Hamilton-Jacobi-Bellman inequality for the average control of piecewise deterministic Markov processes

O. L. V. Costa & F. Dufour

To cite this article: O. L. V. Costa & F. Dufour (2019) Hamilton-Jacobi-Bellman inequality for the average control of piecewise deterministic Markov processes, Stochastics, 91:6, 817-835, DOI: 10.1080/17442508.2018.1546305

To link to this article: https://doi.org/10.1080/17442508.2018.1546305

Published online: 13 Nov 2018.

Submit your article to this journal

Article views: 86

View related articles

View Crossmark data
Hamilton-Jacobi-Bellman inequality for the average control of piecewise deterministic Markov processes

O. L. V. Costa and F. Dufour

ABSTRACT
The main goal of this paper is to study the infinite-horizon long run average continuous-time optimal control problem of piecewise deterministic Markov processes (PDMPs) with the control acting continuously on the jump intensity $\lambda$ and on the transition measure $Q$ of the process. We provide conditions for the existence of a solution to an integro-differential optimality inequality, the so called Hamilton-Jacobi-Bellman (HJB) equation, and for the existence of a deterministic stationary optimal policy. These results are obtained by using the so-called vanishing discount approach, under some continuity and compactness assumptions on the parameters of the problem, as well as some non-explosive conditions for the process.

1. Introduction
In [8,10] it was introduced a general family of nondiffusion stochastic models, namely, piecewise deterministic Markov processes (PDMPs). These models have found applications in many optimization problems in queuing and inventory systems, maintenance-replacement models, and many other areas of engineering and operations research. Three local parameters characterize the PDMPs, the flow $\phi$, the jump rate $\lambda$, and the transition measure $Q$. Roughly speaking, the motion of PDMPs goes as follows. Considering $x_0$ the initial state, the motion of the process follows the flow $\phi(x_0, t)$ until the first jump time $T_1$, which occurs either spontaneously in a Poisson-like fashion with rate $\lambda$ or when the flow $\phi(x_0, t)$ hits the boundary of the state space. The post-jump location of the process, in either case, is selected by the transition measure $Q(\cdot | \phi(x, T_1))$ and the motion restarts from this new point afresh. As presented in [10], a suitable choice of the state space and the local characteristics $\phi$, $\lambda$, and $Q$ can cover a great deal of problems in operations research, engineering systems and management science.

A common approach to tackle the discounted and long run average control problems of PDMPs is to characterize the value function as a solution to the so called Hamilton-Jacobi-Bellman (HJB) equation associated with an imbedded discrete-stage Markov decision model, with the stages defined by the jump times $T_n$ of the process. In this case the decision
is to find, at each stage, a control function that solves an imbedded deterministic optimal control problem. Usually the control strategy is chosen among the set of piecewise open loop policies, that is, stochastic kernels or measurable functions that depend only on the last jump time and post jump location. We can cite [2,3,5,6,9,10,14,26,27] as a sample of works following this technique. Another important approach for these problems, which we will call the infinitesimal approach, is to characterize the optimal value function as the viscosity solution of the corresponding integro-differential HJB equation. As a sample of works using this kind of approach we can mention [7,10–13,28].

In this paper we adopt the infinitesimal approach to study the long run average control problem of PDMPs. We provide conditions for the existence of a solution to an integro-differential HJB inequality, and for the existence of a deterministic stationary optimal policy. These conditions are essentially related to continuity and compactness assumptions on the parameters of the problem, as well as some non-explosive conditions for the controlled process. To derive these results we apply the so-called vanishing discounted approach by adapting and combining arguments used in the context of continuous-time Markov decision processes (see [29]), and results from the authors obtained for the infinite-horizon discounted optimal control problem obtained in [7]. As far as the authors are aware of, this is the first time that this kind of result is presented in the literature for long run average control problems of PDMPs considering the broader class of controls mentioned above. It should be stressed that, when compared with the PDMPs long run average control literature, we consider a broader class of control strategies (possibly depending on the past-history of the process and taking values in state-dependent action spaces) instead of the open loop policies with fixed action set considered in previous papers. It is also important to point out that, differently from the so-called continuous-time Markov decision processes (see, for instance, [15–18,23,29,30]), the PDMPs are characterized by a drift motion between jumps, and forced jumps whenever the process hits the boundary, so that the available results for the continuous-time Markov decision processes cannot be applied to our case.

The paper is organized as follows. In Section 2 we present the notation, the parameters defining the model, the construction of the controlled process, the definition of the admissible strategies, and the problem formulation. In Section 3 we give the main assumptions and some auxiliary results. In Section 4 we obtain the main results of our paper (see Theorems 4.1 and 4.2) that provide sufficient conditions for the existence of a solution to a HJB inequality, and for the existence of a deterministic stationary optimal policy.

2. Problem formulation for the controlled PDMP

In this section we introduce the notation, the parameters defining the model, the construction of the controlled process, the definition of the admissible strategies, and the problem formulation. It follows closely Sections 2 and 3 in [7] and, due to that, some details will be skipped.

2.1. Notation

We will denote by \( \mathbb{N} \) the set of natural numbers including 0, \( \mathbb{N}^* = \mathbb{N} \setminus \{0\} \), \( \mathbb{R} \) the set of real numbers, \( \mathbb{R}_+^* = \mathbb{R}_+ \setminus \{0\} \), \( \hat{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \).
For $x \in \mathbb{R}$, $|x|$ denotes the greatest integer that is less than or equal to $x$. The term *measure* will always refer to a countably additive, $\mathbb{R}_+$-valued set function. For $X$ a Borel space (i.e. a Borel-measurable subset of a complete and separable metric space) we denote by $\mathcal{B}(X)$ its associated Borel $\sigma$-algebra, and by $\mathcal{M}(X)$ ($\mathcal{P}(X)$ respectively) the set of measures (probability measures) defined on $(X, \mathcal{B}(X))$, endowed with the weak topology. We represent by $\mathcal{P}(X \mid Y)$ the set of stochastic kernels on $X$ given $Y$ where $Y$ denotes a Borel space. For any set $A$, $I_A$ denotes the indicator function of the set $A$, and for any point $x \in X$, $\delta_x$ denotes the Dirac measure defined by $\delta_x(\Gamma) = 1_{\Gamma}(x)$ for any $\Gamma \in \mathcal{B}(X)$. The space of Borel-measurable (bounded, lower semicontinuous respectively) real-valued functions defined on the Borel space $X$ will be denoted by $\mathcal{M}(X)$ ($\mathbb{B}(X)$, $\mathcal{L}(X)$ respectively) and we set $\mathcal{L}_b(X) = \mathcal{L}(X) \cap \mathbb{B}(X)$. Moreover, the space of Borel-measurable, lower semicontinuous, $\mathbb{R}$-valued functions defined on the Borel space $X$ will be denoted by $\hat{\mathcal{L}}(X)$. In all the previous cases we introduce the subscript $+$ for the case of non-negative functions. Finally, the infimum over an empty set is understood to be equal to $+\infty$, and we set $e^{-\infty} = 0$.

### 2.2. Parameters of the model

The control model we will consider will depend on the following elements:

- The state space $X$, which we assume to be an open subset of $\mathbb{R}^d$ ($d \in \mathbb{N}^*$) with boundary represented by $\partial X$.
- The flow $\phi(x, t) : \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}^d$, associated with a given Lipschitz continuous vector field in $\mathbb{R}^d$, that is, $\phi(x, 0) = x$ and $\phi(x, t + s) = \phi(\phi(x, s), t)$ for all $x \in \mathbb{R}^d$ and $(t, s) \in \mathbb{R}^2$.
- The so-called active boundary defined as $\Xi = \{x \in \partial X : x = \phi(y, t) \text{ for some } y \in X \text{ and } t \in \mathbb{R}_+\}$. With some abuse of notation, we set $\overline{X}$ as $X \cup \Xi$, and for $x \in \overline{X}$, we define $t^*(x) = \inf\{t \in \mathbb{R}_+ : \phi(x, t) \in \Xi\}$. Since the flow $\phi$ outside the space $\overline{X}$ plays no role, it can be defined arbitrarily.
- The action space $A$, assumed to be a Borel space, and defined as $A = A^i \cup A^g$ with $A^i \cap A^g = \emptyset$, and $A^i \in \mathcal{B}(A)$ (respectively $A^g \in \mathcal{B}(A)$) representing the set of *impulsive* (respectively *gradual*) actions, assumed to be nonempty.
- The set of feasible actions in state $x \in \overline{X}$ is $A(x)$, which is a nonempty measurable subset of $A$, satisfying $A(x) \subset A^g$ for all $x \in X$ and $A(x) \subset A^i$ for all $x \in \Xi$. Define the set $K = K^i \cup K^g$ with

$$
K^g = \{(x, a) \in X \times A : a \in A(x)\} \in \mathcal{B}(X \times A^g),
$$

$$
K^i = \{(x, a) \in \Xi \times A : a \in A(x)\} \in \mathcal{B}(\Xi \times A^i).
$$

It is assumed that $K^g$ (respectively, $K^i$) contains the graph of a measurable function from $X$ (respectively, $\Xi$) to $A$.
- The transition rate (infinitesimal generator) $q$ is a signed kernel on $X$ given $K^g$. This means that $\Gamma \mapsto q(\Gamma \mid x, a)$ is a signed measure on $(X, \mathcal{B}(X))$ for all $(x, a) \in K^g$, and that $(x, a) \mapsto q(\Gamma \mid x, a)$ is measurable for all $\Gamma \in \mathcal{B}(X)$. It satisfies $q(\Gamma \mid x, a) \geq 0$ for all $\Gamma \in \mathcal{B}(X)$ such that $x \notin \Gamma$, and it is conservative, i.e. $q(X \mid x, a) = 0$ and stable in that $\sup_{a \in A(x)} \lambda(x, a) < \infty$, where

$$
\lambda(x, a) = -q(\{x\} \mid x, a) = q(X \setminus \{x\} \mid x, a).
$$
The stochastic kernel \( Q \in \mathcal{P}(X \mid K^i) \), meaning that for any \((z, b) \in K^i\), \( Q(\cdot \mid z, b) \) is the distribution of the state immediately after the jump from the boundary when an impulsive action \( b \in A(z) \) is applied. We call such jumps as ‘forced’ jumps, while the jumps governed by the generator \( q \) are called ‘natural’ jumps.

Note that \( \lambda(x, a) \), as previously introduced, is the natural jumps’ intensity. For an arbitrary \( \Gamma \in \mathcal{B}(X) \) and \( x \notin \Gamma \), the ratio defined by \( Q(\Gamma \mid x, a) = q(\Gamma \mid x, a) / \lambda(x, a) \) whenever \( \lambda(x, a) > 0 \) gives the probability that the state belongs to \( \Gamma \) immediately after a natural jump. In case \( \lambda(x, a) = 0 \), \( Q \) can be defined in an arbitrary measurable way. By doing this we obtain a stochastic kernel \( Q \in \mathcal{P}(X \mid K^i) \) satisfying \( Q(X \setminus \{x\} \mid x, a) = 1 \) for any \((x, a) \in K^i\). Clearly,

\[
q(dy \mid x, a) = \lambda(x, a) [ Q(dy \mid x, a) - \delta_x(dy) ].
\]

Combining with the points \((x, a) \in K^i\) and the previously defined stochastic kernel \( Q \in \mathcal{P}(X \mid K^i) \), we obtain the stochastic kernel \( Q \in \mathcal{P}(X \mid K) \).

We conclude this section with the following definitions that will be used in the sequel.

**Definition 2.1:** We define \( \mathcal{P}^a \) (respectively, \( \mathcal{P}^i \)) as the set of stochastic kernels \( \pi \in \mathcal{P}(A^a \mid X) \) (respectively \( \gamma \in \mathcal{P}(A^i \mid X) \)) satisfying \( \pi(A(x) \mid x) = 1 \) for any \( x \in X \) (respectively, \( \gamma(A(z) \mid z) = 1 \) for any \( z \in \mathcal{E} \)).

**Definition 2.2:** We define \( A(X) \) as the set of functions \( g \in \mathcal{M}(X) \) which are absolutely continuous with respect to the flow \( \phi \) on \([0, t^*(x)]\) (that is, the function \( g(\phi(x, \cdot)) \) is absolutely continuous on \([0, t^*(x)] \cap \mathbb{R}_+ \)) and such that \( \lim_{t \to t^*(x)} g(\phi(x, t)) \) exists whenever \( t^*(x) < \infty \). In this case the domain of definition of the mapping \( g \) can be extended to \( \bar{X} \) by setting \( g(z) = \lim_{t \to t^*(x)} g(\phi(x, t)) \) where \( z = \phi(x, t^*(x)) \in \mathcal{E} \). For \( g \in A(X) \) we have, from Lemma 2.2 in [6], that there exists a real-valued measurable function \( \mathcal{X}g \) defined on \( X \) satisfying

\[
g(\phi(x, t)) = g(x) + \int_{[0,t]} \mathcal{X}g(\phi(x, s)) \, ds,
\]

for any \( t \in [0, t^*(x)] \). Notice that for \( g \in A(X) \) the function \( \mathcal{X}g \) satisfying (2) is not necessarily unique. As before, the case of bounded functions in \( A(X) \) will be denoted by \( A_{ab}(X) \).

### 2.3. Construction of the controlled process \( \xi_t \)

We define the canonical space \( \Omega = \bigcup_{n=0}^{\infty} \Omega_n \bigcup (X \times (\mathbb{R}_+^n \times X)^\infty) \) where \( \Omega_n = X \times (\mathbb{R}_+^n \times X)^n \times (\{\infty\} \times \{x_\infty\})^\infty \) and \( x_\infty \) is an isolated artificial point corresponding to the case when no jumps occur in the future, endowed with its Borel \( \sigma \)-algebra denoted by \( \mathcal{F} \). We set \( \omega \in \Omega \) as

\[
\omega = (x_0, \theta_1, x_1, \theta_2, x_2, \ldots),
\]

where \( x_0 \in X \) represents the initial state of the controlled point process \( \xi \), and for \( n \in \mathbb{N}^* \), the components \( \theta_n > 0 \) and \( x_n \) correspond to the time interval between two consecutive jumps and the value of the process \( \xi \) immediately after the jump. For the case \( \theta_n < \infty \)
and \( \theta_{n+1} = \infty \), the trajectory of the controlled point process has only \( n \) jumps, and we put \( \theta_n = \infty \) and \( x_n = x_\infty \) (artificial point) for all \( m \geq n + 1 \). Between jumps, the state of the process \( \xi \) moves according to the flow \( \phi \). The path up to \( n \in \mathbb{N} \) is denoted by \( h_n = (x_0, \theta_1, x_1, \theta_2, \ldots, \theta_n, x_n) \), and the collection of all such paths is denoted by \( \mathbf{H}_n \). We denote by \( H_n = (X_0, \Theta_1, X_1, \ldots, \Theta_n, X_n) \) the \( n \)-term random history process taking values in \( \mathbf{H}_n \) for \( n \in \mathbb{N} \). For \( n \in \mathbb{N} \), introduce the mappings \( X_n : \Omega \to \mathbb{X}_\infty \) by \( X_n(\omega) = x_n \), where \( \mathbb{X}_\infty = \mathbb{X} \cup \{ x_\infty \} \) and, for \( n \geq 1 \), the mappings \( \Theta_n : \Omega \to \mathbb{R}_+^n \) by \( \Theta_n(\omega) = \theta_n \); \( \Theta_0(\omega) = 0 \). The sequence \( (T_n)_{n \in \mathbb{N}}^n \) of \( \mathbb{R}_+^n \)-valued mappings is defined on \( \Omega \) by \( T_n(\omega) = \sum_{i=1}^n \Theta_i(\omega) = \sum_{i=1}^n \theta_i \) and \( T_\infty(\omega) = \lim_{n \to \infty} T_n(\omega) \). The random measure \( \mu \) associated with \( (\Theta_n, X_n)_{n \in \mathbb{N}} \) is a measure defined on \( \mathbb{R}_+^n \times \mathbb{X} \) by

\[
\mu(\omega; dt, dx) = \sum_{n \geq 1} I_{[T_n(\omega) < \infty]} \delta(T_n(\omega), X_n(\omega))(dt, dx).
\]

For notational convenience the dependence on \( \omega \) will be suppressed and, instead of \( \mu(\omega; dt, dx) \), it will be written \( \mu(dt, dx) \). Following the notation as in [10], we set

\[
p^*(dt) = I_{[\xi_t \in \mathbb{X}]} \mu(dt, X)
\]

so that \( p^*(t) \) counts the number of jumps from the boundary of the controlled process \( \xi_t \) (see [10], sub-section 26). For \( t \in \mathbb{R}_+ \), define \( \mathcal{F}_t = \sigma[H_0] \vee \sigma[\mu[0,s] \times B] : s \leq t, B \in B(\mathbb{X}) \). Finally, we define the controlled process \( \{ \xi_t \}_{t \in \mathbb{R}_+} \) as:

\[
\xi_t(\omega) = \begin{cases} 
\phi(X_n, t - T_n) & \text{if } T_n \leq t < T_{n+1} \text{ for } n \in \mathbb{N}; \\
x_\infty, & \text{if } T_\infty \leq t.
\end{cases}
\]

Clearly \( \{ \xi_t \}_{t \in \mathbb{R}_+} \) can be equivalently described by the sequence \( (\Theta_n, X_n)_{n \in \mathbb{N}} \).

### 2.4. Admissible strategies

An admissible control strategy is a sequence \( u = (\pi_n, \gamma_n)_{n \in \mathbb{N}} \) such that, for any \( n \in \mathbb{N} \),

- \( \pi_n \in \mathcal{P}(\mathcal{A}^\delta | \mathbb{H}_n \times \mathbb{R}_+^n) \) and satisfies, for \( h_n = (x_0, \theta_1, x_1, \ldots, \theta_n, x_n) \in \mathbb{H}_n, \pi_n(\mathcal{A}(\phi(x_n, t) | h_n, t)) = 1 \) for any \( t \in ]0, t^*(x_n) [ \) in the case \( x_n \neq x_\infty \), and \( \pi_n(\cdot | h_n, t) \) is an arbitrary stochastic kernel on \( \mathcal{P}(\mathcal{A}^\delta | \mathbb{H}_n \times \mathbb{R}_+^n) \) in the case \( x_n = x_\infty \).
- \( \gamma_n \in \mathcal{P}(\mathcal{A}^I | \mathbb{H}_n) \) and satisfies, for \( h_n = (x_0, \theta_1, x_1, \ldots, \theta_n, x_n) \in \mathbb{H}_n \) and \( t^*(x_n) < \infty \) with \( x_n \neq x_\infty, \gamma_n(\mathcal{A}(\phi(x_n, t^*(x_n))) | h_n) = 1 \) and, otherwise, \( \gamma_n(\cdot | h_n) \) is an arbitrarily fixed kernel in \( \mathcal{P}(\mathcal{A}^I | \mathbb{H}_n) \).

The set of admissible control strategies is denoted by \( \mathcal{U} \). When an admissible control strategy \( u = (\pi_n, \gamma_n)_{n \in \mathbb{N}} \) is considered we denote by \( \pi \) and \( \gamma \) the random processes with values in \( \mathcal{P}(\mathcal{A}^\delta) \) and \( \mathcal{P}(\mathcal{A}^I) \) correspondingly as

\[
\pi(da | t) = \sum_{n \in \mathbb{N}} I_{[T_n < t \leq T_{n+1}]} \pi_n(da | H_n, t - T_n)
\]

and

\[
\gamma(da | t) = \sum_{n \in \mathbb{N}} I_{[T_n < t \leq T_{n+1}]} \gamma_n(da | H_n),
\]
for $t \in \mathbb{R}_+^*$. The processes $\pi$ and $\gamma$ are $\{\mathcal{F}_t\}_{t \in \mathbb{R}_+^*}$-predictable random processes with values in $\mathcal{P}(\mathcal{A}^s)$ and $\mathcal{P}(\mathcal{A}^j)$ correspondingly. The following classes of admissible strategies will be considered along the paper. A control strategy $u \in \mathcal{U}$ is called deterministic stationary, if

$$\pi_n(\cdot \mid h_n, t) = \delta_{\varphi^s(\phi(x_n^0, t))}(\cdot)$$

and

$$\gamma_n(\cdot \mid h_n) = \delta_{\varphi^j(\phi(x_n^0, t^*(x_n)))(\cdot)},$$

where $\varphi^s : \overline{X} \to \mathcal{A}$ is a measurable mapping satisfying $\varphi^s(y) \in \mathcal{A}(y)$ for any $y \in \overline{X}$. By a slight abuse of notation, such strategy will be just denoted by $u = \varphi^s$.

Consider an admissible strategy $u \in \mathcal{U}$ and an initial state $x_0 \in X$. From Theorem 3.6 in [21] (or Remark 3.43, page 87 in [21]), there exists a probability $\mathbb{P}^u_{x_0}$ on $(\Omega, \mathcal{F})$ such that the restriction of $\mathbb{P}^u_{x_0}$ to $(\Omega, \mathcal{F}_0)$ is given by (see [7] for further details) $\mathbb{P}^u_{x_0}(\{X_0 = x_0\}) = 1$, and (see Lemma 3.1 in [7]) the predictable projection of the random measure $\mu$ with respect to $\mathbb{P}^u_{x_0}$ is given by $\nu = \nu_0 + \nu_1$, where, for $\Gamma \in \mathcal{B}(\mathbb{R}_+^* \times X)$,

$$\nu_0(\Gamma) = \int_G \int_{\mathcal{A}(\xi_x)} Q(dx \mid \xi_x, a) \lambda(\xi_x, a) \tau(da \mid s) ds,$n

$$\nu_1(\Gamma) = \int_G \sum_{n \in \mathbb{N}^*} I_{[x_n \in \Gamma]} \int_{\mathcal{A}(\xi_{T_n})} Q(dx \mid \xi_{T_n}, a) \gamma(da \mid T_n^-) \delta_T(ds).$$

### 2.5. Problems formulation

We present in this subsection the infinite-horizon long run average continuous-time optimal control problem we will consider in this paper, with the control acting continuously on the jump intensity $\lambda$ and on the transition measure $Q$ of the process (but not on the deterministic flow $\phi$). In order to obtain our results we will use the so-called vanishing discount approach, so that we also need to introduce the infinite-horizon discounted optimal control problem, as analysed in [7]. In what follows the cost rate $C^s$ associated with a gradual action is a real-valued measurable mapping defined on $\mathcal{K}^s$ and the cost $C^j$ associated with an impulsive action on the boundary $\mathcal{K}$ is a real-valued measurable mapping defined on $\mathcal{K}^j$. The associated long run average criteria corresponding to an admissible control strategy $u \in \mathcal{U}$ is defined by

$$\mathcal{A}(u, x_0) = \lim_{t \to \infty} \frac{1}{t} \left\{ \mathbb{E}^u_{x_0} \left[ \int_{[0, t]} \int_{\mathcal{A}(\xi_x)} C^s(\xi_x, a) \tau(da \mid s) ds \right] + \mathbb{E}^u_{x_0} \left[ \int_{[0, t]} \int_{\mathcal{A}(\xi_x)} C^j(\xi_x, a) \gamma(da \mid s) p^s(ds) \right] \right\}.$$ 

Similarly, the associated infinite-horizon discounted criteria corresponding to an admissible control strategy $u \in \mathcal{U}$ is defined by

$$\mathcal{V}_u(u, x_0) = \mathbb{E}^u_{x_0} \left[ \int_{[0, +\infty]} e^{-\alpha s} \int_{\mathcal{A}(\xi_x)} C^s(\xi_x, a) \tau(da \mid s) ds \right] + \mathbb{E}^u_{x_0} \left[ \int_{[0, +\infty]} e^{-\alpha s} \int_{\mathcal{A}(\xi_x)} C^j(\xi_x, a) \gamma(da \mid s) p^s(ds) \right].$$

In the previous expression, $\alpha > 0$ is the discount factor. In the next section, after introducing our main assumptions, we show in Remark 3.2 that, for any control strategy $u \in \mathcal{U}$, the functions $\mathcal{V}_u(u, \cdot)$ and $\mathcal{A}(u, \cdot)$ are well defined.
**Definition 2.3:** The optimization problems consist in minimizing the performance criterion $V_\alpha(u, x_0)$ and $A(u, x_0)$ within the class of admissible strategies $u \in \mathcal{U}$, where $x_0$ is the initial state. The optimal value functions will be denoted respectively by $V^*_\alpha(x_0)$ and $A^*(x_0)$, and $u \in \mathcal{U}$ will be said to be an optimal strategy for the discounted (respectively, long run average) problem if $V_\alpha(u, x_0) = V^*_\alpha(x_0)$ (respectively, $A(u, x_0) = A^*(x_0)$).

### 3. Main assumptions and preliminary results

In this section we present the main assumptions we will consider in the paper as well as some auxiliary results that will be employed in the next section.

#### 3.1. Main assumptions

We introduce in this subsection the main assumptions required in this paper. Assumption A will be used, among other things, to guarantee that the process is non-explosive and more importantly, to provide an affine upper bound for the expected value of the number of boundary jumps up to a time $t \in \mathbb{R}_+$. Assumptions B and C will be mainly required to obtain the existence of an optimal selector.

**Assumption A:** There are constants $K \geq 0$ and $\varepsilon_1 > 0$ such that

(A1) For any $(x, a) \in \mathcal{K}^g$, $\lambda(x, a) \leq K$.

(A2) For any $(z, b) \in \mathcal{K}^i$, $Q(A_{\varepsilon_1} | z, b) = 1$ where

$$A_{\varepsilon_1} = \{x \in \mathcal{X} : t^x(x) > \varepsilon_1\}.$$

(A3) For any $(x, a) \in \mathcal{K}^g$, $Q(A(x) | x, a) = 1$ where

$$A(x) = \{y \in \mathcal{X} : t^y(x) \geq \min\{t^x(x), \varepsilon_1\}\}.$$

**Assumption B:** (B1) The set $A(y)$ is compact for every $y \in \overline{\mathcal{X}}$.  
(B2) The kernel $Q$ is weakly continuous (also called weak-Feller Markov kernel).  
(B3) The function $\lambda$ is continuous on $\mathcal{K}^g$.  
(B4) The flow $\phi$ is continuous on $\mathbb{R}_+ \times \mathbb{R}^p$.  
(B5) The function $t^x$ is continuous on $\overline{\mathcal{X}}$.

**Assumption C:** (C1) The multifunction $\Psi^g$ from $\mathcal{X}$ to $\mathcal{A}$ defined by $\Psi^g(x) = A(x)$ is upper semicontinuous. The multifunction $\Psi^i$ from $\overline{\mathcal{X}}$ to $\mathcal{A}$ defined by $\Psi^i(z) = A(z)$ is upper semicontinuous.  
(C2) The cost function $C^g$ (respectively, $C^i$) is bounded and lower semicontinuous on $\mathcal{K}^g$ (respectively, $\mathcal{K}^i$).

Without loss of generality, we assume that inequalities $|C^g| \leq K$ and $|C^i| \leq K$ are valid where $K$ is the constant from (A1).
3.2. Auxiliary results

We present in this subsection some auxiliary results that will also be useful for the long-run average cost problem. Some of these results are based on the ones obtained in [7] for the infinite-horizon discounted problem.

**Lemma 3.1:** Suppose Assumption A is satisfied. Then there exist positive numbers \( M < \infty \), \( c_0 < \infty \) such that, for any control strategy \( u \in \mathcal{U} \) and for any \( x_0 \in X \)

\[
\mathbb{E}_{x_0}^{u}\left[ \sum_{n \in \mathbb{N}^{*}} e^{-\alpha T_n} \right] \leq M, \quad \mathbb{P}_{x_0}^{u}(T_{\infty} < +\infty) = 0, \tag{5}
\]

and for any \( t \in \mathbb{R}_+ \),

\[
\mathbb{P}_{x_0}^{u}\left[ \sum_{n \in \mathbb{N}^{*}} I_{\{T_n \leq t, \xi_n \in \mathcal{E}\}} \right] \leq Mt + c_0. \tag{6}
\]

**Proof:** For the proof of (5), see Lemma 4.1 in [7]. For the proof of (6), let us define \( \hat{T}_i \) as the time of the \( i^{th} \) jump of the process \( \xi_t \) from the boundary \( \mathcal{E} \), that is, \( \hat{T}_i = \inf\{t > \hat{T}_{i-1}; \xi_t \in \mathcal{E}\} \), with \( \hat{T}_0 = 0 \) and \( i \in \mathbb{N}^{*} \). Define also, for \( i \in \mathbb{N}^{*} \), \( \hat{\Theta}_i = \hat{T}_i - \hat{T}_{i-1} \) on the set \( \{\hat{T}_{i-1} < \infty\} \), otherwise \( \hat{\Theta}_i = \infty \) for \( \hat{T}_{i-1} = \infty \). Clearly we have

\[
\mathbb{E}_{x_0}^{u}\left[ \sum_{n \in \mathbb{N}^{*}} I_{\{T_n \leq \hat{T}_i, \xi_n \in \mathcal{E}\}} \right] = \mathbb{E}_{x_0}^{u}\left[ \sum_{i \in \mathbb{N}^{*}} I_{\{\hat{\Theta}_i \leq t\}} \right]. \tag{7}
\]

Moreover, for any \( t \in \mathbb{R}_+ \) and \( i \in \mathbb{N}^{*} \) we have that

\[
\{\hat{T}_i \leq t\} = \{\hat{T}_{i-1} \leq t - \epsilon_1\} \cap \{\hat{T}_i \leq t\} \cup \{\hat{T}_{i-1} > t - \epsilon_1\} \cap \{\hat{T}_i \leq t\}. \tag{8}
\]

Clearly \( \{\hat{T}_{i-1} \leq t - \epsilon_1\} \cap \{\hat{T}_i \leq t\} \subset \{\hat{T}_{i-1} \leq t - \epsilon_1\} \) and

\[
\{\hat{T}_{i-1} > t - \epsilon_1\} \cap \{\hat{T}_i \leq t\} \subset \{\hat{T}_{i-1} > t - \epsilon_1\} \cap \{\hat{\Theta}_i \leq \epsilon_1\}. \tag{9}
\]

where (9) follows from the fact that \( \omega \in \{\hat{T}_{i-1} > t - \epsilon_1\} \cap \{\hat{T}_i \leq t\} \) implies that

\[
t - \epsilon_1 + \hat{\Theta}_i(\omega) < \hat{T}_{i-1}(\omega) + \hat{\Theta}_i(\omega) = \hat{T}_i(\omega) \leq t.
\]

From (8) and (9) we have for any \( i \in \mathbb{N}^{*} \) that

\[
I_{\{\hat{T}_i \leq t\}} \leq I_{\{\hat{T}_{i-1} \leq t - \epsilon_1\}} + I_{\{\hat{T}_{i-1} > t - \epsilon_1\}} I_{\{\hat{\Theta}_i \leq \epsilon_1\}}
\]

and thus

\[
\mathbb{P}_{x_0}^{u}(I_{\{\hat{T}_i \leq t\}}) = \mathbb{P}_{x_0}^{u}(\hat{T}_i \leq t) \leq \mathbb{P}_{x_0}^{u}(\hat{T}_{i-1} \leq t - \epsilon_1) + \mathbb{P}_{x_0}^{u}(\{\hat{T}_i \leq t\} \cap \{\hat{\Theta}_i \leq \epsilon_1\}). \tag{10}
\]

We claim now that for \( i \geq 2 \) we have that \( \mathbb{P}_{x_0}^{u}(\{\hat{\Theta}_i \leq \epsilon_1\}) = 0 \). The idea is to show that with probability 1, after jumping from the frontier, the process will take a time greater than \( \epsilon_1 \) to reach again the frontier, even considering possible spontaneous jumps in between. Indeed, first we notice that, by Assumption (A2), with probability 1, the process will jump to the
set \( A_{\epsilon_1} \) after touching the boundary, so that the time for the flow \( \phi(\xi, t) \) to hit the frontier again from the post jump location \( \zeta = \xi \tau \) will be \( t_\epsilon(\zeta) > \epsilon_1 \). If we have a spontaneous jump, say, at time \( \tau < t_\epsilon(\zeta) \), and the process jumps from the point \( \phi(\xi, \tau) \) to a point \( y \), we have, from Assumption (A3), that with probability 1, \( t_\epsilon(y) \geq \min\{t_\epsilon(\phi(\xi, \tau)), \epsilon_1\} \). If \( \min\{t_\epsilon(\phi(\xi, \tau)), \epsilon_1\} = t_\epsilon(\phi(\xi, \tau)) \) then \( \tau + t_\epsilon(y) \geq \tau + t_\epsilon(\phi(\xi, \tau)) = \tau + t_\epsilon(\xi) - \tau = t_\epsilon(\zeta) > \epsilon_1 \), since \( t_\epsilon(\phi(\xi, \tau)) = t_\epsilon(x) - \tau \). Otherwise it is clear that \( \tau + t_\epsilon(y) \geq \tau + \epsilon_1 > \epsilon_1 \). By repeating this procedure and since we have a countable number of spontaneous jumps until reaching the boundary again, we have that with probability 1 the time to reach the boundary will be greater than \( \epsilon_1 \), completing the proof of the initial claim.

Thus we have, for \( i \geq 2 \), that (10) simplifies to \( \mathbb{P}_{x_0}^{u}(\hat{T}_i \leq t) \leq \mathbb{P}_{x_0}^{u}(\hat{T}_{i-1} \leq t - \epsilon_1) \), so that by recurrence we obtain that

\[
\mathbb{P}_{x_0}^{u}(\hat{T}_i \leq t) \leq \mathbb{P}_{x_0}^{u}(\hat{T}_{i-k} \leq t - k\epsilon_1),
\]

for \( k = 0, \ldots, i - 2 \). Set \( p = \lfloor t/\epsilon_1 \rfloor \). If \( i > p + 3 \) then we obtain that for \( k = i - 2 > p + 1 \) that \( t - k\epsilon_1 = \epsilon_1 (t/\epsilon_1 - k) < \epsilon_1 (t/\epsilon_1 - (p + 1)) < 0 \) since \( p \leq t/\epsilon_1 < p + 1 \). Thus from (11) we get that \( \mathbb{P}_{x_0}^{u}(\hat{T}_i \leq t) \leq \mathbb{P}_{x_0}^{u}(\hat{T}_2 \leq 0) = 0 \) for all \( i > p + 3 \). From this and (7) we get that

\[
\mathbb{P}_{x_0}^{u}\left[\sum_{n \in \mathbb{N}^*} I\{T_n \leq k, T_n - \epsilon \zeta \}ight] = \sum_{i=1}^{p+3} \mathbb{P}_{x_0}^{u}(\hat{T}_i \leq t) \leq p + 3 \leq \frac{t}{\epsilon_1} + 3
\]

showing (6), after taking an appropriate constant \( M \).

**Remark 3.2:** From Lemma 3.1 and recalling Assumption (C2) we have that

\[
|V_\alpha(u, x_0)| \leq K \left( \frac{1}{\alpha} + \mathbb{E}_{x_0}^{u}\left[\sum_{n \in \mathbb{N}^*} e^{-\alpha T_n}\right]\right) \leq K \left( \frac{1}{\alpha} + M \right)
\]

and

\[
|A(u, x_0)| \leq K \left( 1 + \lim_{t \to \infty} \frac{1}{t} \mathbb{E}_{x_0}^{u}\left[\sum_{n \in \mathbb{N}^*} I\{T_n \leq t, T_n - \epsilon \zeta \}ight]\right) \leq K(1 + M).
\]

Theorem 3.6 below presents sufficient conditions based on the three local characteristics of the process \( \phi, \lambda, Q \), and the semi-continuity properties of the set valued action space, for the existence of a solution for an integro-differential HJB optimality equation associated with the problem as well as conditions for the existence of an optimal selector. Moreover it shows that the solution of the integro-differential HJB optimality equation is in fact unique and coincides with the optimal value for the \( \alpha \)-discounted problem, and the optimal selector derived in Theorem 3.6 yields an optimal deterministic stationary strategy for the discounted control problem. Before that we present some auxiliary results in Lemmas 3.3, 3.4 and 3.5 that will also be useful for the long-run average cost problem.
Lemma 3.3: Consider a bounded from below real-valued measurable function $F$ (respectively, $G$) defined on $X$ (respectively, $\Xi$) and a real number $\beta > 0$ satisfying
\[
\int_{[0,t^*(x)]} e^{-\beta s} F(\phi(x,s)) \, ds < +\infty,
\]
for any $x \in \Xi$. Then the real-valued mapping $V$ defined on $\Xi$ by
\[
V(x) = \int_{[0,t^*(x)]} e^{-\beta s} F(\phi(x,s)) \, ds + e^{-\beta t^*(x)} G(\phi(x,t^*(x)))
\]
belongs to $A(\Xi)$. Moreover, there exists a bounded from below measurable function $XV$ satisfying
\[
-\beta V(x) + XV(x) = -F(x),
\]
for any $x \in X$. Furthermore, $V(z) = G(z)$ for any $z \in \Xi$.

Proof: The proof follows from a slight adaptation of the proof of Lemma 5.1 in [7].

Let us introduce for any $V \in M(\Xi)$ with $V$ bounded from below, the $\hat{\mathbb{R}}$-valued function $\mathcal{R}V$ defined on $X$ by
\[
\mathcal{R}V(x) = \inf_{a \in A(x)} \left\{ C^g(x,a) + qV(x,a) + KV(x) \right\},
\]
where the constant $K$ has been introduced in Assumption (A1) and the transition kernel $q$ in equation (1). We recall, from (1), and the definition of the stochastic kernel $Q$, that
\[
QV(x,a) = \int_X V(y)Q(dy|x,a), \quad (x,a) \in K,
\]
\[
qV(x,a) = \lambda(x,a) [QV(x,a) - V(x)], \quad (x,a) \in K^i.
\]
Observe that in the previous equation $qV$ is well defined but may take the value $+\infty$. Define also the $\hat{\mathbb{R}}$-valued function $\mathcal{T}V$ on $\Xi$ by
\[
\mathcal{T}V(z) = \inf_{b \in A(z)} \left\{ C^i(z,b) + QV(z,b) \right\},
\]
(again $QV$ is well defined but may take the value $+\infty$), and the $\hat{\mathbb{R}}$-valued function $\mathcal{B}_\alpha V$ defined on $X$ by
\[
\mathcal{B}_\alpha V(y) = \int_{[0,t^*(y)]} e^{-(K+\alpha)t} \mathcal{R}V(\phi(y,t)) \, dt + e^{-(K+\alpha)t^*(y)} \mathcal{T}V(\phi(y,t^*(y))),
\]
for $\alpha \in [0,1]$. Since $V$ is bounded from below the integral in (14) is well defined but may take the value $+\infty$ and, moreover, recalling that $|C^g| \leq K$ and $|C^i| \leq K$, we can find a constant $c_0 > 0$ such that $\mathcal{R}V(x) \geq -Kc_0$ and $\mathcal{T}V(z) \geq -c_0$ so that from (12)–(14) it follows that $\mathcal{B}_\alpha V(y) \geq -c_0(1 - e^{-Kt^*(y)}) - c_0 e^{-Kt^*(y)} = -c_0$ for any $\alpha \in [0,1]$. 

Lemma 3.4: Suppose Assumptions A, B and C are satisfied. If \( V \in L(\mathbb{X}) \) is bounded from below then \( \mathcal{R}V \in \hat{L}(\mathbb{X}), \mathcal{S}V \in \hat{L}(\mathcal{Z}), \mathcal{B}_\alpha V \in \hat{L}(\mathbb{X}) \) for any \( \alpha \in [0,1] \) and all functions are bounded from below.

Proof: Define \( V^n(x) = \min\{V(x), n\} \) so that \( V^n \in L_b(\mathbb{X}) \) and therefore, the results hold for \( V^n \) from Lemma 5.2 in [7], that is, \( \mathcal{R}V^n \in L_b(\mathbb{X}), \mathcal{S}V^n \in L_b(\mathcal{Z}), \) and \( \mathcal{B}_\alpha V^n \in L_b(\mathbb{X}) \). From Proposition 10.1 in [24], it follows that \( \mathcal{R}V = \lim_{n \to \infty} \mathcal{R}V^n \in \hat{L}(\mathbb{X}) \) and similarly, \( \mathcal{S}V = \lim_{n \to \infty} \mathcal{S}V^n \in \hat{L}(\mathcal{Z}) \). Now, from the monotone convergence theorem, we have \( \mathcal{B}_\alpha V = \lim_{n \to \infty} \mathcal{B}_\alpha V^n \), and so \( \mathcal{B}_\alpha V \in \hat{L}(\mathbb{X}) \). Clearly, these functions are bounded from below, giving the result.

Let us introduce the constants \( K_\alpha \) and \( K_C \) (see Assumption A) as follows:

\[
K_\alpha = \frac{K(1 + K)(1 - e^{-(K+\alpha)\varepsilon_1}) + (K + \alpha)Ke^{-(K+\alpha)\varepsilon_1}}{\alpha(1 - e^{-(K+\alpha)\varepsilon_1})},
\]

\[
K_C = \frac{2K(1 + K)}{1 - e^{-K\varepsilon_1}}.
\]

Notice that for any \( 0 < \alpha < 1 \) we have that

\[
0 < \alpha K_\alpha \leq K_C.
\]

Lemma 3.5: Suppose Assumptions A, B and C hold. Consider a function \( V \in L_b(\mathbb{X}) \) satisfying \( |V(y)| \leq K_a I_{A_1}(y) + (K_\alpha + K_1)I_{A_1} \) for any \( y \in \mathbb{X} \) then the function \( \mathcal{B}_\alpha V \in A_b(\mathbb{X}) \) and \( |\mathcal{B}_\alpha V(y)| \leq K_a I_{A_1} + (K_\alpha + K)I_{A_1} \) for any \( y \in \mathbb{X} \).

Proof: See Lemma 5.4 in [7].

The next theorem provides sufficient conditions for the existence of a solution for the HJB equation associated with the discounted optimization problem as well as conditions for the existence of an optimal selector for this equation. Moreover it shows that the solution of the HJB equation is in fact unique and coincides with the optimal value for the discounted optimal control problem, and provides the existence of an optimal deterministic stationary strategy for the problem.

Theorem 3.6: Suppose Assumptions A, B and C are satisfied. Then there exist \( W \in A_b(\mathbb{X}) \) and \( \mathcal{X}W \in B(\mathbb{X}) \) satisfying

\[
-\alpha W(x) + \mathcal{X}W(x) + \inf_{\alpha \in A_{\alpha}(x)} \left\{ C_{\alpha}(x, a) + qW(x, a) \right\} = 0,
\]

for any \( x \in \mathbb{X} \), and

\[
W(z) = \inf_{b \in A(z)} \left\{ C(z, b) + QW(z, b) \right\},
\]

for any \( z \in \mathcal{Z} \). Moreover there is a measurable mapping \( \widehat{\varphi}_\alpha : \mathbb{X} \to A \) such that \( \widehat{\varphi}_\alpha(y) \in A(y) \) for any \( y \in \mathbb{X} \) and satisfying

\[
C_{\alpha}(x, \widehat{\varphi}_\alpha(x)) + qW(x, \widehat{\varphi}_\alpha(x)) = \inf_{a \in A(x)} \left\{ C_{\alpha}(x, a) + qW(x, a) \right\},
\]

(18)
for any \( x \in X \), and

\[
C^i(z, \hat{\varphi}_\alpha(z)) + QW(z, \hat{\varphi}_\alpha(z)) = \inf_{b \in A(z)} \left\{ C^i(z, b) + QW(z, b) \right\},
\]

(19)

for any \( z \in \mathcal{Z} \). The deterministic stationary strategy \( \hat{\varphi}_\alpha \) is optimal for the \( \alpha \)-discounted problem and the function \( W \in A_{\hat{\varphi}}(\mathcal{X}) \), solution of (16)–(17), is unique and coincides with \( \mathcal{V}^*_\alpha(x) = \inf_{u \in \mathcal{U}} \mathcal{V}_\alpha(u, x) \). Furthermore we have that

\[
|\mathcal{V}^*_\alpha(x)| \leq K_\alpha + K I_{\mathcal{A}_1}(x).
\]

(20)

**Proof:** See Theorem 5.5 and Proposition 5.6 in [7].

### 4. Main results

In this section we present the main results of this paper (see Theorems 4.1 and 4.2). In Theorem 4.1 we obtain sufficient conditions for the existence of a solution and optimal selector to an integro-differential HJB inequality, which in turns is derived by using the so-called vanishing discount approach. Based on the solution of this integro-differential HJB inequality we derive in Theorem 4.2 a deterministic stationary optimal policy for the infinite-horizon long run average continuous-time control problem, as defined in Definition 2.3. Define for all \( x \in \mathcal{X} \),

\[
m_\alpha = \inf_{x \in \mathcal{X}} \mathcal{V}^*_\alpha(x), \quad \rho_\alpha = \alpha m_\alpha,
\]

(21)

\[
h_\alpha(x) = \mathcal{V}^*_\alpha(x) - m_\alpha \geq 0.
\]

(22)

We will consider the following final assumption (see the Appendix for the definition of the generalized inferior limit \( \lim^g \)):

**Assumption D:** \( \lim^g_{\alpha \to 0} h_\alpha(x) < \infty \) for all \( x \in \mathcal{X} \).

From (15), (20) and (21) we get for any \( 0 < \alpha < 1 \) that

\[
|\rho_\alpha| = |\alpha m_\alpha| \leq \alpha \left| \inf_{x \in \mathcal{X}} \mathcal{V}^*_\alpha(x) \right| \leq \alpha \sup_{x \in \mathcal{X}} |\mathcal{V}^*_\alpha(x)| \leq \alpha K_\alpha + K \leq K_C + K
\]

(23)

and thus we can find a subsequence \( \{\alpha_n\} \) such that \( \lim_{n \to \infty} \alpha_n = 0 \) and \( \lim_{n \to \infty} \rho_{\alpha_n} = \rho \) for some \( |\rho| \leq K_C + K \). We define

\[
h_*(x) = \lim_{n \to \infty} h_{\alpha_n}(x).
\]

(24)

Clearly \( h_*(x) \geq 0 \) since from (22) we have that \( h_\alpha(x) \geq 0 \). From Assumption D we have that \( h_*(x) < \infty \), and from Proposition A.2, \( h_* \in \mathbb{L}_+(\mathcal{X}) \).
From Theorem 3.6 we have that \( W(x) = V^*_\alpha(x) \) satisfies (16) and (17), and thus from (21), (22) and after some algebraic manipulations we obtain that

\[
-(\alpha + K)h_\alpha(x) + \mathcal{X}h_\alpha(x) + \inf_{a \in A^p(x)} \left\{ C^g(x, a) + qh_\alpha(x, a) + Kh_\alpha(x) \right\} - \rho_\alpha = 0, \tag{25}
\]

for any \( x \in X \),

\[
h_\alpha(z) = \inf_{b \in A^l(z)} \left\{ C^i(z, b) + Qh_\alpha(z, b) \right\}, \tag{26}
\]

for any \( z \in \mathcal{Z} \). Moreover, according to Theorem 3.6 there exists a measurable selector \( \hat{\varphi}_\alpha : \overline{X} \to A \) satisfying \( \hat{\varphi}_\alpha(y) \in A(y) \) for any \( y \in \overline{X} \) reaching the infimum in (25) and (26). Thus,

\[
-(\alpha + K)h_\alpha(x) + \mathcal{X}h_\alpha(x) + C^g(x, \hat{\varphi}_\alpha(x)) + qh_\alpha(x, \hat{\varphi}_\alpha(x)) + Kh_\alpha(x) - \rho_\alpha = 0, \tag{27}
\]

for any \( x \in X \),

\[
h_\alpha(z) = C^i(z, b) + Qh_\alpha(z, \hat{\varphi}_\alpha(z)) \tag{28}
\]

for any \( z \in \mathcal{Z} \).

The following theorem provides sufficient conditions for the existence of a solution and optimal selector to an integro-differential HJB inequality, associated to the long run average control problem.

**Theorem 4.1:** Suppose that Assumptions A, B, C and D are satisfied. Then the following holds:

(a) There exist \( H \in A(\overline{X}) \cap L(\overline{X}) \) bounded from below satisfying

\[
\rho \geq \mathcal{X}H(x) + \inf_{a \in A^p(x)} \left\{ C^g(x, a) + qH(x, a) \right\}, \tag{29}
\]

for any \( x \in X \), and

\[
H(z) \geq \inf_{b \in A^l(z)} \left\{ C^i(z, b) + QH(z, b) \right\}, \tag{30}
\]

for any \( z \in \mathcal{Z} \).

(b) There is a measurable mapping \( \hat{\varphi} : \overline{X} \to A \) such that \( \hat{\varphi}(y) \in A(y) \) for any \( y \in \overline{X} \) and satisfying

\[
C^g(x, \hat{\varphi}(x)) + qH(x, \hat{\varphi}(x)) = \inf_{a \in A(x)} \left\{ C^g(x, a) + qH(x, a) \right\}, \tag{31}
\]

for any \( x \in X \), and

\[
C^i(z, \hat{\varphi}(z)) + QH(z, \hat{\varphi}(z)) = \inf_{b \in A(z)} \left\{ C^i(z, b) + QH(z, b) \right\}, \tag{32}
\]

for any \( z \in \mathcal{Z} \).

**Proof:** The proof of (a) will follow steps S1 and S2 below:
(S1) $h_s(x)$ defined in (24) satisfies the following inequality:

$$ h_s(x) \geq \int_{[0,t^*(x)]} e^{-Ks} (\mathcal{R}h_s(\phi(x,s)) - \rho) \, ds + e^{-Kt^*(x)} \mathcal{T}h_s(\phi(x,t^*(x))). \quad (33) $$

(S2) Define $H(x)$, for all $x \in \overline{X}$ as follows:

$$ H(x) = \int_{[0,t^*(x)]} e^{-Ks} (\mathcal{R}h_s(\phi(x,s)) - \rho) \, ds + e^{-Kt^*(x)} \mathcal{T}h_s(\phi(x,t^*(x))). \quad (34) $$

Then $H \in A(\overline{X}) \cap L(\overline{X})$ and satisfies (29) and (30).

For the proof of S1, taking the integral of (27) along the flow $\phi(x,t)$, we get from (27) and (28) (see [6]) that for any $x \in X$,

$$ h_\alpha(x) = \int_{[0,t^*(x)]} e^{-(K+\alpha)t} (\mathcal{R}h_\alpha(\phi(x,t)) - \rho_\alpha) \, dt + e^{-(K+\alpha)t^*(x)} \mathcal{T}h_\alpha(\phi(x,t^*(x))). \quad (35) $$

Notice that, according to Theorem 3.6, we have

$$ \mathcal{R}h_\alpha(y) = C^b(y, \widehat{\varphi}_\alpha(y)) + qh_\alpha(y, \widehat{\varphi}_\alpha(y)) + Kh_\alpha(y), \quad y \in X, \quad (36) $$

$$ \mathcal{T}h_\alpha(z) = C^i(z, \widehat{\varphi}_\alpha(z)) + Qh_\alpha(z, \widehat{\varphi}_\alpha(z)), \quad z \in \Xi. \quad (37) $$

From Proposition A.2(i), there exists a sequence $\{x_n\} \subset X$ such that $x_n \to x$ and $\lim_{n \to \infty} h_\alpha_n(x_n) = h_\alpha(x)$. Since $\mathcal{R}h_\alpha$ is bounded from below and $\rho_\alpha$ is bounded, we can apply the Fatou's lemma in (35) to get that

$$ h_\alpha(x) = \lim_{n \to \infty} h_\alpha_n(x_n) \geq \int_{[0,\infty]} \lim_{n \to \infty} \left( I_{[0,t_n^*(x)]}(t) e^{-(K+\alpha_n)t} \mathcal{R}h_\alpha_n(x_n(t)) - \rho_\alpha_n \right) \, dt $$

$$ + \lim_{n \to \infty} e^{-(K+\alpha_n)t_n^*} \mathcal{T}h_\alpha_n(x_n(t_n^*)), \quad (38) $$

where for notational simplicity, we set

$$ x_n(t) = \phi(x_n, t), \quad x(t) = \phi(x, t), \quad a_n(t) = \widehat{\varphi}_\alpha_n(x_n(t)), \quad t_n^* = t^*(x_n). $$

From Assumption (B5) and the convergence of $\rho_\alpha_n$ to $\rho$ we get that

$$ \lim_{n \to \infty} I_{[0,t_n^*(x)]}(t) e^{-(K+\alpha_n)t} \mathcal{R}h_\alpha_n(x_n(t)) - \rho_\alpha_n = I_{[0,t^*(x)]}(t) e^{-Kt} \left( \lim_{n \to \infty} \mathcal{R}h_\alpha_n(x_n(t)) - \rho \right), \quad (39) $$

a.s. on $[0, \infty)$ and

$$ \lim_{n \to \infty} e^{-(K+\alpha_n)t_n^*} \mathcal{T}h_\alpha_n(x_n(t_n^*)) = e^{-Kt^*(x)} \lim_{n \to \infty} \mathcal{T}h_\alpha_n(x_n(t_n^*)). \quad (40) $$

For a fixed $t \in (0, t^*(x))$, there is no loss of generality in assuming that $t < t_n^*$ for any $n \in \mathbb{N}$ and thus $x_n(t) \in X$. Let us consider first the term $\lim_{n \to \infty} \mathcal{R}h_\alpha_n(x_n(t))$ in (39). We can find a subsequence $\{n_j\}$ of $\{n\}$ such that we have $\lim_{n \to \infty} \mathcal{R}h_\alpha_n(x_n(t)) =$
\[
\lim_{j \to \infty} \mathcal{R} h_{\alpha_n}(x_n(t_j)) \quad \text{for } t_j \to t^*(x) < \infty.
\]
Since the multifunction \( \Psi^\delta \) is compact valued and upper semicontinuous (see Assumptions (B1) and (C1)) and \( x_{n_j} \to x(t) \) (see Assumption (B4)), we can find a subsequence \( \{ a_{n_j}(t) \} \in A(x_{n_j}(t)) \), still denoted by \( a_{n_j}(t) \) such that \( a_{n_j}(t) \to a \in A(x(t)) \) (see Theorem 17.16 in [1]) as \( j \to \infty \). Thus, according to (36) we have that
\[
\lim_{n \to \infty} \mathcal{R} h_{\alpha_n}(x_n(t)) = \lim_{j \to \infty} \left( C^\delta(x_n(t), a_{n_j}(t)) + q h_{\alpha_n}(x_{n_j}(t), x_n(t)) + K h_{\alpha_n}(x_n(t)) \right),
\]
and therefore
\[
\lim_{n \to \infty} \mathcal{R} h_{\alpha_n}(x_n(t)) \geq \lim_{j \to \infty} C^\delta(x_{n_j}(t), a_{n_j}(t)) + \lim_{j \to \infty} \left( q h_{\alpha_n}(x_{n_j}(t), x_n(t)) + K h_{\alpha_n}(x_n(t)) \right). \tag{41}
\]
From the fact that \( C^\delta \) is lower semicontinuous on \( K^\delta \) we get that
\[
\lim_{j \to \infty} C^\delta(x_{n_j}(t), a_{n_j}(t)) \geq C^\delta(x(t), a). \tag{42}
\]
From Proposition A.2(i) and (iii), the fact that \( Q \) is weakly continuous on \( K^\delta \), and the continuity of \( \lambda \), we get that
\[
\lim_{j \to \infty} \lambda(x_{n_j}(t), a_{n_j}(t)) Q h_{\alpha_n}(x_{n_j}(t), a_{n_j}(t)) \geq \lambda(x(t), a) Q h_{\delta}(x(t), a) \tag{43}
\]
and, recalling that \( K - \lambda(x_{n_j}(t), a_{n_j}(t)) \geq 0 \) from Assumption (A1), we conclude that
\[
\lim_{j \to \infty} \left[ K - \lambda(x_{n_j}(t), a_{n_j}(t)) \right] h_{\alpha_n}(x_{n_j}(t), a_{n_j}(t)) \geq \left[ K - \lambda(x(t), a) \right] h_{\delta}(x(t), a). \tag{44}
\]
Combining (36), (41), (42), (43), (44), we get that
\[
\lim_{n \to \infty} \mathcal{R} h_{\alpha_n}(x_n(t))) \geq C^\delta(x(t), a) + q h_{\delta}(x(t), a) + K h_{\delta}(x(t)) \geq \mathcal{R} h_{\delta}(x(t)). \tag{45}
\]
Let us consider now, for \( t^*(x) < \infty \), the term \( \lim_{n \to \infty} \mathcal{H} h_{\alpha_n}(x_n(t^n_n)) \) in (40). Similarly as before, since \( \Psi^\delta \) is compact valued and upper semicontinuous and \( x_{n}(t^n_n) \to x(t^*(x)) \), we can find a subsequence \( \{ a_{n_j}(t^n_n) \} \in A(x_{n}(t^n_n)) \) such that \( a_{n_j}(t^n_n) \to b \in A(x(t^*(x))) \) (see again Theorem 17.16 in [1]). Using similar arguments as above (in particular equation (37)) we can show that there exists \( b \in A(x(t^*(x))) \) such that
\[
\lim_{n \to \infty} \mathcal{H} h_{\alpha_n}(x_n(t^n_n)) \geq C^\delta(x(t^*(x)), b) + q h_{\delta}(x(t^*(x)), b) \geq \mathcal{H} h_{\alpha}(x(t^*(x))). \tag{46}
\]
From (38), (39), (45) and (46) we conclude that (33) holds, showing step (S1).

For step (S2), we first notice, by considering in Lemma 3.4 the operator \( \mathcal{H} \) with the cost \( C^\delta(x, a) - \rho \) instead of \( C^\delta(x, a) \), that \( H(x) = \mathcal{B}_0 h_{\delta}(x) \) and, since \( h_{\delta} \in L(\overline{X}) \), we obtain from Lemma 3.4 that \( H \in L(\overline{X}) \) and it is bounded from below. Moreover, it follows from (33) that \( H(x) \leq h_{\delta}(x) \), implying that \( H \in L(\overline{X}) \).
From Lemma 3.3 we get that \( H(x) \in \mathbb{A}(\overline{X}) \) and the existence of a bounded from below measurable function \( \mathcal{X}H \) satisfying

\[
- KH(x) + \mathcal{X}H(x) + \inf_{a \in \mathcal{A}(x)} \left\{ C^g(x,a) + qh(x,a) + Kh(x) \right\} = \rho
\]

(47)

for any \( x \in X \) and

\[
H(z) = \inf_{b \in \mathcal{A}(z)} \left\{ C^l(z,b) + Qh(z,b) \right\},
\]

(48)

for any \( z \in \Xi \). Since \( h^* - H(x) \geq 0 \), we have from (47) and (48) that for any \( x \in X \),

\[
\mathcal{X}H(x) + \inf_{a \in \mathcal{A}(x)} \left\{ C^g(x,a) + qH(x,a) \right\} \\
= -KH(x) + \mathcal{X}H(x) + \inf_{a \in \mathcal{A}(x)} \left\{ C^g(x,a) + qh(x,a) + Kh(x) \right\} \\
\leq -KH(x) + \mathcal{X}H(x) + \inf_{a \in \mathcal{A}(x)} \left\{ C^g(x,a) + qh(x,a) + Kh(x) \right\} = \rho
\]

(49)

and for any \( z \in \Xi \),

\[
\inf_{b \in \mathcal{A}(z)} \left\{ C^l(z,b) + QH(z,b) \right\} \leq \inf_{b \in \mathcal{A}(z)} \left\{ C^l(z,b) + Qh(z,b) \right\} = H(z).
\]

(50)

From (49), (50), we conclude that (29) and (30) hold, showing step (S2) and completing the proof of part (a) of the theorem.

Part (b) follows from the fact that \( H \) is lower semicontinuous on \( \overline{X} \), Assumptions A, B, C and Proposition D.5 in [19].

Theorem 4.2: Suppose that Assumptions A, B, C and D are satisfied and consider \( \hat{\varphi} \) as in (31), (32). Then the deterministic stationary strategy \( \hat{\varphi} \) is optimal for the average cost problem and for any \( x \in X \),

\[
\rho = A(\hat{\varphi}, x) = A^* (x).
\]

(51)

Proof: We first notice that, from Proposition 4.6 in [6],

\[
\lim_{\alpha \downarrow 0} \alpha \mathcal{V}^*_\alpha (x) \leq A^* (x).
\]

Thus it follows that

\[
\rho = \lim_{n \to \infty} \alpha_n \inf_{x \in \overline{X}} \mathcal{V}^*_n (x) \leq \lim_{n \to \infty} \alpha_n \mathcal{V}^*_n (x) \leq A^* (x).
\]
To show the other side of the inequality, define for a function $W \in \mathbb{M}(X)$ bounded from below and $\hat{\varphi}$ as in b),

$$
G\hat{\varphi} W(x) = \int_{[0, t^*(x)]} e^{-\Lambda\hat{\varphi}(x,s)} \lambda\hat{\varphi} Q\hat{\varphi} W(\phi(x, s))) \, ds + e^{\Lambda\hat{\varphi}(x,t^*(x))} Q\hat{\varphi} W(\phi(x, t^*(x))),
$$

$$
L\hat{\varphi} W(x) = \int_{[0, t^*(x)]} e^{-\Lambda\hat{\varphi}(x,s)} W(\phi(x, s)) \, ds,
$$

$$
\mathcal{L}\hat{\varphi}(x) = \int_{[0, t^*(x)]} e^{-\Lambda\hat{\varphi}(x,s)} \, ds,
$$

$$
P\hat{\varphi} W(x) = e^{-\Lambda\hat{\varphi}(x,t^*(x))} W(\phi(x, t^*(x))).
$$

From (29) and (30) we get that

$$
H(x) \geq T\varphi(\rho, H)(x) = -\rho \mathcal{L}\hat{\varphi}(x) + L\hat{\varphi} Cg(x) + P\hat{\varphi} Ci(x) + G\hat{\varphi} H(x).
$$

From Proposition 3.3 in [6] we get that

$$
\hat{J}_{\varphi} m(t, x) \leq H(x)
$$

where

$$
\hat{J}_{\varphi} m(t, x) = \mathbb{E}_{\hat{X}} \left[ \int_{[0, t \wedge T_m]} \left[ Cg(\xi_s, \hat{\varphi}) - \rho \right] \, ds + \int_{[0, t \wedge T_m]} C^i((\xi_s - \hat{\varphi}) \, dp^*(s) + T\varphi(\rho, H)(\xi_{t \wedge T_m}) \right].
$$

Since $T\varphi(\rho, H)$ is bounded from below by, say, $-c_0$, we get from above that

$$
- c_0 + \mathbb{E}_{\hat{X}} \left[ \int_{[0, t \wedge T_m]} \left[ Cg(\xi_s, \hat{\varphi}) \right] \, ds + \int_{[0, t \wedge T_m]} C^i((\xi_s - \hat{\varphi}) \, dp^*(s) \right] 
\leq H(x) + \rho \mathbb{E}_{\hat{X}} (t \wedge T_m).
$$

Taking the limit as $m$ goes to infinity we obtain that

$$
- c_0 + \mathbb{E}_{\hat{X}} \left[ \int_{[0, t]} \left[ Cg(\xi_s, \hat{\varphi}) \right] \, ds + \int_{[0, t]} C^i((\xi_s - \hat{\varphi}) \, dp^*(s) \right] \leq H(x) + \rho t,
$$

and thus $A^*(x) \leq A(x, \hat{\varphi})(x) \leq \rho$.

### Disclosure statement

No potential conflict of interest was reported by the authors.

### Funding

The first author received financial support from CNPq (Brazilian National Council for Scientific and Technological Development), grant 304091/2014-6, FAPESP (São Paulo Research Foundation)/ Shell through the Research Centre for Gas Innovation, FAPESP Grant Proc. 2014/50279-4, project INCT, grant CNPq 465755/2014-3, FAPESP 2014/50851-0 and FUSP (University of São Paulo Support Foundation).
References


Appendix

As in [25] the following definition of the generalized inferior limit will be used along the paper:

**Definition A.1:** Let $X$ be a Borel space and let $\{w_n\}$ be a family of functions in $\mathcal{M}(X)$. The generalized inferior limit of the sequence $\{w_n\}$, denoted by $\liminf_{n \to \infty} w_n$, is defined as

$$\liminf_{n \to \infty} w_n(x) = \sup_k \left( \sup_{\epsilon > 0} \inf_{m \geq k} \inf \left( \inf_{d(y,x) < \epsilon} w_m(y) \right) \right) \tag{A1}$$

where $d(\cdot, \cdot)$ is the metric in $X$. For notational convenience, $\liminf_{n \to \infty} w_n$ will be denoted by $w^\ast$.

The following properties from the generalized inferior limit will be used in the sequel.

**Proposition A.2:** Let $\{w_n\}$ be a sequence of nonnegative functions in $\mathcal{M}(X)$ and consider an arbitrary $x \in X$. In this case, $w^\ast(x)$ as defined in (A.1) satisfies the following properties:

(i) For any sequence $\{x_n\}$ such that $x_n \to x$, it follows that $\liminf_{n \to \infty} w_n(x_n) \geq w^\ast(x)$, and there exists a sequence $\{x_n\}$ such that $x_n \to x$ and $\liminf_{n \to \infty} w_n(x_n) = w^\ast(x)$.

(ii) $w^\ast \in \mathcal{L}_+(X)$.

(iii) [Generalized Fatou’s Lemma] Suppose that $\{\mu_n\}$ is a sequence of probability measures in $\mathcal{P}(X)$ and that $\{\mu_n\}$ converges weakly to a $\mu \in \mathcal{P}(X)$. Then

$$\liminf_{n \to \infty} \int_{\mathcal{S}} w_n(x) \mu_n(dx) \geq \int_{\mathcal{S}} w^\ast(x) \mu(dx). \tag{A2}$$

**Proof:** For the proof of (i) see Lemma 4.1 in [4]. For (ii) see Lemma 3.1 in [22] and for (iii) see Lemma 3.2 in [22].