( A_k \).

In Figure 4, we illustrate the time-dependence of the solution structure using a simple example with two hypotheses \( A_1 : \Lambda = \lambda_1 \) and \( A_2 : \Lambda = \lambda_2 \) on the arrival rate only. This problem was solved in Peskir and Shiryaev [2000] on infinite horizon, and the authors show that the immediate stopping is optimal if and only if \( \mu_{2,1,2}(\lambda_2 - \lambda_1) \leq \mu_{2,1} + \mu_{1,2} \). Hence, the inequality \( \mu_{2,1,2}(\lambda_2 - \lambda_1) > \mu_{2,1} + \mu_{1,2} \) has to be satisfied also in any finite-horizon problem with non-trivial solution.

In Figure 4, the arrival rates are \( \lambda_1 = 1 \) and \( \lambda_2 = 5 \). For the Bayes risk given in (5.4), we select \( \mu_{1,2} = \mu_{2,1} = 2 \) for the penalty costs. This numerical example matches Peskir and Shiryaev [2000, Figures 2-3]. The left panel of Figure 4 shows the value functions \( V(T, \cdot) \) with horizons \( T = 0.1, T = 0.2, T = 0.4 \) and \( T = 2 \) respectively, and the terminal reward \( H(\bar{\pi}) = \min\{\mu_{1,2} \pi_2 : \mu_{2,1}(1-\pi_2)\} \) on the state space of \( \pi_2 \in [0, 1] \). We see that as \( T \) increases, the value function decreases, as expected. The right panel of Figure 4 shows that the continuation region widens as time to maturity increases. We also observe that the boundary curves approach the solution structure of problem with infinite horizon. Peskir and Shiryaev [2000] obtained a continuation region of \([0.22, 0.70]\), very close to ours of \([0.230, 0.705]\) for \( T > 1 \).

Let us define the lower boundary curve \( T \mapsto b_1(T) \triangleq \sup\{\pi_2 \in [0, 1] : V(T, \bar{\pi}) = 2\pi_2\} \). Clearly \( b_1(0) = 0.5 \). In the right panel, we remarkably observe that the lower boundary curve \( b_1(\cdot) \) has a discontinuity at \( T = 0 \) and then remains constant until about \( T = 0.2 \). Note that the point \( \bar{\pi} = (\pi_1, \pi_2) = (0.5, 0.5) \) is the global maximum of the terminal cost function \( H(\bar{\pi}) \). Starting at the point \((0.5 + \varepsilon, 0.5 - \varepsilon)\), for \( \varepsilon \geq 0 \) and small, as long as there is no jump, the conditional likelihoods process \( \bar{\Pi} \) drifts (quickly) toward \( (1, 0) \) and away from this maximum. Intuitively speaking, for very small values of \( T \), the probability of observing a jump is low and thus it is optimal to continue. Therefore, the lower curve in Figure 4 is discontinuous around \( T = 0 \). The drift of the process \( \bar{\Pi} \) towards \((1, 0)\) decreases as \( \pi_2 \) decreases and approaches \((1, 0)\) (see (2.14)). As a result, at points \( \bar{\pi} \) where \( \pi_2 \) is small, the effect of waiting cost becomes dominant and it is optimal to stop even if \( T \) is small.
Figure 4. Bayesian regime detection example of Section 5.2. The left panel shows the value functions $V(T, \bar{\pi})$ for various time horizons $T$. The right panel shows the stopping regions $\Gamma_{T,k}$ (namely $\Gamma_{T,1}$ below the lower curve and $\Gamma_{T,2}$ above the higher curve) for $T = 2$.

The following proposition summarizes our discussion on this example and states that this behavior of the lower boundary curve around $T = 0$ holds for any set of parameters $\lambda_2 > \lambda_1$, $\mu_{1,2}$, $\mu_{2,1}$.

**Proposition 5.1.** Consider the hypothesis-testing problem in (5.4) with two simple hypotheses on the arrival rate: $A_1 : \Lambda = \lambda_1$ and $A_2 : \Lambda = \lambda_2$ (with $\lambda_2 > \lambda_1$). The continuation region $C_T$ is non-empty (for $T > 0$) if and only if $\mu_{2,1}\mu_{1,2}(\lambda_2 - \lambda_1) > \mu_{2,1} + \mu_{1,2}$. The boundary curve $T \mapsto b_1(T) \triangleq \sup\{\pi_2 \in [0, 1] : V(T, \bar{\pi}) = \mu_{1,2} \pi_2\}$ is discontinuous at $T = 0$, and there is an interval around $T = 0$ at which $b_1(\cdot)$ is constant.

**Remark 5.2.** As a final note, we would like to add that our analysis in Sections 3 and 4 can also be applied easily to solve the finite horizon change-detection problem. In this problem, the local parameters $(\lambda, \nu)$ of an observed compound Poisson process change at some unobservable time $\theta$ when the process $M$ hits one of its absorbing states, and the objective is to find the best time $\tau$ that minimizes $E(\tau - \theta)^+ + cP(\tau < \theta)$, which is another special case of (1.1).

In the infinite horizon setting, Dayanik and Sezer [2005] and Bayraktar and Sezer [2009] show that the stopping region consists of closed convex regions (again with non-empty interior) around the corners of $D$ corresponding to absorbing states. In the finite-horizon formulation, Remarks 4.1 and 4.2 and Lemma 4.2 imply that there are time-dependent closed convex sets around these corners, and the hitting time of $\Pi$ to those regions is an optimal alarm time (thanks to Corollary 4.1). Moreover, Lemma 4.1 implies that it is never optimal to stop around the remaining corners of $D$ corresponding to non-absorbing states.

**Acknowledgments**

The authors would like to thank the editors and anonymous referees for many helpful comments and remarks that improved the presentation in the paper.
Appendix

Appendix A. Discrete Information Costs

The objective function in (1.1) is applicable to a variety of economic settings. This has allowed us to provide a unified treatment of many disparate models. Returning to the economic interpretation of the running costs appearing in the first term in (1.1), in a typical setting they represent information acquisition expenses, or opportunity costs. Alternatively, observation costs may be incurred only when new information arrives. This, for example, happens if new information corresponds to opportunities lost (e.g. deals signed by competitors), leading to a cost structure of the form \( \sum_{j=1}^{N_{\tau}} e^{-\rho_{\tau}j}K(Y_j) \). Here, \( N_{\tau} \) is the number of arrivals by time \( \tau \), \( (\sigma_j, Y_j) \) are the arrival times and marks respectively, and \( K(Y_j) \) is the cost incurred upon an arrival of size \( Y_j \) (with \( K: \mathbb{R}^d \mapsto \mathbb{R} \) satisfying \( \nu_i K^+ \triangleq \int_{\mathbb{R}^d} K^+(y) \nu_i(dy) < \infty, \forall i \in E \)).

In the third case, one deals with the objective function

\[
(A.1) \quad \hat{U}(T, \pi) \triangleq \sup_{\tau \leq T, \pi \in \mathcal{F}^\pi} \mathbb{E}^\pi \left[ \sum_{j=1}^{N_{\tau}} e^{-\rho_{\tau}j}K(Y_j) + e^{-\rho_T} \sum_{k=1}^{\alpha} \mathbb{1}_{\{d=k\}} \left( \sum_{i \in E} \mu_{k,i} \cdot 1\{M_i = i\} \right) \right],
\]

by solving the equivalent stopping problem \( \hat{V}(T, \pi) \triangleq \sup_{\tau \leq T} \mathbb{E}^\pi \left[ \sum_{j=1}^{N_{\tau}} e^{-\rho_{\tau}j}K(Y_j) + e^{-\rho_T} H \left( \hat{\Pi}_\tau \right) \right], \)

as in Proposition 2.3. In this case, one can verify that the sequential approximation method of Section 3 holds for the value function \( \hat{V} \). Namely, if we define the sequence of functions \( \{\hat{V}_m(\cdot, \cdot)\}_{m \geq 0} \), where \( \hat{V}_m(s, \pi) \triangleq \sup_{\tau \leq s} \mathbb{E}^\pi \left[ \sum_{j=1}^{m N_{\tau}} e^{-\rho_{\tau}j}K(Y_j) + e^{-\rho_T} H \left( \hat{\Pi}_{\tau} \right) \right], \) it can be shown (see (3.5-3.7), Proposition 3.2) that we have \( \hat{V}_{m+1}(s, \pi) = \hat{J}_0 \hat{V}_m(s, \pi) \) where the operator \( \hat{J}_0 \) is defined as

\[
\hat{J}_0 w(s, \pi) = \sup_{t \in [0, s]} \mathbb{E}^\pi \left[ e^{-I(t)} \cdot e^{-\rho t} \cdot H \left( \hat{x}(t, \pi) \right) \right] + \int_0^t e^{-\rho u} \sum_{i \in E} m_i(t, \pi) \cdot \lambda_i \left( \int_{\mathbb{R}^d} K(y) \nu_i(dy) + S_i w(s - u, \hat{x}(u, \pi)) \right) du,
\]

for a bounded function \( w : [0, T] \times D \rightarrow \mathbb{R} \).

Clearly \( \{V_m\}_{m \geq 0} \) is an increasing sequence. Using the inequality \( \mathbb{E} \left[ \sum_{j=1}^{N_{\tau}} K^+(Y_j) \right] \leq (\max_{i \in E} \lambda_i) T \cdot (\max_{i \in E} \nu_i K^+) \) and the truncation arguments in the proof of Proposition 3.1, one can show that the sequence converges to \( \hat{V} \) uniformly with the error bound

\[ 0 \leq V - V_m \leq \left( \max_{i \in E} \lambda_i \right) T \cdot (\max_{i \in E} \nu_i K^+) + 2\|H\| \left( \frac{\lambda T}{m - 1} \right)^{1/2} \left( \frac{\lambda}{2\rho + \lambda} \right)^{m/2}. \]

Arguments in Sections 3 and 4 can then be replicated to conclude that

\[
\mathbb{E}^\pi \left[ \sum_{j=1}^{N_{\tau}(s, \pi)} e^{-\rho_{\tau}j}K(Y_j) + e^{-\rho_{\tau}j}H \left( \hat{\Pi}_{\tau}(s, \pi) \right) \right] \geq \hat{V}(s, \pi) - \varepsilon,
\]

for the stopping time \( \hat{U}_\varepsilon(s, \pi) \triangleq \inf \left\{ t \in [0, s] : \hat{V}(s - t, \hat{\Pi}_t) - \varepsilon \leq H(\hat{\Pi}_t) \right\} \). Hence, the admissible strategy \( (\hat{U}_\varepsilon(s, \pi), d(\hat{U}_\varepsilon(s, \pi))) \) is an optimal strategy for the problem in (A.1), as expected.

Furthermore, other results of Section 4 can be adjusted for this new objective function. Below, we summarize these results in a remark, and we conclude our discussion here.

**Remark A.1.** Let \( \nu_j K \triangleq \int_{\mathbb{R}^d} K(y) \nu_j(dy), \) for \( j \in E \).
(i) For a given index \( i \in E \), Define \( A^*(i) \triangleq \{ k \in A : \mu_{k,i} = \max_{j \in A} \mu_{j,i} \} \) as in Lemma 4.1. If 
\[-\rho \mu_{k,i} + \lambda_i \cdot \nu_i K + \sum_{j \neq i} (\mu_{k,j} - \mu_{j,i}) q_{i,j} > 0 \]
holds for all \( k \in A^*(i) \), then there exists some \( \hat{\pi}_i > 1 \) (for all \( T > 0 \)) such that it is optimal to continue on the time-
\( \tau \leq \hat{\tau}_i \). Lemma B.1.

(ii) Assume \( \nu_j K \leq 0 \) for all \( j \in E \), and \( \pi \triangleq \max_{k,i} \mu_{k,i} > 0 \), and let \( I^* \)
be as in (4.16). For \( i \in I^* \), if \( \nu_i K < 0 \) or \( \rho > 0 \) there exists a number \( \hat{\pi}_i > 1 \) (free of \( T \)) such that it is optimal
to stop at the points \( \pi \) for which \( \pi_i \geq \hat{\pi}_i \). That is: \( \Gamma_{T,i} \subseteq \{ [0,T] \times D : \pi_i \geq \hat{\pi}_i \} \) for all 
\( T \geq 0 \).

(iii) In the case where \( \nu_j K \leq 0 \) for all \( j \in E \), and \( H(\cdot) \geq 0 \), the stopping region is monotone in \( \rho \) and \( \nu_j K \), for \( j \in E \). Namely, if we increase one of these factors in absolute terms
(keeping everything else fixed), the stopping region expands, and the DM is forced to make
a decision sooner.

(iv) For a given \( \varepsilon > 0 \), let \( m \in \mathbb{N} \) such that \( \| \hat{V}(T,\cdot) - \hat{V}(T,\cdot) \| \leq \varepsilon / 2 \). Then the stopping time
\( \hat{T}^{(m)}(s,\pi) \triangleq \inf \{ t \in [0,T] : \hat{V}_m(T-t,\hat{\Pi}_t) - \varepsilon \leq H(\hat{\Pi}_t) \} \) gives an \( \varepsilon \)-optimal strategy.

(v) If “\( \rho > 0 \)” or “\( K(\cdot) \leq 0 \) with \( \max_{i \in E} \nu_i K(\cdot) < 0 \)”, then \( V(T,\cdot) \searrow \hat{V}(\infty,\cdot) \) uniformly as in
(B.2) if we redefine

\[
Err(T) \triangleq \begin{cases} 
\rho^T \left( \max_{i \in E} \lambda_i \cdot \max_{i \in E} \nu_i K^+ + 2 \cdot \| H(\cdot) \| \right), & \text{if } \rho > 0 \\
\frac{2 \cdot \| H(\cdot) \|}{T} \left( \min_{i \in E} \nu_i K^+ \right), & \text{if } \rho = 0, \ K(\cdot) \leq 0 \text{ and } \max_{i \in E} \nu_i K < 0.
\end{cases}
\]

APPENDIX B. REMARKS ON THE INFINITE HORIZON PROBLEM

In general, if there is a strict penalty for waiting, it is likely that the DM will make a decision
prior to the final time \( T \) for moderate or large values of \( T \). In this case, the constraint \( \tau \leq T \) in
(2.3) is of less importance, and one essentially faces an infinite horizon stopping problem. Solving
the infinite horizon problem can be computationally more appealing since we eliminate the time-
dimension of the state space \( [0,T] \times D \). Below, we show that the value function of the finite-horizon
problem converges uniformly to that of the infinite horizon under the assumption

\[
(\text{B.1}) \quad \text{either } \rho > 0 \text{ or } \max_{i \in E} c_i < 0.
\]

The infinite horizon problem is defined as in (2.3) (and (1.1)) by removing the constraint \( \tau \leq T \).
With the notation in (2.3), let \( V(\infty,\pi) \) be the value function of this stopping problem.

**Lemma B.1.** As \( T \searrow \infty \), the function \( V(T,\pi) \) converges to \( V(\infty,\pi) \) uniformly on \( D \), and we have

\[
(\text{B.2}) \quad V(T,\pi) \leq V(\infty,\pi) \leq V(T,\pi) + Err(T), \quad \text{for all } \pi \in D \text{ and } T \geq 0,
\]

where

\[
Err(T) \triangleq \begin{cases} 
\rho^T \left( \| C \| + 2 \cdot \| H(\cdot) \| \right), & \text{if } \rho > 0 \\
\frac{2 \cdot \| H(\cdot) \|}{T} \left( \min_{i \in E} \nu_i K^+ \right), & \text{if } \rho = 0, \ K(\cdot) \leq 0 \text{ and } \max_{i \in E} \nu_i K < 0.
\end{cases}
\]
Proof. The first inequality in (B.2) is obvious. To show the second inequality let \( \tau \) be an \( \mathbb{F} \)-stopping time. Then, we have

\[
\mathbb{E}^\mathbb{F} \left[ \int_0^T e^{-\rho t} (\Pi_t) dt + e^{-\rho T} H(\Pi_T) \right] \leq \mathbb{E}^\mathbb{F} \left[ \int_0^{\tau \wedge T} e^{-\rho t} C(\Pi_t) dt + e^{-\rho \tau \wedge T} H(\Pi_{\tau \wedge T}) \right] \\
+ \mathbb{E}^\mathbb{F} \left[ 1_{\{\tau \geq T\}} \left( \int_T^\infty e^{-\rho t} C(\Pi_t) dt + e^{-\rho t} H(\Pi_t) - e^{-\rho T} H(\Pi_T) \right) \right]
\]

If \( \rho > 0 \), the last expectation above is bounded above by \( e^{-\rho T} (\|C\| + 2 \cdot \|H\|) \). Then taking the supremum over all \( \tau \)'s on both sides we obtain (B.2).

On the other hand, if \( \rho = 0 \) and \( \max_{i \in E} c_i < 0 \), we may safely restrict ourselves to the set of stopping times \( \tau \) for which \( \mathbb{E}[\tau] \leq (\min_{k,i} \mu_{k,i} - \max_{k,i} \mu_{k,i}) / \max_{i \in E} c_i \): the expected reward associated with any stopping time having a higher expected value is dominated by the reward achieved upon stopping immediately. Then, the second expectation in (B.3) is bounded above by

\[
2 \cdot \|H\| \cdot \mathbb{P}\{\tau > T\} \leq 2 \cdot \|H\| \cdot \mathbb{E}[\tau] \leq \frac{2 \cdot \|H\|}{T} \left( \frac{\min_{k,i} \mu_{k,i} - \max_{k,i} \mu_{k,i}}{\max_{i \in E} c_i} \right),
\]

thanks to Markov's inequality. Then, the inequality in (B.2) follows after taking the supremums over \( \tau \) again.

\[\square\]

The explicit error bounds for the rate of convergence allows us to approximate \( V(T, \cdot) \) with the value function of the infinite horizon problem when \( T \) is large. The function \( V(\infty, \tilde{\pi}) \) can be computed sequentially as in Section 3. That is, if we define the non-decreasing sequence

\[
V_m(\infty, \tilde{\pi}) \triangleq \sup_{\tau \geq 0} \mathbb{E}^\mathbb{F} \left[ \int_0^{\tau \wedge \sigma_{m}} e^{-\rho t} C(\Pi_t) dt + e^{-\rho \tau \wedge \sigma_{m}} H(\Pi_{\tau \wedge \sigma_{m}}) \right], \quad m \in \mathbb{N},
\]

then it can be shown that the elements of this sequence can be computed by applying a functional operator \( \hat{J}_0 \), which is obtained from the operator \( J_0 \) in (3.6) after replacing the constraint \( t \in [0, s] \) with \( t \geq 0 \). Also, note that the new operator \( \hat{J}_0 \) is defined on the domain of functions defined on \( D \) only. The proof of these statements can be obtained by modifying the arguments of Section 3, or those in Dayanik et al. [2008a, Section 3]. Moreover, following the proof of Proposition 3.1 and the arguments of Section 4.1, we have

\[
\|V_m(\infty, \cdot) - V(\infty, \tilde{\pi})\| \leq \text{Err}_\infty(m) \triangleq \begin{cases} \left( \frac{\lambda}{\rho + \lambda} \right)^m, & \text{if } \rho > 0, \\ \frac{\max_{k,i} \mu_{k,i}}{\max_{i \in E} c_i} \cdot \frac{\lambda}{m - 1}^{1/2}, & \text{if } \rho = 0 \text{ and } \max_{i \in E} c_i < 0, \end{cases}
\]

and the stopping time

\[
U^{(m)}(\infty, \tilde{\pi}) \triangleq \inf \left\{ t \geq 0 : V_m(\infty, \Pi_t) - \varepsilon \leq H(\Pi_t) \right\}
\]

is \( \varepsilon \)-optimal for the infinite horizon problem (see also Dayanik et al. [2008a, Section 4.1]).

Note that for large \( m \), the function \( V_m(\infty, \cdot) \) approximates the function \( V(\infty, \cdot) \), and for large \( T, V(\infty, \cdot) \) is a good approximation for \( V(T, \cdot) \). However, the stopping rule in (B.5) is not a good substitute for the optimal time \( U_0(T, \tilde{\pi}) \) since the former may not be less than \( T \) almost surely. Moreover, since \( U^{(m)}(\infty, \tilde{\pi}) \) may be greater than \( U_0(T, \tilde{\pi}) \), Proposition 4.1 is not necessarily true.
and the martingale property (4.5) may fail. Nevertheless, if we apply the rule $U^{(m)}_\varepsilon(\infty, \bar{\pi}) \land T$, we can still control the error for large $T$. Indeed, it can be shown (Ludkovski and Sezer [2007]) that

(B.6) $V(T, \bar{\pi}) \leq \mathbb{E}^\bar{\pi} \left[ \int_0^{U^{(m)}_\varepsilon(\infty, \bar{\pi}) \land T} e^{-\rho t} C(\bar{\Pi}_t) \, dt + e^{-\rho (U^{(m)}_\varepsilon(\infty, \bar{\pi}) \land T)} H \left( \bar{\Pi}, U^{(m)}_\varepsilon(\infty, \bar{\pi}) \land T \right) \right] + \varepsilon + \text{Err}_\infty(m) + \text{Err}_\infty(0) \cdot \text{Err}(T)$.

Hence, if $T$ is large enough (so that $\text{Err}_\infty(0) \cdot \text{Err}(T)$ is small), by taking $\varepsilon$ in (B.5) small for a large value of $m$, the error associated with applying $U^{(m)}_\varepsilon \land T$ can be reduced to acceptable levels.

**Appendix C. Supplementary Proofs**

**Proof of Remark 2.1.** In order to establish (2.10-2.11), let $\mathbb{E}_j[\cdot]$ denote the expectation operator $\mathbb{E}^\bar{\pi}[\cdot \mid \mathcal{M}_0 = j]$, and let $t_m \leq t < t + u < t_{m+1}$. Here $t_m$ and $t_{m+1}$ can be considered as the sample realization $\sigma_m(\omega)$ and $\sigma_{m+1}(\omega)$ of the $m$'th and $m+1$'st arrival times respectively. Using the definition of $L^\bar{\pi}_i$ in (2.7) we have $L^\bar{\pi}_i(t + u, m : (t_k, y_k), k \leq m) = \sum_{j \in E} \pi_j \cdot \mathbb{E}_j \left[ 1_{\{M_{t+u}=i\}} \cdot e^{-I(t+u)} \cdot \prod_{k=1}^m \ell(t_k, y_k) \right]$. Using the Markov property of $M$, the expression in (C.1) can be written as

$$= \sum_{j \in E} \pi_j \cdot \mathbb{E}_j \left[ e^{-I(t)} \prod_{k=1}^m \ell(t_k, y_k) \right] \cdot \mathbb{E}_l \left[ 1_{\{M_u=i\}} \cdot e^{-I(u)} \right] \cdot L^\bar{\pi}_l(t, m : (\sigma_k, y_k), k \leq m).$$

Then the explicit form of $\bar{\Pi}$ in (2.9) implies that for $\sigma_m \leq t \leq t + u < \sigma_{m+1}$, we have

(C.2) $\Pi^{(i)}(t + u) = \frac{\sum_{j \in E} L^\bar{\pi}_l(t, m : (\sigma_k, y_k), k \leq m)}{\sum_{j \in E} L^\bar{\pi}_l(t, m : (\sigma_k, y_k), k \leq m)} \cdot \mathbb{E}_l \left[ 1_{\{M_u=i\}} \cdot e^{-I(u)} \right] = \frac{\mathbb{E}^{\bar{\pi}} \left[ 1_{\{M_u=i\}} \cdot e^{-I(u)} \right]}{\mathbb{E}^{\bar{\pi}} \left[ 1_{\{M_u=j\}} \cdot e^{-I(u)} \right]} = \frac{\mathbb{P}^{\bar{\pi}} \{ \sigma_1 > u, M_u = i \}}{\mathbb{P}^{\bar{\pi}} \{ \sigma_1 > u \}} \bar{\pi} = \bar{\Pi}_t$.

On the other hand, the expression in (2.7) gives

(C.3) $L^\bar{\pi}_i(\sigma_{m+1}, m + 1 : (\sigma_k, Y_k), k \leq m + 1) = \mathbb{E}^\bar{\pi} \left[ 1_{\{M_{t+u}=i\}} \prod_{k=1}^{m+1} \ell(t_k, y_k) \right] = \lambda_i f_i(Y_{m+1}) \mathbb{E}^\bar{\pi} \left[ 1_{\{M_{t+u}=i\}} \prod_{k=1}^{m+1} \ell(t_k, y_k) \right]$. Hence, at arrival times $\sigma_1, \sigma_2, \ldots$ of $X$, the process $\bar{\Pi}$ exhibits a jump behavior and satisfies the recursive relation $\Pi^{(i)}(\sigma_{m+1}) = \frac{\lambda_i f_i(Y_{m+1}) L^\bar{\pi}_i(\sigma_{m+1}+, m : (\sigma_k, Y_k), i \leq m)}{\sum_{j \in E} \lambda_j f_j(Y_{m+1}) L^\bar{\pi}_j(\sigma_{m+1}+, m : (\sigma_k, Y_k), k \leq m)}$ and

(C.4) $\frac{\lambda_i f_i(Y_{m+1}) L^\bar{\pi}_i(\sigma_{m+1}+, m : (\sigma_k, Y_k), i \leq m)}{\sum_{j \in E} \lambda_j f_j(Y_{m+1}) L^\bar{\pi}_j(\sigma_{m+1}+, m : (\sigma_k, Y_k), k \leq m)} = \frac{\lambda_i f_i(Y_{m+1}) \Pi^{(i)}(\sigma_{m+1}+) \Pi^{(i)}(\sigma_{m+1}-)}{\sum_{j \in E} \lambda_j f_j(Y_{m+1}) \Pi^{(j)}(\sigma_{m+1}-)}$ for $m \in \mathbb{N}$. 
The identities in (C.2) and (C.4) give (2.10-2.11). By repeating (C.1-C.2) with $m = 0$ (i.e., with no arrivals on $[0, t + s]$), we see that the paths $t \mapsto \bar{x}(t, \bar{\pi})$ have the semigroup property $ar{x}(t + u, \bar{\pi}) = \bar{x}(u, \bar{x}(t, \bar{\pi}))$. \hfill \Box

**Proof of Proposition 3.1.** The inequality $V_m(s, \bar{\pi}) \leq V(s, \bar{\pi})$ is immediate. To show the second inequality, let $\tau$ be an $\mathbb{F}$-stopping time less than $s$ $\mathbb{P}$-a.s.. Then we have

\begin{equation}
\mathbb{E}^\pi \left[ \int_0^\tau e^{-\rho t} C(\bar{\Pi}_t) dt + e^{-\rho \tau} H \left( \bar{\Pi}_{\tau} \right) \right] = \mathbb{E}^\pi \left[ \int_0^\tau e^{-\rho t} C(\bar{\Pi}_t) dt + e^{-\rho \tau} H \left( \bar{\Pi}_{\tau} \right) \right] \leq \mathbb{E}^\pi \left[ \int_0^\tau e^{-\rho t} C(\bar{\Pi}_t) dt + e^{-\rho \tau} H \left( \bar{\Pi}_{\tau} \right) \right] + (T\|C\| + 2\|H\|) \cdot \mathbb{E}^\pi \left[ e^{-\rho \tau} m 1_{\{T > \tau\}} \right]
\end{equation}

where the last line follows since $s \leq T$ and $\{\tau > \sigma_m\} \subseteq \{T > \sigma_m\}$. Using the Cauchy-Schwarz inequality and the inequalities $\mathbb{P}^\pi \{ T > \sigma_m \} \leq \mathbb{P}^\pi [1_{\{T > \sigma_m\}}/T] \leq T \cdot \mathbb{E}^\pi [1/\sigma_m]$ we obtain

$$\mathbb{E}^\pi \left[ e^{-\rho \tau} m 1_{\{T > \tau\}} \right] \leq \sqrt{T \mathbb{E}^\pi [1/\sigma_m] \mathbb{E}^\pi [e^{-2\rho \tau}].}$$

Conditioning on $M$, $\mathbb{E}^\pi [e^{-u \sigma_1} | M] = \mathbb{E}^\pi \left[ \int_0^\infty u \cdot e^{-ut} dt \right] = \int_0^\infty \mathbb{P}^\pi [\sigma_1 \leq t | M] u \cdot e^{-ut} dt = \int_0^\infty \left[ 1 - e^{-\lambda t} \right] u \cdot e^{-ut} dt \leq \int_0^\infty \left[ 1 - e^{-\lambda t} \right] u \cdot e^{-ut} dt = \frac{\lambda}{u + \lambda}.$

The process $X$ has independent increments conditioned on $M$. Then, by induction we have

\begin{equation}
\mathbb{E}^\pi \left[ e^{-u \sigma_m} \right] \leq \left( \frac{\lambda}{u + \lambda} \right)^m,
\end{equation}

for all $m \in \mathbb{N}$. Moreover, since $1/\sigma_m = \int_0^\infty e^{-\sigma_m u} du$, the inequality in (C.6) gives $\mathbb{E}^\pi [1/\sigma_m] \leq \int_0^\infty (\lambda^m / u + \lambda)^m du = \lambda / (m - 1)$, for $m \geq 2$. By using these upper bounds in (C.5) and taking the supremum of both sides we obtain (3.3). \hfill \Box

**Proof of Lemma 3.1.** To establish the convexity, we will show that expression in (3.6) is convex (in $\bar{\pi}$) for each $t$ and $s$.

We first note that $\mathbb{E}^\pi \left[ e^{-L(t)} \right] = \sum_{j \in E} \pi_j \mathbb{E}_j \left[ e^{-L(t)} \right]$ and $m_i(t, \bar{\pi}) = \sum_{j \in E} \pi_j \mathbb{E}_j \left[ 1_{\{\lambda_i = 1\}} e^{-L(t)} \right]$ are linear in $\bar{\pi}$ where $m_i(t, \bar{\pi})$ is defined in (2.13) for $i \in E$ and $\mathbb{E}_j$ is the expectation operator $\mathbb{E}_j[\cdot | M_j = j]$ for $j \in E$. Then we see that the expression $\mathbb{E}^\pi \left[ e^{-L(t)} \right] e^{-\rho t} H(\bar{x}(t, \bar{\pi})) = \max_{k \in A} e^{-\rho t} \sum_{i \in E} \mu_{k,i} m_i(t, \bar{\pi})$ is convex as the upper envelope of convex functions. Next we let $\bar{\pi} \mapsto w(s, \bar{\pi})$ be a convex mapping for each $s \geq 0$. Then we have $w(s, \bar{\pi}) = \sup_{k \in K_s} \beta_{k,0}(s) + \beta_{k,1}(s) \pi_1 + \ldots + \beta_{k,n}(s) \pi_n$, for some index set $K_s$, and each $\beta_{k,i}(s)$ is a function in $s$. Using this characterization with the definition of the operator $S_i$ in (3.7) we obtain $\int_0^t e^{-\rho u} \sum_{i \in E} \mathbb{E}^\pi \left[ 1_{\{\lambda_i = 1\}} e^{-L(u)} \right] \cdot \lambda_i S_i w(s - u, \bar{x}(u, \bar{\pi})) du = \int_0^t e^{-\rho u} \sum_{i \in E} \lambda_i m_i(u, \bar{\pi}) \cdot \left[ \int_{\mathbb{R}^d} \sup_{k \in K_{s-u}} \left( \frac{\lambda_j f_j(y) m_j(u, \bar{\pi})}{\sum_{l \in E} \lambda_l f_l(y) m_l(u, \bar{\pi})} \right) f_i(y) \nu(dy) \right] du
\end{equation}

\begin{align*}
&= \int_0^t e^{-\rho u} \left[ \int_{\mathbb{R}^d} \sup_{k \in K_{s-u}} \left( \sum_{j \in E} \beta_{k,j}(s - u) + \beta_{k,0}(s - u) \right) \frac{\lambda_j f_j(y) m_j(u, \bar{\pi})}{\sum_{l \in E} \lambda_l f_l(y) m_l(u, \bar{\pi})} \right] \nu(dy) \right] du.
\end{align*}
Since the expression inside the supremum operator are linear in π, the integrand in the inner integral is convex, and therefore so is the expression above. Also note that \( \int_0^t e^{-\rho u} \sum_{i \in E} m_i(u, \vec{π}) C(\vec{x}(u, \vec{π})) du = \int_0^t e^{-\rho u} \sum_{i \in E} \epsilon_i m_i(u, \vec{π}) du \), where both the integrand and the integral are linear in \( \vec{π} \). Finally, as the sum of three convex functions \( \vec{π} \mapsto Jw(t, s, \vec{π}) \) is convex. Since \( J_0w(s, \vec{π}) \) is the supremum of convex functions, it is again convex.

Boundedness and monotonicity are immediate by the definition of the operator \( J \) in (3.4). Finally, the continuity of \( J_0w \) for a (bounded) continuous function \( w(\cdot, \cdot) \) on \([0,T] \times D\) follows from the continuity of \( S_tw, H(\cdot) \) and regularity of paths \( t \mapsto \vec{x}(t, \vec{π}) \).

**Proof of Proposition 3.2.** For every \( \varepsilon \geq 0 \), let us define

\[
(C.7) \quad r^\varepsilon_m(s, \vec{π}) \triangleq \inf\{ t \in [0, s] : Jv_m(t, s, \vec{π}) \geq J_0v_m(s, \vec{π}) - \varepsilon \}, \quad \vec{π} \in D,
\]

and

\[
S^\varepsilon_1(s, \vec{π}) \triangleq r^\varepsilon_0(s, \vec{π}) \wedge \sigma_1 \quad \text{and} \quad S^\varepsilon_{m+1}(s, \vec{π}) \triangleq \begin{cases} r^\varepsilon_m(s, \vec{π}) & \text{if } \sigma_1 > r^\varepsilon_m(s, \vec{π}), \\ \sigma_1 + S^\varepsilon_m(s - \sigma_1, \vec{π}) & \text{if } \sigma_1 \leq r^\varepsilon_m(s, \vec{π}). \end{cases}
\]

Then, for every \( m \geq 1 \) we have Dayanik et al. [2008a, Prop. 3.5]

\[
(C.8) \quad \mathbb{E}^D \left[ \int_0^{S^\varepsilon_m(s, \vec{π})} e^{-\rho t} C(\vec{π}_t) dt + e^{-\rho S^\varepsilon_m(s, \vec{π})} H \left( \vec{π}_{S^\varepsilon_m(s, \vec{π})} \right) \right] \geq v_m(s, \vec{π}) - \varepsilon.
\]

The inequality \( V_m \geq v_m \) follows from (C.8) since \( S^\varepsilon_m(s, \vec{π}) \leq s \wedge \sigma_m \) by construction. The reverse inequality \( V_m \leq v_m \) is obtained analogously to Dayanik et al. [2008a, Proof of Prop. 3.2]. Namely, we show

\[
(C.9) \quad \mathbb{E} \left[ \int_0^{\tau \wedge \sigma_m} e^{-\rho t} C(\vec{π}_t) dt + e^{-\rho \tau \wedge \sigma_m} H \left( \vec{π}_{\tau \wedge \sigma_m} \right) \right] \leq v_m(s, \vec{π}),
\]

for every bounded stopping time \( \tau \leq s \) and \( m \in \mathbb{N} \) using the strong Markov property of \( \vec{π} \). Taking the infimum of both sides in (C.9), we arrive at the desired inequality \( V_m \leq v_m \).

**Proof of Proposition 3.3.** Using (3.8) and Corollary 3.1 we get \( V(s, \vec{π}) = v(s, \vec{π}) = \sup_{n \geq 1} v_n(s, \vec{π}) = \sup_{n \geq 1} \sup_{t \in [0, s]} Jv_{n-1}(t, s, \vec{π}) = \sup_{t \in [0, s]} \sup_{n \geq 1} Jv_{n-1}(t, s, \vec{π}) = \sup_{t \in [0, s]} \sup_{n \geq 1} \mathbb{E}^D \left[ e^{-I(t)} \right] e^{-\rho t} H(\vec{x}(t, \vec{π}))

\[
= \sup_{t \leq s} \mathbb{E}^D \left[ e^{-I(t)} \right] e^{-\rho t} H(\vec{x}(t, \vec{π})) + \int_0^t e^{-\rho u} \sum_{i \in E} m_i(u, \vec{π}) \left( C(\vec{x}(u, \vec{π})) + \lambda_i S_i v_{n-1}(s - u, \vec{x}(u, \vec{π})) \right) du
\]

where the fifth equality is from (3.6) and the sixth equality is by the bounded convergence theorem since we have \( \|v_m(\cdot, \cdot)\| \leq \|v(\cdot, \cdot)\| \leq \|H(\cdot)\| + T \|C(\cdot)\| \) for all \( m \in \mathbb{N} \).

Let \( W(\cdot, \cdot) \) be another solution of \( W(s, \vec{π}) = J_0W(s, \vec{π}) \), such that \( W(s, \vec{π}) \geq H(\vec{π}) = v_0(s, \vec{π}) \). Then \( W(s, \vec{π}) = J_0W(s, \vec{π}) \geq \sup_{t \in [0, s]} Jv_0(t, s, \vec{π}) = v_1(s, \vec{π}) \). By induction, \( W(s, \vec{π}) \geq v_n(s, \vec{π}) \) for all \( n \), and hence \( W(s, \vec{π}) \geq \lim_{n \to \infty} v_n(s, \vec{π}) = V(s, \vec{π}) \).

**Proof of Lemma 4.1.** Let \( \vec{e}_i \in D \) denote the point whose \( i \)th component is equal to 1. To establish the result it is sufficient to find a closed ball with strictly positive radius around \( \vec{e}_i \) (e.g.,
a region of the form \{\bar{\pi} \in D : ||\bar{\pi} - \bar{c}_i|| \leq \delta\} for some \delta > 0, where ||\cdot|| denotes the Euclidean norm on \mathbb{R}^n) such that \( H(\bar{\pi}) < v_1(s, \bar{\pi}) \leq V(s, \bar{\pi}) \) for all points on this closed ball.

We first note that there exists a closed ball \( B_0 \) around \( \bar{\pi} \) with positive radius such that \( H(\bar{\pi}) = \max_{k \in A^*(i)} H_k(\bar{\pi}) \), for \( \bar{\pi} \in B_0 \). Then on \( B_0 \) and for small \( s > 0 \) we have \( v_1(s, \bar{\pi}) = \sup_{t \leq s} J_0(t, s, \bar{\pi}) = \max_{k \in A^*(i)} \sup_{t \in [0, s]} J_0^{(k)} H(t, \bar{\pi}) \), where \( J_0^{(k)} H(t, \bar{\pi}) \triangleq \)

\[
\mathbb{E}[e^{-\rho t} H_k(\bar{x}(t, \bar{\pi})) + \int_0^t e^{-\rho u} \sum_{j \in E} m_j(u, \bar{\pi})(C(\bar{x}(u, \bar{\pi})) + \lambda_j S_j H(\bar{x}(u, \bar{\pi}))) du].
\]

Then, using (2.14) we have \( dJ_0^{(k)} H(t, \bar{\pi})/dt\big|_{t=0} \geq \left( -\rho - \sum_{j \in E} \lambda_j \pi_j \right) H_k(\bar{\pi}) + \sum_{j \in E} \mu_{k,j} \left( \sum_{l \in E} q_{l,j} \pi_l - \lambda_j \pi_j + \pi_j \sum_{l \in E} \lambda_l \pi_l \right) + C(\bar{\pi}) + \sum_{j \in E} \lambda_j \pi_j S_j H_k(\bar{\pi}). \)

The right hand side of the inequality above is uniformly continuous on the compact set \( D \). Its value at the point \( \bar{\pi} \) equals \( c_i - \rho \mu_{k,i} + \sum_{j \neq i} (\mu_{k,j} - \mu_{k,i}) q_{k,j} > 0 \). Hence for some \( \delta_k > 0 \) there exists an open ball (contained in \( B_0 \)) with radius \( \delta_k \) around \( \bar{\pi} \) such that \( dJ_0^{(k)} H(t, \bar{\pi})/dt\big|_{t=0} > 0 \) for all the points in this ball. Let \( B_k \) be the closed ball around the same point \( \bar{\pi} \) with radius \( \delta_k/2 \). Then on the intersection set \( \cap_{k \in A^*(i)} B_k \) the mapping \( \bar{\pi} \mapsto dJ_0^{(k)} H(t, \bar{\pi})/dt\big|_{t=0} \) is strictly positive and \( \sup_{t \geq 0} J_0^{(k)} H(t, \bar{\pi}) > H_k(t, \bar{\pi}) \) for all \( k \in A^*(i) \). This implies that \( v_1(s, \bar{\pi}) > H(\bar{\pi}) \) for all \( s > 0 \) on \( \cap_{k \in A^*(i)} B_k \).

**Proof of Lemma 4.2.** Let \( i \in I^* \) for \( I^* \) defined in (4.16). To establish the result, we will find \( \pi_i^* < 1 \) such that \( H(\bar{\pi}) = J_0 w(s, \bar{\pi}) \) on \( \{(s, \bar{\pi}) \in [0, T] \times D : \pi_i^* \leq \pi_i < 1\} \) for a bounded function \( w(\cdot) \equiv ||H|| = \bar{\mu} \triangleq \max_{k \in A} \mu_{k,i} \). Since \( V \) is bounded by the same upper bound (recall that \( c_i \leq 0 \) for \( i \in E \) by assumption) and satisfies \( V(s, \bar{\pi}) = J_0 V(s, \bar{\pi}) \) we will have \( H(\cdot) = V(\cdot) \) on this region.

**Part 1:** Let us first define

\[
F_k(t, \bar{\pi}) \triangleq \mathbb{E}[e^{-\rho t} H_k(\bar{x}(t, \bar{\pi})) + \int_0^t e^{-\rho u} \sum_{j \in E} m_j(u, \bar{\pi})(C(\bar{x}(u, \bar{\pi})) + \lambda_j S_j H(\bar{x}(u, \bar{\pi}))) du].
\]

Since \( H(\bar{\pi}) \leq J_0 w(s, \bar{\pi}) = \sup_{t \in [0, s]} J_0 w(t, s, \bar{\pi}) \leq \sup_{t \in [0, s]} \max_{k \in A} F_k(t, \bar{\pi}) = \max_{k \in A} \sup_{t \in [0, s]} F_k(t, \bar{\pi}) \) (see (3.6)), it is enough to show that for some \( \pi_i^* < 1 \) we have \( \sup_{t \geq 0} F_k(t, \bar{\pi}) = H_k(\bar{\pi}) \) for all \( k \in A \).

Let \( \hat{\pi}_i < 1 \) be a value such that \( H(\hat{\pi}) = \max_{k \in A^*} h_k(\bar{\pi}) \), where \( A^* \triangleq \{k \in A : \mu_{k,i} = \bar{\mu}\} \). That is, we have \( \mu_{k,i} = \bar{\mu} \) for all \( k \in A^* \) (and \( i \in I^* \)). Note that \( \hat{\pi}_i \) can for instance be selected as

\[
\hat{\pi}_i = \max_{k \in A^*} \frac{\bar{\mu} - \min_{k \in A} \mu_{k,i}}{2\bar{\mu} - \min_{k \in A} \mu_{k,i} - a_{k,i}}.
\]

Let us then define the hitting time \( T(\bar{\pi}, \hat{\pi}_i) \triangleq \inf \{t \geq 0 : x_i(t, \bar{\pi}) \leq \hat{\pi}_i\} \). For \( t \leq T(\bar{\pi}, \hat{\pi}_i) \), we have \( \max_{k \in A} H_k(\bar{x}(t, \bar{\pi})) = \max_{k \in A^*} H_k(\bar{x}(t, \bar{\pi})) \), which implies \( \max_{k \in A} F_k(t, \bar{\pi}) = \max_{k \in A^*} F_k(t, \bar{\pi}) \). Note that we have \( dF_k(t, \bar{\pi})/dt = \)

\[
\sum_{i \in E} \mathbb{E}[1_{\{\mathcal{M}_i = i\}} e^{-\rho t} \left\{-\lambda_i + \rho \cdot H_k(\bar{x}(t, \bar{\pi})) + \frac{dH_k(\bar{x}(t, \bar{\pi}))}{dt} \right\} + C(\bar{x}(t, \bar{\pi})) + \lambda_i ||H|| \}
\]

where

\[
dH_k(\bar{x}(t, \bar{\pi})) = \sum_{i \in E} \mu_{k,i} \left( \sum_{j} q_{ji} x_j(t, \bar{\pi}) - \lambda_i x_i(t, \bar{\pi}) + x_i(t, \bar{\pi}) \sum_{j} \lambda_j x_j(t, \bar{\pi}) \right)
\]
due to (2.14). Let us denote \( \mu \triangleq \min_{k,i} \mu_{k,i} \). For \( k \in A^* \), we have \( H_k(x(t, \pi)) = \overline{\pi} x_k(t, \pi) + \sum_{i \neq i} \mu_{k,i} x_i(t, \pi) \geq \overline{\pi} x_k(t, \pi) + \mu (1 - x_k(t, \pi)) \). Using this inequality, we get an upper bound for the derivative in (C.11) as \( dF_k(t, \pi)/dt \)

\[
\sum \lambda x_j(t) - \sum \lambda x_j(t) \sum \lambda_j x_j(t) \leq \left( 1 - x_i(t, \pi) \right) \cdot \left( 3 \cdot \overline{\pi} \cdot \lambda + n \cdot \left( \max_{l,j} |q_{l,j}| \right) \cdot \overline{\pi} \right)
\]

thanks to the inequality \( \sum_{i \in E} \mu_{k,i} dq_i \leq 0 \) (recall that \( \mu = \mu_{k,i} = \max_{k,i} \mu_{k,i} \) and \( q_{ii} = - \sum_{i \neq i} q_{ii} \)). The inequalities (C.13) and (C.14) then imply that for \( t < T(\pi, \pi_i) \), and for \( k \in A^* \);

\[
\frac{dF_k(t, \pi)}{dt} \leq \mathbb{E}^\pi \left[ e^{-I(t)-\rho t} \cdot \left( \mu (1 - x_i(t, \pi)) + c_i x_i(t, \pi) \right) \cdot \overline{\pi} \cdot \lambda \right]
\]

where \( G \triangleq 4 \cdot \overline{\pi} \cdot \lambda + n \cdot (\max_{l,j} |q_{l,j}|) \cdot \overline{\pi} - (\rho + \lambda) \cdot \overline{\mu} \). Note that the assumption \( \rho > 0 \) or \( c_i > 0 \) in Lemma 4.2 assures that \( dF_k(t, v_p)/dt \big|_{t=0} \) is negative as \( \pi_i \to 1 \). Therefore, if we define

\[
\hat{\pi}_i \triangleq \max \left\{ \hat{\pi}_i, \frac{G}{\rho \overline{\mu} - c_i + G} \right\} = \max \left\{ \hat{\pi}_i, \frac{4 \overline{\pi} \lambda + n (\max_{l,j} |q_{l,j}|) \overline{\pi} - (\rho + \lambda) \mu}{-c_i + n \overline{\pi} (\max_{l,j} |q_{l,j}|) + 3 \lambda \overline{\pi} + (\overline{\pi} - \mu)(\rho + \lambda)} \right\} < 1,
\]

we have \( dF_k(t, \pi)/dt \leq 0 \) on \( t \in [0, T(\pi, \pi_n)] \) for all \( k \in A^* \) and for all \( \pi \) such that \( \pi_i > \hat{\pi}_i \). This implies that \( JH(t, s, \pi) \leq H(\pi) \) on this region.

**Part II:** Next, let \( T(\pi, \hat{\pi}_i) \) be the hitting time of the deterministic path \( x_i(t, \pi) \) to the level \( \hat{\pi}_i \).

Below we show that there exists \( \pi_i^* \) such that

\[
F_k(t, \pi) \leq \mathbb{E}^\pi \left[ \int_0^{T(\hat{\pi}_i)} e^{-\rho t} c dt + e^{-\rho t - \gamma} \overline{\pi} \cdot \lambda \right] \leq \overline{\pi} \pi_i^* + m(1 - \pi_i^*) \leq H(\pi)
\]

for all \( k \in A \) (not just \( A^* \)) and for all \( t \geq T(\pi, \hat{\pi}_i) \) on the region \( \{ \pi \in D; \pi_i \geq \pi_i^* \} \). This will further imply that \( JH(t, s, \pi) \leq H(\pi) \) for all \( t \geq 0 \) for a point \( \pi \) falling on the latter region, and we will have \( H(\pi) \leq J_0 H(s, \pi) = \sup_{\tau \in [0, \infty]} JH(t, s, \pi) \leq H(\pi) \).

Note that the first inequality in (C.16) follows from \( C(\cdot) \leq c \) and \( H(\cdot) \leq \overline{\mu} \). For a given value \( \pi_i^* \) the last inequality is true for all the points on \( \{ \pi \in D; \pi_i \geq \pi_i^* \} \) since \( H(\pi) = \sup_{k \in A^*} H_k(\pi) = \overline{\pi} \pi_i^* + \sup_{k \in A^*} \sum_{i \neq i} \mu_{k,i} \pi_i \geq \overline{\pi} \pi_i^* + m(1 - \pi_i^*) \geq \overline{\pi} \pi_i^* + m(1 - \pi_i^*) \). Hence, it remains to show that the second inequality holds for some \( \pi_i^* \).

For \( \pi_i > \hat{\pi}_i \) we have \( \hat{\pi}_i = \pi_i + \int_0^{T(\hat{\pi}_i)} \frac{d(x_i(t, \pi))}{dt} \) \( dt \). Then, thanks to (2.14) we get \( 0 \geq \hat{\pi}_i - \pi_i = \int_0^{T(\hat{\pi}_i)} \left( \sum_{j \in E} q_{ji} x_j(t, \pi) - \lambda_i x_i(t, \pi) + x_i(t, \pi) \sum \lambda_j x_j(t, \pi) \right) dt \geq \int_0^{T(\hat{\pi}_i)} (q_{ii} - \lambda_i) dt \)

\[
= (q_{ii} - \lambda_i) \cdot T(\pi, \hat{\pi}_i),
\]

which further implies

\[
T(\pi, \hat{\pi}_i) \geq (\pi_i - \hat{\pi}_i)/(q_{ii} - \lambda_i).
\]
Case I: $\rho > 0$. By (C.17) we get the inequality $E^\pi \exp \left( -\rho \cdot T(\widehat{\pi}, \widehat{\pi}_i) \wedge \sigma_1 \right) \leq$

$$E^\pi \exp \left( -\rho \left[ \frac{\pi_i - \widehat{\pi}_i}{-q_{ii} + \lambda_i} \wedge \sigma_1 \right] \right) = \int_0^\infty \exp \left( -\rho \left[ \frac{\pi_i - \widehat{\pi}_i}{-q_{ii} + \lambda_i} \wedge u \right] \right) \sum_{k \in E} E^\pi \left[ 1_{(M_n = i)} e^{-\gamma(u)} \lambda_k \right] du.$$

The last expression above is strictly decreasing in $\pi_i$ and equals 1 at $\pi_i = \widehat{\pi}_i$. Moreover, $\pi_i \rightarrow \pi \pi_i + \mu(1 - \pi_i)$ is increasing and equals $\pi$ at $\pi_i = 1$. Therefore, there exists a unique $\pi_i^* \in [\widehat{\pi}_i, 1)$ defined as

(C.18) $\pi_i^* = \inf \left\{ \pi_i \geq \widehat{\pi}_i : \pi \in \mathbb{E}^\pi \exp \left( -\rho \left[ \frac{\pi_i - \widehat{\pi}_i}{-q_{ii} + \lambda_i} \wedge \sigma_1 \right] \right) \leq \pi \pi_i + \mu(1 - \pi_i) \right\} < 1,$

such that the inequality in (C.18) holds for all $\pi_i \in [\pi_i^*, 1]$. The definition of $\pi_i^*$ implies that for all the points $\pi$ with $\pi_i \geq \pi_i^*$ and for $t \geq T(\pi, \pi)$ we have $E^\pi \left[ e^{-\rho T(\pi, \pi) \wedge \sigma_1} \right] \leq \pi \mathbb{E}^\pi \left[ e^{-\rho T(\pi, \pi) \wedge \sigma_1} \right] \leq \pi \mathbb{E}^\pi \exp \left( -\rho \left[ \frac{\pi_i - \widehat{\pi}_i}{-q_{ii} + \lambda_i} \wedge \sigma_1 \right] \right) \leq \pi \pi_i + \mu(1 - \pi_i) \leq H(\pi)$.

This establishes (C.16) and concludes the proof when $\rho > 0$.

Case II: $\rho > 0$. If $\rho > 0$, arguments given for Case I still holds. Hence we assume that $\rho = 0$. Using (C.17) again, we obtain

$$E^\pi \left[ T(\pi, \pi)_i \wedge \sigma_1 \right] \geq E^\pi \left[ \frac{\pi_i - \widehat{\pi}_i}{-q_{ii} + \lambda_i} \wedge \sigma_1 \right] = \int_0^\infty \left[ \frac{\pi_i - \widehat{\pi}_i}{-q_{ii} + \lambda_i} \wedge u \right] \sum_{j \in E} \lambda_j m_j(u, \pi) du.$$

The last expression above equals to 0 at $\pi_i = \widehat{\pi}_i$ and it is strictly increasing in $\pi_i$ for $\pi_i \geq \widehat{\pi}_i$. Hence, there exists a unique point $\pi_i^* \triangleq \inf \left\{ \pi_i \geq \widehat{\pi}_i : c E^\pi \left[ \pi_i - \widehat{\pi}_i \wedge \sigma_1 \right] + \pi = \pi \pi_i + \mu(1 - \pi_i) \right\} < 1.$

Then, for the points $\pi$ with $\pi_i \geq \pi_i^*$ and for $t \geq T(\pi, \pi)$ we have $E^\pi \left[ e^{-\rho T(\pi, \pi) \wedge \sigma_1} \right] \leq \pi \mathbb{E}^\pi \left[ e^{-\rho T(\pi, \pi) \wedge \sigma_1} \right] \leq \pi \mathbb{E}^\pi \exp \left( -\rho \left[ \frac{\pi_i - \widehat{\pi}_i}{-q_{ii} + \lambda_i} \wedge \sigma_1 \right] \right) \leq \pi \pi_i + \mu(1 - \pi_i) \leq H(\pi)$,

and this concludes the proof.

Proof of Proposition 5.1. The first claim on immediate stopping if $\mu_{2,1,1.2}(\lambda_2 - \lambda_1) \leq \mu_{2,1} + \mu_{1,2}$ is an immediate corollary of Peskir and Shiryaev [2000, Theorem 2.1].

Let us now assume that $\mu_{2,1,1.2}(\lambda_2 - \lambda_1) > \mu_{2,1} + \mu_{1,2}$. For the problem with two hypotheses, we have $H(\pi) = \min \{ \mu_{1,2} \pi_2 ; \mu_{2,1} \pi_1 \}$, and recall that $v_1(T, \pi) = \inf_{t \in [0, \delta]} JH(t, \pi)$. For $\pi = (\pi_1, \pi_2)$ with $\pi_2 \in (\lambda_1 \mu_{2,1} / (\lambda_2 \mu_{2,1} + \lambda_1 \mu_{2,1}), \mu_{2,1} / (\mu_{2,1} + \mu_{1,2}))$ and for small $t > 0$, evaluating the expression $JH(t, \pi)$ gives

$$\left[ \pi_1 e^{-\lambda_1 t} + \pi_2 e^{-\lambda_2 t} \right] \mu_{1,2} x_2(t, \pi) + \int_0^t \frac{2}{\lambda_1} \pi_1 e^{-\lambda_1 u} \left( 1 + \lambda_j \left( \mu_{2,1} \frac{\lambda_1 x_1(u, \pi)}{\lambda_1 x_1(u, \pi) + \lambda_2 x_2(u, \pi)} \right) \right) du,$$

and using the dynamics of $t \mapsto \lambda \pi(t, \pi)$ in (2.14) we obtain

(C.19) $\frac{d J(t, \pi)}{dt} = \left[ 1 + \mu_{2,1} \lambda_1 \right] \pi_1 e^{-\lambda_1 t} + \left[ 1 - \mu_{1,2} \lambda_2 \right] \pi_2 e^{-\lambda_2 t} \left( \frac{1}{\mu_{2,1} + \mu_{1,2}} \left[ \mu_{2,1} + \mu_{1,2} + \mu_{2,1} \mu_{1,2} (\lambda_1 - \lambda_2) \right] + \delta (\mu_{2,1} \lambda_1 + \mu_{1,2} \lambda_2) \right).$

when $t = 0$ and $\pi = (\mu_{2,1} / (\mu_{2,1} + \mu_{1,2}) + \delta, \mu_{2,1} / (\mu_{2,1} + \mu_{1,2}) - \delta)$ for $\delta > 0$ small. Under the assumption $\mu_{2,1} \mu_{1,2} (\lambda_2 - \lambda_1) > \mu_{2,1} + \mu_{1,2}$, the last expression is negative for $\delta$ sufficiently small.
This implies that $v_1(T, \pi) < H(\pi)$ for small values of $T > 0$ at points $\pi$, for which $\pi_2 = \mu_{2,1}/(\mu_{2,1} + \mu_{1,2}) - \delta$ where

$$\delta < \frac{\mu_{2,1}\mu_{1,2}(\lambda_2 - \lambda_1) - \mu_{2,1} - \mu_{1,2}}{(\mu_{2,1} + \mu_{1,2})(\mu_{2,1}\lambda_1 + \mu_{1,2}\lambda_2)}.$$  

Since $b_1(0) = \mu_{2,1}/(\mu_{2,1} + \mu_{1,2})$, it follows that the boundary curve $T \mapsto b_1(T)$ is discontinuous at $T = 0$ (see the lower curve in Figure 4). The expression in (C.19) with $t = 0$ indicates that $dJH(t, \pi)/dt|_{t=0}$ is decreasing in $\pi_2$ and vanishes at the point $\pi$ with

$$\pi_2 = \frac{1 + \mu_{2,1}\lambda_1}{\mu_{2,1}\lambda_1 + \mu_{1,2}\lambda_2}$$

where the inequality is due to the assumption $\mu_{2,1}\mu_{1,2}(\lambda_2 - \lambda_1) > \mu_{2,1} + \mu_{1,2}$. This implies that

$$\{ (T, \pi) : \pi_2 \leq \frac{\mu_{2,1}}{\mu_{2,1} + \mu_{1,2}} \text{ and } V_1(T, \pi) = H(\pi) \} \subseteq \{ (T, \pi) : \pi_2 \leq \frac{1 + \mu_{2,1}\lambda_1}{\mu_{2,1}\lambda_1 + \mu_{1,2}\lambda_2} \}.$$ 

At the point $\pi$ with $\pi_2 = (1 + \mu_{2,1}\lambda_1)/(\mu_{2,1}\lambda_1 + \mu_{1,2}\lambda_2)$ the expression for $dJH(t, \pi)/dt$ in (C.19) is strictly positive for small $t > 0$. Since the region $\{ \pi \in D : V(T, \pi) = H(\pi) \}$ is convex for each $T$ (see Remark 4.2) and we have $v_1(T, \pi) = H(\pi)$, for all $T > 0$ at $\pi = (1, 0)$, this implies that $\exists u > 0$ such that

$$v_1(T, \pi) = H(\pi) \text{ on } \{ (T, \pi) : T \in [0, u] \text{ and } \pi_2 \leq \frac{1 + \mu_{2,1}\lambda_1}{\mu_{2,1}\lambda_1 + \mu_{1,2}\lambda_2} \}.$$ 

Recall that the deterministic part $t \mapsto \overline{x}(t, \pi)$ drifts towards the point $(1, 0)$. Then, by induction we conclude that $v_n(T, \pi) = H(\pi)$ for all $n \in \mathbb{N}$, which implies that $\lim_{n \to \infty} v_n(T, \pi) = V(T, \pi) = H(\pi)$ on the same region. As a result, we see that if the solution of the problem is not trivial, the lower boundary curve $b_1(T)$ is discontinuous at $T = 0$, and there is an initial region over which the curve stays flat at level $\pi_2 = (1 + \mu_{2,1}\lambda_1)/(\mu_{2,1}\lambda_1 + \mu_{1,2}\lambda_2)$ as in Figure 4.  

\[ \square \]

\textbf{References}


