CrossMark

# Impulsive Control for Continuous-Time Markov Decision Processes: A Linear Programming Approach

**F. Dufour**[1,2,3] · **A. B. Piunovskiy**[4]

**Abstract** In this paper, we investigate an optimization problem for continuous-time Markov decision processes with both impulsive and continuous controls. We consider the so-called constrained problem where the objective of the controller is to minimize a total expected discounted optimality criterion associated with a cost rate function while keeping other performance criteria of the same form, but associated with different cost rate functions, below some given bounds. Our model allows multiple impulses at the same time moment. The main objective of this work is to study the associated linear program defined on a space of measures including the occupation measures of the controlled process and to provide sufficient conditions to ensure the existence of an optimal control.

---

✉ F. Dufour
dufour@math.u-bordeaux1.fr

A. B. Piunovskiy
piunov@liv.ac.uk

[1] Bordeaux INP, IMB, UMR CNRS 5251, Talence, France

[2] Universite de Bordeaux, IMB, UMR CNRS 5251, Talence, France

[3] Inria Bordeaux-Sud-Ouest, 200 Avenue de la Vieille Tour, 33405 Talence Cedex, France

[4] Department of Mathematical Sciences, University of Liverpool, Liverpool L69 7ZL, United Kingdom

# 1 Introduction

The objective of this work is to study continuous-time Markov decision processes (CTMDP's) with constraints for the infinite horizon discounted cost with both impulsive and continuous controls. CTMDP's is a general family of controlled stochastic processes suitable for the modeling of sequential decision-making problems. They appear in many fields such as engineering, computer science, economics and operational research among others. Our goal is to study the two types of control for CTMDPs as described by Davis in his book [6]: continuous control acting at all times on the process through the transition rate, and impulsive control, used to describe control actions that move the process to a new point of the state space at some specific times. Continuous control for CTMDP's has been extensively studied in the literature, see for example the recent books [9,10,27] and the references therein. Meanwhile, impulsive control for CTMDP's has received less attention and was studied in [6,8,15,16,25,29–32]; see also the recent work [4], where a general Markov model was considered with application to financial mathematics. We do not attempt to present here an exhaustive panorama on this topic, but refer the interested reader to [8] for a brief survey of CTMDP with impulsive control.

It is important to emphasize that in the framework of impulsive control for CTMDPs, there exist two rather distinct families of problems. The first class is related to models allowing only one impulsive action at a time. The second family is more general and studies models with possibly multiple impulses at the same time moment. This latter set of problems is much more delicate for the analysis. Indeed, if the process may take different values at the same time moment then it leads to non standard paths for the controlled process. Most of the works in the literature are concerned with the first class of problems. The second family of problems have been addressed mainly by Yushkevich in [29–32]. Yushkevich has introduced a new class of stochastic models, the so called T-processes where roughly speaking the processes are indexed by a parameter representing the natural current time and the number of the impulsive actions at that time moment. In [8], another approach has been developed by the authors in order to use the standard theory of marked point processes [17,23]. Roughly speaking, the model discussed in [8] is defined by the following components: the state space $\mathbf{X}$, the set of continuous actions $\mathbf{A}^g$, the space of impulsive actions $\mathbf{A}^i$, a transition rate $q$ on $\mathbf{X}$ given $\mathbf{X} \times \mathbf{A}^g$ and a stochastic kernel $Q$ on $\mathbf{X}$ given $\mathbf{X} \times \mathbf{A}^i$. The model is given by a marked point process $(\Theta_n, Y_n)_{n \in \mathbb{N}}$ where $\Theta_n$ represents the sojourn time between two consecutive epochs induced either by a natural jump or by an intervention of the decision-maker. The natural jumps are generated by the transition rate $q$. The state vector $Y_n$ represents the successive jumps of the process and the associated impulsive actions at the $n$-$th$ epoch. More precisely, $Y_n$ is of the form $(x_0, a_0, x_1, a_1, \ldots, x_k, a_k, x_{k+1}, \Delta, \Delta, \ldots)$ where $x_0$ corresponds to a possibly natural jump or to the value of the process just before the intervention. The triple $(x_j, a_j, x_{j+1})$ indicates that the impulsive action $a_j$ has been applied to the system at state $x_j$ leading to a new jump $x_{j+1}$ having distribution $Q(\cdot|x_j, a_j)$. The special impulsive action $\Delta$ means that the impulses are over and the artificial state $\Delta$ means the same.

The necessity of an immediate sequence of impulses appears naturally in many mathematical models describing real-life problems. For example, in heavy traffic control problems, the approximation of the real physical system may lead to control problems with simultaneous multiple impulses in the limit model. For the original physical system, the impulses are separated in time but for the limit model, impulses can occur at the same time moment due to the fact that the time is rescaled and compressed. For a detailed exposition of such phenomenon, the reader is referred for example to Sect. 8 of the book [21] (and the references therein) where an example of a production system in heavy traffic with impulsive control is discussed.

From a theoretical point of view, constrained CTMDPs are substantially different and more difficult to study than unconstrained CTMDPs. The linear programming technique has proved to be a very efficient method for solving such problems. The key idea is to reformulate the original sequential decision-making problem as an infinite dimensional static optimization problem over a space of measures where the admissible solutions are the so-called occupation measures of the controlled process. This technique has been extensively studied in the recent decades for continuous time processes in the context of continuous (gradual) control. We do not pretend to present here an exhaustive panorama of this approach but the interested reader may consult the following works and the references therein: [11,21,24,26] for CTMDPs, [1,3,22] for diffusion processes and [2,12,20,28] for controlled martingale problems. However, the linear programming approach has been considerably less studied for impulsive control for unconstrained and constrained problems. To the best of our knowledge, this paper seems to be the first attempt to tackle such problems. By using the Lagrange multiplier approach, impulsive control problem with constraints has been addressed in [25] for a model with no continuous control, finite state and action spaces and where a single impulsive action is allowed at a time.

In this paper, we investigate a constrained optimization problem for a CTMDP with general state and action spaces, and with both impulsive and continuous controls where the performance and the constraint criteria are given in terms of infinite-horizon discounted functionals. Our model allows multiple impulses at the same time moment. To the best of our knowledge, this paper can be seen as the first attempt to solve such general CTMDP's. A distinguished feature of this work with respect to [8] is that in the present paper we consider the constrained control problem while in [8], the unconstrained case has been studied by using the dynamic programming technique. The main objective of the present work is to provide sufficient conditions to ensure the existence of an optimal control. First, we study the properties of the occupation measures. It is shown in Theorem 4.5 that for any admissible control strategy, the corresponding occupation measure satisfies a specific linear equation. It is then proved that this linear equation characterizes the optimal control problem under consideration, in the sense that, from any measure $\eta$ satisfying such linear equation one can construct a control strategy $u$ such that the corresponding occupation measure $\eta_u$ is *smaller* than $\eta$ (for a precise mathematical statement, see Theorem 4.9). Based on these properties, one can introduce a linear program, labeled $\mathbb{PLP}$ and prove that the solvability of the constrained optimization problem is equivalent to the solvability of the $\mathbb{PLP}$ and that these two optimization problems give the same value. By introducing a set of weak hypotheses, the solvability of the $\mathbb{PLP}$ is proved in Theorem 5.5. Finally, it is shown

in Theorem 5.6 the existence of an optimal randomized control strategy and that the class of such strategies is a sufficient set for the constrained optimization problem under consideration.

The rest of the paper is organized as follows. In Sect. 2, the CTMDP under consideration is discussed. Section 3 is devoted to the presentation of the performance criteria and the introduction of the main assumptions. The properties of the occupation measures are derived in Sect. 4. Finally, the linear program is studied in Sect. 5 where the existence of an optimal control strategy is shown. Several auxiliary results are presented in the Appendices 1 and 2 to streamline the presentation.

## 2 The Continuous-Time Markov Control Process

The main goal of this section is to introduce the notations, as well as the parameters defining the model, and to present the construction of the controlled process. In particular, having defined the class of admissible strategies, we introduce a probability measure $\mathbb{P}_{x_0}^u$ with respect to which the controlled process $(\Theta_n, Y_n)_{n \in \mathbb{N}}$ has the required conditional distributions.

The following basic notation will be used in this paper: $\mathbb{N}$ is the set of natural numbers including 0, $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$, $\mathbb{R}$ denotes the set of real numbers, $\mathbb{R}_+$ the set of non-negative real numbers, $\mathbb{R}_+^* = \mathbb{R}_+ \setminus \{0\}$, $\overline{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{+\infty\}$ and $\overline{\mathbb{R}}_+^* = \mathbb{R}_+^* \cup \{+\infty\}$. For any $p \in \mathbb{N}$, $\mathbb{N}_p$ is the set $\{0, 1, \ldots, p\}$ and for any $p \in \mathbb{N}^*$, $\mathbb{N}_p^*$ is the set $\{1, \ldots, p\}$. The term *measure* will always refer to a countably additive, $\overline{\mathbb{R}}_+$-valued set function. A finite (respectively, signed) measure is a countably additive, $\mathbb{R}_+$-valued (respectively, $\mathbb{R}$-valued) set function. Let $\mathbf{X}$ be a Borel space and denote by $\mathcal{B}(\mathbf{X})$ its associated Borel $\sigma$-algebra. For any set $A$, $I_A$ denotes the indicator function of the set $A$. The set of measures (respectively, signed measures) defined on $(\mathbf{X}, \mathcal{B}(\mathbf{X}))$ is denoted by $\mathcal{M}(\mathbf{X})$ (respectively, $\mathcal{M}_s(\mathbf{X})$). The set of finite measures on $(\mathbf{X}, \mathcal{B}(\mathbf{X}))$ is denoted by $\mathcal{M}_f(\mathbf{X})$ and $\mathcal{P}(\mathbf{X})$ is the set of probability measures defined on $(\mathbf{X}, \mathcal{B}(\mathbf{X}))$. For any point $x \in \mathbf{X}$, $\delta_x$ denotes the Dirac measure defined by $\delta_x(\Gamma) = I_\Gamma(x)$ for any $\Gamma \in \mathcal{B}(\mathbf{X})$. The set of bounded real-valued measurable functions defined on $X$ is denoted by $\mathbb{B}(\mathbf{X})$ and the set of $\mathbb{R}$-valued (respectively, $\overline{\mathbb{R}}_+$-valued, and $\overline{\mathbb{R}}_+^*$-valued) measurable functions defined on $\mathbf{X}$ is denoted by $\mathbb{M}(\mathbf{X})$ (respectively, $\overline{\mathbb{M}}_+(\mathbf{X})$, and $\overline{\mathbb{M}}_+^*(\mathbf{X})$). The set of continuous functions in $\mathbb{B}(\mathbf{X})$ is denoted by $\mathbb{C}_b(\mathbf{X})$.

Let $\mathbf{X}$ and $\mathbf{Z}$ be two Borel spaces. A kernel (respectively, signed kernel) $T(\cdot|\cdot)$ on $\mathbf{Z}$ given $\mathbf{X}$ is an $\overline{\mathbb{R}}_+$-valued (respectively, $\mathbb{R}$-valued) mapping defined on $\mathcal{B}(\mathbf{Z}) \times \mathbf{X}$ such that for any $A \in \mathcal{B}(\mathbf{Z})$, $T(A|\cdot) \in \overline{\mathbb{M}}_+(\mathbf{X})$ (respectively, $T(A|\cdot) \in \mathbb{M}(\mathbf{X})$) and for any $x \in \mathbf{X}$, $T(\cdot|x) \in \mathcal{M}(\mathbf{Z})$ (respectively, $T(\cdot|x) \in \mathcal{M}_s(\mathbf{Z})$). A kernel $T(\cdot|\cdot)$ on $\mathbf{Z}$ given $\mathbf{X}$ is called stochastic (respectively, finite) if for any $x \in \mathbf{X}$, $T(\cdot|x) \in \mathcal{P}(\mathbf{Z})$ (respectively, $T(\cdot|x) \in \mathcal{M}_f(\mathbf{Z})$). $\mathcal{P}(\mathbf{Z}|\mathbf{X})$ denotes the set of stochastic kernels on $\mathbf{Z}$ given $\mathbf{X}$. A transition rate $q$ on $\mathbf{X}$ given $\mathbf{X} \times \mathbf{Z}$ is a signed kernel on $\mathbf{X}$ given $\mathbf{X} \times \mathbf{Z}$ satisfying $q(\mathbf{X}|x, z) = 0$ and $q(\Gamma \setminus \{x\}|x, z) \geq 0$ for any $\Gamma \in \mathcal{B}(\mathbf{X})$ and any $(x, z) \in \mathbf{X} \times \mathbf{Z}$. To any transition rate $q$ on $\mathbf{X}$ given $\mathbf{X} \times \mathbf{Z}$, we associate a kernel $\overline{q}$ on $\mathbf{X}$ given $\mathbf{X} \times \mathbf{Z}$ defined by $\overline{q}(\Gamma|x, z) = q(\Gamma \setminus \{x\}|x, z)$ for any $\Gamma \in \mathcal{B}(\mathbf{X})$ and any $(x, z) \in \mathbf{X} \times \mathbf{Z}$.

Let $\eta \in \mathcal{M}(\mathbf{X})$, $f \in \overline{\mathbb{M}}(\mathbf{Z})_+$ and a kernel $T$ on $\mathbf{Z}$ given $\mathbf{X}$, then $\eta T$ denotes the measure on $\mathbf{Z}$ defined by $\eta T(\Gamma) = \int_{\mathbf{X}} T(\Gamma|x)\eta(dx)$ for any $\Gamma \in \mathcal{B}(\mathbf{Z})$ and $Tf$ denotes the function defined on $\mathbf{X}$ by $Tf(x) = \int_{\mathbf{Z}} f(z)T(dz|x)$ for any $x \in \mathbf{X}$. Moreover, if $R$ is a kernel on $\mathbf{Y}$ given $\mathbf{Z}$ then $TR$ is the kernel on $\mathbf{Y}$ given $\mathbf{X}$ defined by $TR(\Gamma|x) = \int_{\mathbf{Z}} R(\Gamma|z)T(dz|x)$ for any $\Gamma \in \mathcal{B}(\mathbf{Y})$ and $x \in \mathbf{X}$.

Finally, the infimum over an empty set is understood to be equal to $+\infty$.

## 2.1 Parameters of the Model

We deal with a control model defined through the following elements:

- $\mathbf{X}$ is the state space, assumed to be a Borel space (i.e., a measurable subset of a complete and separable metric space).
- $\mathbf{A}$ is the action space, assumed to be also a Borel space. $\mathbf{A}^i \in \mathcal{B}(\mathbf{A})$ (respectively $\mathbf{A}^g \in \mathcal{B}(\mathbf{A})$) is the set of impulsive (respectively continuous) actions satisfying $\mathbf{A} = \mathbf{A}^i \cup \mathbf{A}^g$ with $\mathbf{A}^i \cap \mathbf{A}^g = \emptyset$.
- The set of feasible actions in state $x \in \mathbf{X}$ is $\mathbf{A}(x)$, which is a nonempty measurable subset of $\mathbf{A}$. Admissible impulsive and continuous actions in the state $x \in \mathbf{X}$ are denoted by $\mathbf{A}^i(x) = \mathbf{A}(x) \cap \mathbf{A}^i$ and $\mathbf{A}^g(x) = \mathbf{A}(x) \cap \mathbf{A}^g$. It is supposed that $\mathbb{K}^g = \{(x, a) \in \mathbf{X} \times \mathbf{A} : a \in \mathbf{A}^g(x)\} \in \mathcal{B}(\mathbf{X} \times \mathbf{A}^g)$ and this set contains the graph of a measurable function from $\mathbf{X}$ to $\mathbf{A}^g$ (necessarily $\mathbf{A}^g(x) \neq \emptyset$ for all $x \in \mathbf{X}$) and that $\mathbb{K}^i = \{(x, a) \in \mathbf{X} \times \mathbf{A}^i : a \in \mathbf{A}^i(x)\} \in \mathcal{B}(\mathbb{X}^i \times \mathbf{A}^i)$ where $\mathbb{X}^i = \{x \in \mathbf{X} : \mathbf{A}^i(x) \neq \emptyset\} \in \mathcal{B}(\mathbf{X})$ and $\mathbb{K}^i$ contains the graph of a measurable function from $\mathbb{X}^i$ to $\mathbf{A}$.
- The stochastic kernel $Q$ on $\mathbf{X}$ given $\mathbb{K}^i$ describes the result of an impulsive action. In other words, if $x \in \mathbb{X}^i$ and an impulsive action $a \in \mathbf{A}^i(x)$ is applied then the state of the process changes instantly according to the stochastic kernel $Q$.
- The signed kernel $q$ on $\mathbf{X}$ given $\mathbb{K}^g$ is the intensity of jumps governing the dynamic of the process between interventions. For notational convenience, let us denote $q(\Gamma \setminus \{x\}|x, a)$ by $\overline{q}(\Gamma|x, a)$ for $\Gamma \in \mathcal{B}(\mathbf{X})$ and $(x, a) \in \mathbb{K}^g$. It satisfies $q(\mathbf{X}|x, a) = 0$, $\overline{q}(\mathbf{X}|x, a) \geq 0$ and $\sup_{a \in \mathbf{A}(x)} \overline{q}(\mathbf{X}|x, a) < \infty$ for any $(x, a) \in \mathbb{K}^g$.

In our model, an intervention consists only of a finite sequence of pairs of impulsive action and associated jump. Actually, this finite sequence can be equivalently described by an infinite sequence of pairs of state and action, where the pairs are set to the fictitious action and state after a finite step. As a result, an intervention is an element of the set

$$\mathbf{Y} = \bigcup_{k \in \mathbb{N}} \mathbf{Y}_k \text{ with } \mathbf{Y}_k = (\mathbf{X} \times \mathbf{A}^i)^k \times (\mathbf{X} \times \{\Delta\}) \times (\{\Delta\} \times \{\Delta\})^\infty,$$

where $\Delta$ will play the role of the fictitious state and action. The dynamic of such sequences is governed by the Markov Decision Process (MDP) $\mathcal{M}^i$ defined by $\mathcal{M}^i = \left(\mathbf{X}_\Delta, \mathbf{A}^i_\Delta, (\mathbf{A}^i_\Delta(x))_{x \in \mathbf{X}_\Delta}, Q_\Delta\right)$ where $\mathbf{X}_\Delta$, $\mathbf{A}^i_\Delta$ and $\left(\mathbf{A}^i_\Delta(x)\right)_{x \in \mathbf{X}_\Delta}$ are the new state and actions spaces augmented by the fictitious state $\Delta$: $\mathbf{X}_\Delta = \mathbf{X} \cup \{\Delta\}$, $\mathbf{A}^i_\Delta = \mathbf{A}^i \cup \{\Delta\}$

and $\mathbf{A}_\Delta^i(x) = \mathbf{A}^i(x) \cup \{\Delta\}$ for $x \in \mathbf{X}$ and $\mathbf{A}_\Delta^i(\Delta) = \{\Delta\}$. The dynamic is given by $Q_\Delta(.|x, a) = Q(.|x, a)$ for any $(x, a) \in \mathbb{K}^i$ and $Q_\Delta(\{\Delta\}|x, a) = 1$ otherwise. For the model $\mathcal{M}^i$, according to the Ionescu Tulcea's Theorem (see Proposition C.10 in [13]), there exists a unique strategic measure $P^\beta(\cdot|x)$ on $(\mathbf{X}_\Delta \times \mathbf{A}_\Delta^i)^\infty$ associated with the policy $\beta$ and the initial distribution $\delta_x$. Here and below, we use the standard terminology for MDP: see for example [13]. A policy is a sequence of past-dependent distributions on the action space. A randomized (non-randomized, respectively) policy is a control policy consisting in choosing randomly (deterministically, respectively) the actions along the time according to a probability law depending on the past history of the state and action processes. A Markov non-randomized policy is a sequence $(\varphi_j^i)_{j\in\mathbb{N}}$ of $\mathbf{A}_\Delta^i$-valued mappings on $\mathbf{X}_\Delta$, and so on. Observe that $P^\beta$ is in fact a stochastic kernel on $(\mathbf{X}_\Delta \times \mathbf{A}_\Delta^i)^\infty$ given $\mathbf{X}$, see Proposition C.10 in [13]. Since we only consider intervention as an element of $\mathbf{Y}$, we introduce $\Xi$ as the set of policies $\beta$ satisfying $P^\beta(\mathbf{Y}|x) = 1$. We consider randomized interventions and consequently an intervention is an element of

$$\mathcal{P}^\mathbf{Y} = \{\gamma \in \mathcal{P}(\mathbf{Y}|\mathbf{X}) : \gamma(\cdot|\cdot) = P^\beta(\cdot|\cdot) \text{ for some } \beta \in \Xi\},$$

and

$$\mathcal{P}^\mathbf{Y}(x) = \{\rho \in \mathcal{P}(\mathbf{Y}) : \rho(\cdot) = P^\beta(\cdot|x) \text{ for some } \beta \in \Xi\}$$

is the set of feasible interventions in state $x \in \mathbf{X}$. Observe that if an intervention is chosen in $\mathbf{Y}_0$, it means actually that the controller has not intervened on the process through impulsive actions. For technical reasons, it appears necessary to introduce the set $\mathbf{Y}^*$ of *real* interventions given by $\mathbf{Y}^* = \bigcup_{k=1}^\infty \mathbf{Y}_k$. The associated sets of real randomized interventions are defined by

$$\mathcal{P}^{\mathbf{Y}^*} = \{\gamma \in \mathcal{P}(\mathbf{Y}|\mathbf{X}) : \gamma(\cdot|\cdot) = P^\beta(\cdot|\cdot) \text{ for some } \beta \in \Xi$$
$$\text{and } P^\beta(\mathbf{Y}^*|x) = 1, \text{ for any } x \in \mathbb{X}^i\}$$

and

$$\mathcal{P}^{\mathbf{Y}^*}(x) = \{\rho \in \mathcal{P}(\mathbf{Y}) : \rho(\cdot) = P^\beta(\cdot|x) \text{ for some } \beta \in \Xi \text{ and } P^\beta(\mathbf{Y}^*|x) = 1\}$$

for $x \in \mathbf{X}$. Note that $\mathcal{P}^{\mathbf{Y}^*}(x) = \emptyset$ if $x \notin \mathbb{X}^i$.

Let us denote by $\mathcal{P}^g(\mathbf{A}^g|\mathbf{X})$, the set of stochastic kernels $\pi \in \mathcal{P}(\mathbf{A}^g|\mathbf{X})$ such that for any $x \in \mathbf{X}$, $\pi(\mathbf{A}^g(x)|x) = 1$ and by $\mathcal{P}^i(\mathbf{A}_\Delta^i|\mathbf{X}_\Delta)$, the set of stochastic kernels $\varphi \in \mathcal{P}(\mathbf{A}_\Delta^i|\mathbf{X}_\Delta)$ such that for any $x \in \mathbf{X}_\Delta$, $\varphi(\mathbf{A}_\Delta^i(x)|x) = 1$. For $\pi \in \mathcal{P}^g(\mathbf{A}^g|\mathbf{X})$, $\overline{q}^\pi$ denotes the kernel on $\mathbf{X}$ given $\mathbf{X}$ defined by $\int_{\mathbf{A}^g} \overline{q}(\Gamma|x, a)\pi(da|x)$ and $q^\pi$ denotes the transition rate on $\mathbf{X}$ given $\mathbf{X}$ defined by $\int_{\mathbf{A}^g} q(\Gamma|x, a)\pi(da|x)$ for any $\Gamma \in \mathcal{B}(\mathbf{X})$ and $x \in \mathbf{X}$. Similarly, for $\varphi \in \mathcal{P}^i(\mathbf{A}_\Delta^i|\mathbf{X}_\Delta)$, $Q_\Delta^\varphi$ denotes the stochastic kernel on $\mathbf{X}_\Delta$ given $\mathbf{X}_\Delta$ defined by $\int_{\mathbf{A}_\Delta^i} Q_\Delta(\Gamma|x, a)\varphi(da|x)$ for any $\Gamma \in \mathcal{B}(\mathbf{X}_\Delta)$ and $x \in \mathbf{X}_\Delta$.

At some point, we need to consider strategic measures for the model $\mathcal{M}^i$ generated by arbitrary randomized stationary policies not necessarily belonging to $\Xi$. Consequently, let us introduce the space

$$\mathbb{K}_\Delta^i = \{(x, a) \in \mathbf{X}_\Delta \times \mathbf{A}_\Delta^i : a \in \mathbf{A}_\Delta^i(x)\} \in \mathcal{B}(\mathbf{X}_\Delta \times \mathbf{A}_\Delta^i).$$

Let $\varphi \in \mathcal{P}^i(\mathbf{A}_\Delta^i | \mathbf{X}_\Delta)$, by a slight abuse of notation, let us denote by $\varphi$ the randomized stationary policy induced by the stochastic kernel $\varphi$ and by $P^\varphi$ the strategic measure for the model $\mathcal{M}^i$ generated by the policy $\varphi$. Clearly, $P^\varphi \in \mathcal{P}((\mathbb{K}_\Delta^i)^\infty | \mathbf{X})$.

Finally, we end this subsection by introducing a *projection* mapping that will be used repeatedly in the paper. If $y \in \mathbf{Y}$ then there exists a unique $k \in \mathbb{N}$ such that $y \in \mathbf{Y}_k$. The $\mathbf{X}$-valued mapping $\bar{x}$ on $\mathbf{Y}$ is defined by

$$\bar{x}(y) = x_{k+1}. \tag{1}$$

### 2.2 Construction of the Process

Having introduced the parameters of the model, we are now in position to construct the Markov controlled process. Let $\mathbf{Y}_\infty = \mathbf{Y} \cup \{y_\infty\}$ where $y_\infty$ is an artificial (isolated) point and $\Omega_n = \mathbf{Y} \times (\mathbb{R}_+^* \times \mathbf{Y})^n \times (\{\infty\} \times \{y_\infty\})^\infty$, for $n \in \mathbb{N}$. The canonical space $\Omega$ is defined as $\Omega = \bigcup_{n=1}^\infty \Omega_n \bigcup (\mathbf{Y} \times (\mathbb{R}_+^* \times \mathbf{Y})^\infty)$ and is endowed with its Borel $\sigma$-algebra denoted by $\mathcal{F}$. For notational convenience, $\omega \in \Omega$ will be represented as

$$\omega = (y_0, \theta_1, y_1, \theta_2, y_2, \ldots).$$

Here, $y_0 = (x_0, \Delta, \Delta, \ldots)$ is the initial state of the controlled point process $\xi$ with values in $\mathbf{Y}$, defined below; $\theta_1 = 0$ and $y_1 \in \mathbf{Y}$ is the result of the initial intervention. The components $\theta_n > 0$ for $n \geq 2$ mean the sojourn times; $y_n$ denotes the result of an intervention (if $y_n \in \mathbf{Y}^*$) or corresponds to a natural jump (if $y_n \in \mathbf{Y} \setminus \mathbf{Y}^*$). In case $\theta_n < \infty$ and $\theta_{n+1} = \infty$, the trajectory has only $n$ jumps and we put $y_m = y_\infty$ for all $m \geq n + 1$.

The path up to $n \in \mathbb{N}$ is denoted by $h_n = (y_0, \theta_1, y_1, \theta_2, y_2, \ldots \theta_n, y_n)$ and the collection of all such paths is denoted by $\mathbf{H}_n$. For $n \in \mathbb{N}$, introduce the mappings $Y_n : \Omega \to \mathbf{Y}_\infty$ by $Y_n(\omega) = y_n$ and, for $n \geq 2$, the mappings $\Theta_n : \Omega \to \overline{\mathbb{R}}_+^*$ by $\Theta_n(\omega) = \theta_n$; $\Theta_1(\omega) = 0$. The sequence $(T_n)_{n \in \mathbb{N}^*}$ of $\overline{\mathbb{R}}_+^*$-valued mappings is defined on $\Omega$ by $T_n(\omega) = \sum_{i=1}^n \Theta_i(\omega) = \sum_{i=1}^n \theta_i$ and $T_\infty(\omega) = \lim_{n \to \infty} T_n(\omega)$. For notational convenience, we denote by $H_n = (Y_0, \Theta_1, Y_1, \ldots, \Theta_n, Y_n)$ the $n$-term history process taking values in $\mathbf{H}_n$ for $n \in \mathbb{N}$.

The random measure $\mu$ associated with $(\Theta_n, Y_n)_{n \in \mathbb{N}}$ is a measure defined on $\mathbb{R}_+^* \times \mathbf{Y}$ by

$$\mu(\omega; dt, dy) = \sum_{n \geq 2} I_{\{T_n(\omega) < \infty\}} \delta_{(T_n(\omega), Y_n(\omega))}(dt, dy).$$

For notational convenience the dependence on $\omega$ will be ignored and instead of $\mu(\omega; dt, dy)$ it will be written $\mu(dt, dy)$. For $t \in \mathbb{R}_+$, define $\mathcal{F}_t = \sigma\{H_1\} \vee \sigma\{\mu(]0, s] \times B) : s \leq t, B \in \mathcal{B}(\mathbf{Y})\}$. Finally, we define the controlled process $\{\xi_t\}_{t \in \mathbb{R}_+}$:

$$\xi_t(\omega) = \begin{cases} Y_n(\omega), & \text{if } T_n \leq t < T_{n+1} \quad \text{for } n \in \mathbb{N}^*; \\ y_\infty, & \text{if } T_\infty \leq t, \end{cases}$$

and $\xi_{0-}(\omega) = Y_0 = y_0$ with $y_0 = (x_0, \Delta, \Delta, \ldots)$. Obviously, the controlled process $\{\xi_t\}_{t \in \mathbb{R}_+}$ can be equivalently described by the sequence $(\Theta_n, Y_n)_{n \in \mathbb{N}}$. The sequence $(T_n)_{n \in \mathbb{N}^*}$ describes the times of jumps of $\{\xi_t\}_{t \in \mathbb{R}_+}$: $T_n$ is the $n$-th jump moment. The state $\xi_t$ is constant between the jump times $T_n$ and $T_{n+1}$ and represents the successive jumps of the process and the associated impulsive actions at time $T_n$. This choice for the process $\{\xi_t\}_{t \in \mathbb{R}_+}$ is motivated by the fact we consider models with possibly multiple impulses at the same time moment. In such a framework, we extend the state space from $\mathbf{X}$ to $\mathbf{Y}$ in order to include the sequence of successive instantaneous jumps and corresponding impulsive actions. The extended state is of the form $(x_0, a_0, x_1, a_1, \ldots, x_k, a_k, x_{k+1}, \Delta, \Delta, \ldots)$ where $x_0$ corresponds to a possibly natural jump or to the value of the process just before the intervention. The triple $(x_j, a_j, x_{j+1})$ indicates that the impulsive action $a_j$ has been applied to the system at state $x_j$ leading to a new state $x_{j+1}$ having distribution $Q(\cdot|x_j, a_j)$. The special impulsive action $\Delta$ means that the impulses are over and the artificial state $\Delta$ means the same. Observe that the last component different from $\Delta$ in $\xi_t$ corresponds to the last position of the process after a sequence of successive instantaneous jumps and impulses. It is given by $\overline{x}(\xi_t)$. It may appear odd to have included the impulsive actions (associated to the successive jumps) in the state process $\xi_t$. However, this choice was made to simplify the model describing the dynamic of the process. Other approaches would have been possible as for example, a model where $\{\xi_t\}_{t \in \mathbb{R}_+}$ would have been only constituted of the successive jumps but it would have induced a much more complicated mathematical description for the dynamic with heavy notation.

### 2.3 Admissible Strategies and Conditional Distribution of the Controlled Process

An admissible (randomized) control strategy is a sequence $u = (u_n)_{n \in \mathbb{N}}$ such that $u_0 \in \mathcal{P}^{\mathbf{Y}}(x_0)$ and, for any $n \in \mathbb{N}^*$, $u_n$ is given by

$$u_n = (\psi_n, \pi_n, \gamma_n^0, \gamma_n^1),$$

where $\psi_n$ is a stochastic kernel on $\overline{\mathbb{R}}_+^*$ given $\mathbf{H}_n$ satisfying $\psi_n(\cdot|h_n) = \delta_{+\infty}(\cdot)$ for any $h_n = (y_0, \theta_1, \ldots \theta_n, y_n) \in \mathbf{H}_n$ with $\overline{x}(y_n) \notin \mathbb{X}^i$, $\pi_n$ is a stochastic kernel on $\mathbf{A}^g$ given $\mathbf{H}_n \times \mathbb{R}_+^*$ satisfying $\pi_n(\mathbf{A}^g(\overline{x}(y_n))|h_n, t) = 1$ for any $t \in \mathbb{R}_+^*$ and $h_n = (y_0, \theta_1, \ldots \theta_n, y_n) \in \mathbf{H}_n$, $\gamma_n^0$ is a stochastic kernel on $\mathbf{Y}$ given $\mathbf{H}_n \times \mathbb{R}_+^* \times \mathbf{X}$ satisfying $\gamma_n^0(\cdot|h_n, t, x) \in \mathcal{P}^{\mathbf{Y}}(x)$ for any $h_n \in \mathbf{H}_n$, $t \in \mathbb{R}_+^*$ and $x \in \mathbf{X}$, and $\gamma_n^1$ is a stochastic kernel on $\mathbf{Y}$ given $\mathbf{H}_n$ satisfying $\gamma_n^1(\cdot|h_n) \in \mathcal{P}^{\mathbf{Y}^*}(\overline{x}(y_n))$ for any

$h_n = (y_0, \theta_1, \ldots \theta_n, y_n) \in \mathbf{H}_n$ with $\overline{x}(y_n) \in \mathbb{X}^i$; if $\overline{x}(y_n) \notin \mathbb{X}^i$ then $\gamma_n^1(\cdot|h_n) = \delta_{(\overline{x}(y_n), \Delta, \Delta, \ldots)}(\cdot)$.

The above conditions apply when $y_n \neq y_\infty$; otherwise, all the values of $\psi_n(\cdot|h_n)$, $\pi_n(\cdot|h_n, t)$, $\gamma_n^0(\cdot|h_n, t, \cdot)$ and $\gamma_n^1(\cdot|h_n)$ may be arbitrary.

The set of admissible control strategies is denoted by $\mathcal{U}$. In what follows, we use notation $\gamma_n = (\gamma_n^0, \gamma_n^1)$. An admissible strategy $u$ with $u_n = (\psi_n, \pi_n, \gamma_n^0, \gamma_n^1)$ for $n \in \mathbb{N}^*$ is called randomized stationary, if there exist $\psi \in \overline{\mathbb{M}}_+^*(\mathbf{X})$, $\pi \in \mathcal{P}^g(\mathbf{A}^g|\mathbf{X})$, a stochastic kernel $\gamma^0$ on $\mathbf{Y}$ given $\mathbf{X}$ such that $\gamma^0(\cdot|x) \in \mathcal{P}^{\mathbf{Y}}(x)$ for any $x \in \mathbf{X}$, and a stochastic kernel $\gamma^1$ on $\mathbf{Y}$ given $\mathbf{X}$ such that $\gamma^1(\cdot|x) \in \mathcal{P}^{\mathbf{Y}^*}(x)$ for any $x \in \mathbb{X}^i$ satisfying $u_0(\cdot) = \gamma^0(\cdot|x_0)$, $\psi_n(\cdot|h_n) = \delta_{\psi(\overline{x}(y_n))}(\cdot)$, $\pi_n(\cdot|h_n, t) = \pi(\cdot|\overline{x}(y_n))$, $\gamma_n^0(\cdot|h_n, t, x) = \gamma^0(\cdot|x)$, and $\gamma_n^1(\cdot|h_n) = \gamma^1(\cdot|\overline{x}(y_n))$ when $\overline{x}(y_n) \in \mathbb{X}^i$.

Roughly speaking, $\psi_n$ represents the conditional time distribution of the next possible intervention after time $T_n$, $\pi_n$ is the usual continuous control influencing the intensity of the jumps $q$ between $T_n$ and $T_{n+1}$, $\gamma_n^0$ is the distribution of the next intervention if it is decided to have an intervention just immediately after a natural jump and $\gamma_n^1$ is the distribution of the next intervention if it is decided to have an intervention before a natural jump.

Suppose a strategy $u = (u_n)_{n \in \mathbb{N}} \in \mathcal{U}$ is fixed with $u_n = (\psi_n, \pi_n, \gamma_n^0, \gamma_n^1)$ for $n \in \mathbb{N}^*$. We introduce the intensity of the natural jumps

$$\lambda_n(\Gamma_x, h_n, t) = \int_{\mathbf{A}^g} \overline{q}(\Gamma_x|\overline{x}(y_n), a)\pi_n(da|h_n, t),$$

and the rate of the natural jumps

$$\Lambda_n(\Gamma_x, h_n, t) = \int_{]0, t[} \lambda_n(\Gamma_x, h_n, s)ds$$

for any $n \in \mathbb{N}^*$, $\Gamma_x \in \mathcal{B}(\mathbf{X})$, $h_n = (y_0, \theta_1, y_1, \ldots, \theta_n, y_n) \in \mathbf{H}_n$ and $t \in \overline{\mathbb{R}}_+^*$. Now, for any $n \in \mathbb{N}^*$, the stochastic kernel $G_n$ on $\mathbf{Y}_\infty \times \overline{\mathbb{R}}_+^*$ given $\mathbf{H}_n$ is defined by

$$G_n(\{+\infty\} \times \{y_\infty\}|h_n) = \delta_{y_n}(\{y_\infty\}) + \delta_{y_n}(\mathbf{Y})e^{-\Lambda_n(\mathbf{X}, h_n, +\infty)}\psi_n(\{+\infty\}|h_n) \quad (2)$$

and

$$G_n(\Gamma_\Theta \times \Gamma_y|h_n) = \delta_{y_n}(\mathbf{Y}) \Big[ \gamma_n^1(\Gamma_y|h_n) \int_{\Gamma_\Theta} e^{-\Lambda_n(\mathbf{X}, h_n, t)}\psi_n(dt|h_n)$$
$$+ \int_{\Gamma_\Theta} \int_{\mathbf{X}} \psi_n([t, \infty]|h_n)\gamma_n^0(\Gamma_y|h_n, t, x)\lambda_n(dx, h_n, t)e^{-\Lambda_n(\mathbf{X}, h_n, t)}dt \Big],$$
$$(3)$$

and

$$G_n(\{+\infty\} \times \Gamma_y|h_n) = 0, \quad (4)$$

where $\Gamma_y \in \mathcal{B}(\mathbf{Y})$, $\Gamma_\Theta \in \mathcal{B}(\mathbb{R}_+^*)$ and $h_n = (y_0, \theta_1, y_1, \ldots, \theta_n, y_n) \in \mathbf{H}_n$. Note that the kernel $\gamma_n^1$ does not appear in the formula for $G_n$ if $\overline{x}(y_n) \notin \mathbb{X}^i$.

Consider an admissible strategy $u \in \mathcal{U}$ and an initial state $x_0 \in \mathbf{X}$. Recalling that $\Omega = \bigcup_{n=1}^\infty \Omega_n \bigcup \left(\mathbf{Y} \times (\mathbb{R}_+^* \times \mathbf{Y})^\infty\right)$ and that $\mathcal{F}$ denotes its associated Borel $\sigma$-algebra, Theorem 3.6 in [17] (or Remark 3.43 in [18] or [19]) implies the existence of a probability $\mathbb{P}_{x_0}^u$ on $(\Omega, \mathcal{F})$ such that the restriction of $\mathbb{P}_{x_0}^u$ to $(\Omega, \mathcal{F}_0)$ is given by

$$\mathbb{P}_{x_0}^u\left(\{Y_0\} \times \{0\} \times \Gamma_y \times (\overline{\mathbb{R}}_+^* \times \mathbf{Y}_\infty)^\infty\right) = u_0(\Gamma_y | x_0) \tag{5}$$

for any $\Gamma_y \in \mathcal{B}(\mathbf{Y})$ and the positive random measure $\nu$ defined on $\mathbb{R}_+^* \times \mathbf{Y}$ by

$$\nu(dt, dy) = \sum_{n \in \mathbb{N}^*} \frac{G_n(dt - T_n, dy | H_n)}{G_n([t - T_n, +\infty] \times \mathbf{Y}_\infty | H_n)} I_{\{T_n < t \leq T_{n+1}\}} \tag{6}$$

is the predictable projection of $\mu$ with respect to $\mathbb{P}_{x_0}^u$.

*Remark 2.1* Observe that $\mathcal{F}_{T_n}$ is the $\sigma$-algebra generated by the random variable $H_n$ for $n \in \mathbb{N}^*$. The conditional distribution of $(Y_{n+1}, \Theta_{n+1})$ given $\mathcal{F}_{T_n}$ under $\mathbb{P}_{x_0}^u$ is determined by $G_n(\cdot | H_n)$ and the conditional survival function of $\Theta_{n+1}$ given $\mathcal{F}_{T_n}$ under $\mathbb{P}_{x_0}^u$ is given by $G_n([t, +\infty] \times \mathbf{Y}_\infty | H_n)$.

## 3 Optimization Problem and Assumptions

The objective of this section is to introduce the infinite-horizon performance criteria we are concerned with. Next we state our assumptions on the parameters of the model. Moreover, in the context of these hypotheses, at the end of this section we recall a technical result from [8] providing a decomposition of the predictable projection $\nu$ of the measure $\mu$ in terms of the distributions $(\gamma_n^0)_{n \in \mathbb{N}^*}$ and $(\gamma_n^1)_{n \in \mathbb{N}^*}$ of the next intervention.

We consider an optimization problem with $p \in \mathbb{N}$ constraints where the performance and the constraint criteria are given in terms of infinite-horizon discounted functionals. In order to define these criteria, we need to introduce the cost rates $\left(C_j^g\right)_{j \in \mathbb{N}_p}$ associated with continuous actions. For any $j \in \mathbb{N}_p$, the real-valued mapping $C_j^g$ is defined on $\mathbb{K}^g$. The costs $\left(C_j^i\right)_{j \in \mathbb{N}_p}$ associated with an intervention $y = (x_0, a_0, x_1, a_1, \ldots) \in \mathbf{Y}$ are given by $C_j^i(y) = \sum_{k \in \mathbb{N}} c_j^i(x_k, a_k)$, where for any $j \in \mathbb{N}_p$, $c_j^i$ is a non-negative real-valued mapping defined on $\mathbb{K}_\Delta^i$ satisfying $c_j^i(x, a) = 0$ if $(x, a) \notin \mathbb{K}^i$. For any $(x, a) \in \mathbb{K}^i$ and $j \in \mathbb{N}_p$, $c_j^i(x, a)$ corresponds to the cost associated with a single jump at $x \in \mathbf{X}$ resulting from the impulsive action $a \in \mathbf{A}^i(x)$. The cost associated with a randomized intervention $\rho \in \mathcal{P}^{\mathbf{Y}}(x)$ for $x \in \mathbf{X}$ is given by $\int_{\mathbf{Y}} C_j^i(y) \rho(dy | x)$ for any $j \in \mathbb{N}_p$. Therefore, the infinite-horizon discounted performance criteria corresponding to an admissible control strategy $u \in \mathcal{U}$ are defined by

$$\mathcal{V}_j(u, x_0) = \int_{\mathbf{Y}} C_j^i(y) u_0(dy|x_0) + \mathbb{E}_{x_0}^u \left[ \int_0^{+\infty} e^{-\alpha s} \int_{\mathbf{A}^g} C_j^g(\overline{x}(\xi_{s-}), a) \pi(da|s) ds \right]$$

$$+ \mathbb{E}_{x_0}^u \left[ \int_{]0,\infty[ \times \mathbf{Y}} e^{-\alpha s} C_j^i(y) \mu(ds, dy) \right], \tag{7}$$

for any $j \in \mathbb{N}_p$. In the previous expression, $\alpha > 0$ is the discount factor. Note that, the performance criteria are well defined under Assumption A imposed below.

**Definition 3.1** The constrained optimization problem consists to minimize $\mathcal{V}_0(u, x_0)$ within the class of admissible strategies $u \in \mathcal{U}$ where $x_0$ is the initial state and such that $\mathcal{V}_j(u, x_0) \leq B_j$, for any $j \in \mathbb{N}_p^*$ where $(B_j)_{j \in \mathbb{N}_p^*}$ are nonnegative real numbers representing the constraint bounds. The class of feasible strategies $u \in \mathcal{U}$ will be denoted by $\mathcal{U}^f = \left\{ u \in \mathcal{U} : \mathcal{V}_j(u, x_0) \leq B_j, \text{ for any } j \in \mathbb{N}_p^* \right\}$.

**Assumption A** There exists a constant $K \in \mathbb{R}_+$ such that for any $x \in \mathbf{X}, a^g \in \mathbf{A}^g(x)$, $a^i \in \mathbf{A}^i(x)$ and $j \in \mathbb{N}_p$

(A1) $\overline{q}(\mathbf{X}|x, a^g) \leq K$.
(A2) $C_j^g(x, a^g) \geq 0$.
(A3) $c_j^i(x, a^i) \geq 0$.

*Remark 3.1* It must be emphasized that Assumption (A2) can be replaced by the following apparently weaker condition $C_j^g(x, a^g) \geq -K$.

The purpose of the next assumption is to avoid infinite simultaneous interventions. This is a classical hypothesis in the framework of impulsive control problems, see for example [6].

**Assumption B** There exists a constant $\underline{c} > 0$ such that $\sum_{j \in \mathbb{N}_p} c_j^i(x, a) \geq \underline{c}$ for any $(x, a) \in \mathbb{K}^i$.

Finally, let us recall the following technical result from [8, Lemma 3.1]

**Lemma 3.2** *The predictable projection of the random measure $\mu$ is given by $\nu = \nu_0 + \nu_1$ where*

$$\nu_0(\Gamma, \Gamma_y) = \int_{\Gamma} \int_{\mathbf{A}^g} \int_{\mathbf{X}} \gamma^0(\Gamma_y|x, s) \overline{q}(dx|\overline{x}(\xi_{s-}), a) \pi(da|s) ds,$$

$$\nu_1(\Gamma, \Gamma_y) = \sum_{n \in \mathbb{N}^*} \gamma_n^1(\Gamma_y|H_n) \int_{\Gamma} I_{\{T_n < s \leq T_{n+1}\}} \frac{\psi_n(ds - T_n|H_n)}{\psi_n([s - T_n, +\infty]|H_n)},$$

$$\gamma^0(dy|x, t) = \sum_{n \in \mathbb{N}^*} I_{\{T_n < t \leq T_{n+1}\}} \gamma_n^0(dy|H_n, t - T_n, x), \quad \pi(da|t) = \sum_{n \in \mathbb{N}^*} I_{\{T_n < t \leq T_{n+1}\}} \pi_n$$

$(da|H_n, t - T_n)$, *for any* $\Gamma \in \mathcal{B}(\mathbb{R}_+^*), \Gamma_y \in \mathcal{B}(\mathbf{Y}), t \in \mathbb{R}_+$.

## 4 Occupation Measures and Their Properties

In this section, we introduce the definition of an occupation measure $\eta_u$ induced by a control strategy $u \in \mathcal{U}$ (see Definition 4.2). The objective of this section is twofold. First, it is shown in Theorem 4.5 that for any control strategy $u \in \mathcal{U}$, the corresponding occupation measure $\eta_u$ satisfies a linear equation depending on the kernels $q$ and $Q$. A measure satisfying such equation will be called admissible. Second, from any admissible measure $\eta$ one can construct a randomized stationary control strategy $u \in \mathcal{U}$ such that the restrictions of $\eta$ and $\eta_u$ to $\mathbb{K}^g$ are equal and the restriction of $\eta_u$ to $\mathbb{K}^i$ is smaller than the restriction of $\eta$ to $\mathbb{K}^i$ (see Theorem 4.9). These two important results will play an important role in the next section to establish a connection between the constrained optimal control problem and the linear program given in Definition 5.1. In order to prove these two results, one needs to provide auxiliary results whose proofs are deferred to the Appendices 1 and 2 to streamline the presentation.

**Definition 4.1** To any $\gamma \in \mathcal{M}(\mathbf{Y})$, we associate the measure $\widetilde{\gamma} \in \mathcal{M}(\mathbf{X}_\Delta \times \mathbf{A}_\Delta^i)$ defined by $\widetilde{\gamma}(\Gamma) = \sum_{j=1}^{\infty} \gamma\big(\{\mathbf{y} \in \mathbf{Y} : \mathbf{y}_j \in \Gamma\}\big)$, for any $\Gamma \in \mathcal{B}(\mathbf{X}_\Delta \times \mathbf{A}_\Delta^i)$ and where $\mathbf{y}_j$ is the $j$th coordinate of $\mathbf{y} \in \mathbf{Y} \subset (\mathbb{K}_\Delta^i)^\infty$. Similarly, if $R$ is a stochastic kernel on $\mathbf{Y}$ given a Borel space $\mathbf{Z}$ then $\widetilde{R}$ is a kernel on $\mathbf{X}_\Delta \times \mathbf{A}_\Delta^i$ given $\mathbf{Z}$ defined by

$$\widetilde{R}(\Gamma|z) = \sum_{j=1}^{\infty} R\big(\{\mathbf{y} \in \mathbf{Y} : \mathbf{y}_j \in \Gamma\}|z\big), \text{ for any } z \in \mathbf{Z} \text{ and } \Gamma \in \mathcal{B}(\mathbf{X}_\Delta \times \mathbf{A}_\Delta^i).$$

These definitions will be naturally extended to probability measures defined on $(\mathbb{K}_\Delta^i)^\infty$ and to stochastic kernels on $(\mathbb{K}_\Delta^i)^\infty$ given a Borel space $\mathbf{Z}$. Now, we introduce the definition of an occupation measure $\eta_u$ induced by an admissible control strategy $u$.

**Definition 4.2** For a strategy $u = (u_n)_{n \in \mathbb{N}} \in \mathcal{U}$, let us introduce the measures $\eta_u^g$ (respectively, $\mu_u^i$ and $\eta_u^i$) defined on $\mathbf{X} \times \mathbf{A}^g$ (respectively, $\mathbf{Y}$ and $\mathbf{X}_\Delta \times \mathbf{A}_\Delta^i$) by

$$\eta_u^g(dx, da) = \alpha \mathbb{E}_{x_0}^u \left[ \int_0^{T_\infty} e^{-\alpha s} \delta_{\overline{x}(\xi_{s-})}(dx) \pi(da|s) ds \right], \tag{8}$$

$$\mu_u^i(dy) = \mathbb{E}_{x_0}^u \left[ \int_{]0, T_\infty[} e^{-\alpha s} \mu(ds, dy) \right] + u_0(dy|x_0), \tag{9}$$

and

$$\eta_u^i(dx, da) = \widetilde{\mu}_u^i(dx, da). \tag{10}$$

Actually, the measure $\eta_u^g$ is supported on $\mathbb{K}^g$ and clearly finite for any $u \in \mathcal{U}$, and the measure $\eta_u^i$ is supported on $\mathbb{K}_\Delta^i$. Then, the measure $\eta_u$ defined on $\mathbf{X} \times \mathbf{A}$ by

$$\eta_u(\Gamma) = \eta_u^g(\Gamma \cap \mathbb{K}^g) + \eta_u^i(\Gamma \cap \mathbb{K}^i), \tag{11}$$

for any $\Gamma \in \mathcal{B}(\mathbf{X} \times \mathbf{A})$ is called the occupation measure of the controlled process induced by the control strategy $u$. Clearly, the measure $\eta_u$ is supported on $\mathbb{K}^g \cup \mathbb{K}^i$.

The infinite-horizon discounted performance criteria corresponding to an admissible control strategy $u \in \mathcal{U}$ satisfying $\mathbb{P}^u_{x_0}(T_\infty = +\infty) = 1$ can be written in terms of the measure $\eta_u$ as follows

$$\mathcal{V}_j(u, x_0) = \eta_u(C_j) \tag{12}$$

with $C_j(x, a) = \frac{1}{\alpha} C^g_j(x, a) I_{\mathbb{K}^g}(x, a) + c^i_j(x, a) I_{\mathbb{K}^i}(x, a)$, for $j \in \mathbb{N}_p$.

In order to show the first main result of this section, Theorem 4.6, we need to derive two intermediate results: Propositions 4.3 and 4.4. For the sake of clarity in the exposition, their proofs are presented in Appendix 1. Roughly speaking, these two technical results establish links between the measures $\eta^g_u$ and $\eta^i_u$.

**Proposition 4.3** *Consider a strategy* $u = (u_n)_{n \in \mathbb{N}} \in \mathcal{U}$ *fixed with* $u_n = (\psi_n, \pi_n, \gamma^0_n, \gamma^1_n)$ *for* $n \in \mathbb{N}^*$ *satisfying* $\eta^i_u(\mathbf{X} \times \mathbf{A}^i_\Delta) < \infty$. *Then, for any* $\Gamma \in \mathcal{B}(\mathbf{X})$,

$$\eta^g_u(\Gamma \times \mathbf{A}^g) = \eta^i_u(\Gamma \times \{\Delta\}) - \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} I_\Gamma(x) \overline{q}(\mathbf{X}|x, a) \eta^g_u(dx, da)$$

$$- \mathbb{E}^u_{x_0} \left[ \sum_{n \in \mathbb{N}^*} e^{-\alpha T_n} I_\Gamma(\overline{x}(Y_n)) \int_{]0,\infty[} e^{-\alpha s} \psi_n(ds|H_n) \right]. \tag{13}$$

*Proof* See Appendix 1. □

**Proposition 4.4** *Consider a strategy* $u = (u_n)_{n \in \mathbb{N}} \in \mathcal{U}$ *fixed with* $u_n = (\psi_n, \pi_n, \gamma^0_n, \gamma^1_n)$ *for* $n \in \mathbb{N}^*$ *satisfying* $\mathbb{P}^u_{x_0}(T_\infty = +\infty) = 1$. *Then, for any* $\Gamma \in \mathcal{B}(\mathbf{X}_\Delta)$

$$\eta^i_u(\Gamma \times \mathbf{A}^i_\Delta) = \delta_{x_0}(\Gamma) + \int_{\mathbf{X}_\Delta \times \mathbf{A}^i_\Delta} Q_\Delta(\Gamma|z, b) \eta^i_u(dz, db)$$

$$+ \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} \overline{q}(\Gamma \cap \mathbf{X}|x, a) \eta^g_u(dx, da)$$

$$+ \mathbb{E}^u_{x_0} \left[ \sum_{n \in \mathbb{N}^*} e^{-\alpha T_n} I_\Gamma(\overline{x}(Y_n)) \int_{]0,\infty[} e^{-\alpha s} \psi_n(ds|H_n) \right]. \tag{14}$$

*Proof* See Appendix 1. □

*Remark 4.1* Observe that in the previous result, if we consider $\Gamma$ in $\mathcal{B}(\mathbf{X})$ then Eq. (14) becomes

$$\eta^i_u(\Gamma \times \mathbf{A}^i_\Delta) = \delta_{x_0}(\Gamma) + \int_{\mathbf{X} \times \mathbf{A}^i} Q_\Delta(\Gamma|z, b) \eta^i_u(dz, db)$$

$$+ \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} \overline{q}(\Gamma|x, a) \eta^g_u(dx, da)$$

$$+ \mathbb{E}_{x_0}^u \Bigg[ \sum_{n \in \mathbb{N}^*} e^{-\alpha T_n} I_\Gamma(\overline{x}(Y_n)) \int_{]0,\infty[} e^{-\alpha s} \psi_n(ds|H_n) \Bigg], \qquad (15)$$

since $Q_\Delta(\Gamma|x, a) = 0$ for any $(x, a) \notin \mathbb{K}^i$.

**Definition 4.5** A measure $\rho \in \mathcal{M}_f(\mathbf{X} \times \mathbf{A})$ is said to be *admissible* if $\rho$ is concentrated on $\mathbb{K}^g \cup \mathbb{K}^i$ and

$$\rho(\Gamma \times \mathbf{A}) = \delta_{x_0}(\Gamma) + \int_{\mathbf{X} \times \mathbf{A}^i} Q_\Delta(\Gamma|z, b)\rho(dz, db) + \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} q(\Gamma|x, a)\rho(dx, da),$$
$$(16)$$

for any $\Gamma \in \mathcal{B}(\mathbf{X})$.

The next result shows that any occupation measure is admissible.

**Theorem 4.6** *Consider a strategy* $u = (u_n)_{n \in \mathbb{N}} \in \mathcal{U}$ *fixed with* $u_n = (\psi_n, \pi_n, \gamma_n^0, \gamma_n^1)$ *for* $n \in \mathbb{N}^*$ *satisfying* $\eta_u^i(\mathbf{X} \times \mathbf{A}_\Delta^i) < \infty$. *Then, the measure* $\eta_u$ *is admissible.*

*Proof* First of all, recall that $\eta_u$ is finite and notice that $q(\Gamma|x, a) = \overline{q}(\Gamma|x, a) - I_\Gamma(x)\overline{q}(\mathbf{X}|x, a)$ for any $\Gamma \in \mathcal{B}(\mathbf{X})$ and $(x, a) \in \mathbb{K}^g$. Now consider $\Gamma \in \mathcal{B}(\mathbf{X})$, then by adding Eqs. (13) and (15), it yields that

$$\eta_u(\Gamma \times \mathbf{A}) = \eta_u^g(\Gamma \times \mathbf{A}^g) + \eta_u^i(\Gamma \times \mathbf{A}^i)$$
$$= \delta_{x_0}(\Gamma) + \int_{\mathbf{X} \times \mathbf{A}^i} Q_\Delta(\Gamma|z, b)\eta_u^i(dz, db) + \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} q(\Gamma|x, a)\eta_u^g(dx, da)$$
$$= \delta_{x_0}(\Gamma) + \int_{\mathbf{X} \times \mathbf{A}^i} Q_\Delta(\Gamma|z, b)\eta_u(dz, db) + \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} q(\Gamma|x, a)\eta_u(dx, da),$$

showing the result.                                                                                           □

Let $\varphi$ be a randomized stationary policy for the model $\mathcal{M}^i$. Introduce the set

$$\mathcal{S}_\varphi = \big\{ x \in \mathbf{X} : \widetilde{P}^\varphi(\mathbf{X} \times \{\Delta\}|x) = 1 \big\}, \qquad (17)$$

and the stochastic kernel $R^\varphi$ on $\mathbf{Y}$ given $\mathbf{X}$ by

$$R^\varphi(dy|x) = P^\varphi(dy|x)I_{\mathcal{S}_\varphi}(x) + \delta_{(x, \Delta, \Delta, \dots)}(dy)I_{\mathcal{S}_\varphi^c}(x). \qquad (18)$$

We introduce now a special class of randomized stationary control strategies.

**Definition 4.7** Consider $\pi \in \mathcal{P}^g(\mathbf{A}^g|\mathbf{X})$ and $\varphi \in \mathcal{P}^i(\mathbf{A}_\Delta^i|\mathbf{X}_\Delta)$. The strategy $u^{\pi,\varphi} = (u_n^{\pi,\varphi})_{n \in \mathbb{N}}$ is defined by $u_0^{\pi,\varphi}(\cdot) = R^\varphi(\cdot|x_0)$ and by $u_n^{\pi,\varphi} = (\psi_n, \pi_n, \gamma_n^0, \gamma_n^1)$ for $n \in \mathbb{N}^*$ where for any $x \in \mathbf{X}$, $t \in \mathbb{R}_+$ and $h_n = (y_0, \theta_1, \dots \theta_n, y_n) \in \mathbf{H}_n$, $\psi_n(\cdot|h_n) = \delta_{+\infty}(\cdot)$, $\pi_n(\cdot|h_n, t) = \pi(\cdot|\overline{x}(y_n))$, $\gamma_n^0(\cdot|h_n, t, x) = R^\varphi(\cdot|x)$. Finally, $\gamma_n^1$ is defined by $\gamma_n^1(\cdot|h_n) = \gamma^1(\cdot|\overline{x}(y_n))$ where $\gamma^1$ is an arbitrary stochastic kernel on $\mathbf{Y}$ given $\mathbf{X}$ satisfying $\gamma^1(\cdot|x) \in \mathcal{P}^{\mathbf{Y}^*}(x)$ for $x \in \mathbb{X}^i$ and $\gamma^1(\cdot|x) = \delta_{(x, \Delta, \Delta, \dots)}(\cdot)$ otherwise.

Clearly, the strategy so defined satisfies $u^{\pi,\varphi} \in \mathcal{U}$. The following proposition provides important properties of the measures $\eta^g_{u^{\pi,\varphi}}$ and $\eta^i_{u^{\pi,\varphi}}$ that will be needed in the proof of Theorem 4.9.

**Proposition 4.8** *Consider* $\pi \in \mathcal{P}^g(\mathbf{A}^g|\mathbf{X})$ *and* $\varphi \in \mathcal{P}^i(\mathbf{A}^i_\Delta|\mathbf{X}_\Delta)$. *Then*

$$\eta^i_{u^{\pi,\varphi}}(dx, da) = \left[ \delta_{x_0} + \frac{1}{\alpha} \widehat{\eta}^g_{u^{\pi,\varphi}} \overline{q}^\pi \right] \widetilde{R}^\varphi(dx, da), \tag{19}$$

*and for any* $\Gamma \in \mathcal{B}(\mathbf{X})$

$$\widehat{\eta}^g_{u^{\pi,\varphi}}(\Gamma) = \frac{1}{\alpha} \widehat{\eta}^g_{u^{\pi,\varphi}} r^\pi_\varphi(\Gamma) + \widetilde{R}^\varphi(\Gamma \times \{\Delta\}|x_0), \tag{20}$$

*where* $\widehat{\eta}^g_{u^{\pi,\varphi}}(dx)$ *denotes* $\eta^g_{u^{\pi,\varphi}}(dx, \mathbf{A}^g)$ *and the transition rate* $r^\pi_\varphi$ *on* $\mathbf{X}$ *given* $\mathbf{X}$ *is defined by*

$$r^\pi_\varphi(\Gamma|x) = \overline{q}^\pi \widetilde{R}^\varphi(\Gamma \times \{\Delta\}) - I_\Gamma(x) \overline{q}^\pi(\mathbf{X}|x). \tag{21}$$

*Proof* See Appendix 2. ☐

The following theorem is the second main result of this section. Roughly speaking, it can be seen as a converse of Theorem 4.6. In particular, it shows that an admissible measure $\eta$ may not be necessarily an occupation measure but one can construct from $\eta$ a randomized stationary control strategy $u$ such that the corresponding occupation measure $\eta_u$ is *smaller* than $\eta$ (see Eq. 26).

**Theorem 4.9** *Let* $\eta$ *be an admissible measure. Let us define the measure* $\eta^g$ *on* $\mathbf{X} \times \mathbf{A}^g$ *by*

$$\eta^g(\Gamma) = \eta(\Gamma \cap \mathbb{K}^g), \tag{22}$$

*for any* $\Gamma \in \mathcal{B}(\mathbf{X} \times \mathbf{A}^g)$ *and the measure* $\eta^i$ *on* $\mathbf{X} \times \mathbf{A}^i_\Delta$ *by*

$$\eta^i(\Gamma) = \eta(\Gamma \cap \mathbb{K}^i), \tag{23}$$

*for any* $\Gamma \in \mathcal{B}(\mathbf{X} \times \mathbf{A}^i)$ *and*

$$\eta^i(\Gamma \times \{\Delta\}) = \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} I_\Gamma(x) \overline{q}(\mathbf{X}|x, a) \eta^g(dx, da) + \eta^g(\Gamma \times \mathbf{A}^g),$$

*for any* $\Gamma \in \mathcal{B}(\mathbf{X})$. *Then, there exist a stochastic kernel* $\pi \in \mathcal{P}^g(\mathbf{A}^g|\mathbf{X})$ *satisfying*

$$\eta^g(\Gamma) = \int_\Gamma \pi(da|x) \eta^g(dx, \mathbf{A}^g), \tag{24}$$

*for any* $\Gamma \in \mathcal{B}(\mathbf{X} \times \mathbf{A}^g)$ *and a stochastic kernel* $\varphi \in \mathcal{P}^i(\mathbf{A}_\Delta^i | \mathbf{X}_\Delta)$ *satisfying*

$$\eta^i(\Gamma) = \int_\Gamma \varphi(da|x)\eta^i(dx, \mathbf{A}_\Delta^i) \tag{25}$$

*for any* $\Gamma \in \mathcal{B}(\mathbf{X} \times \mathbf{A}_\Delta^i)$ *and* $\varphi(\{\Delta\}|\Delta) = 1$. *Then,*

$$\eta_{u^{\pi,\varphi}}^g = \eta^g \ and \ \eta_{u^{\pi,\varphi}}^i \leq \eta^i \tag{26}$$

*and*

$$\eta_{u^{\pi,\varphi}}^i = \left[\delta_{x_0} + \frac{1}{\alpha}\eta_{u^{\pi,\varphi}}^g \overline{q}\right]\widetilde{P}^\varphi. \tag{27}$$

*Proof* First, notice that $\eta^i \in \mathcal{M}_f(\mathbf{X} \times \mathbf{A}_\Delta^i)$ and since $Q_\Delta(\mathbf{X}|x, a) = 1$ for any $(x, a) \in \mathbb{K}^i$ and $q(\mathbf{X}|x, a) = 0$ for any $(x, a) \in \mathbb{K}^g$, we obtain from Eq. (16) that $\eta^g \in \mathcal{P}(\mathbf{X} \times \mathbf{A}^g)$. Consequently, Proposition D.8 in [13, p. 184] ensures the existence of $\pi \in \mathcal{P}^g(\mathbf{A}^g|\mathbf{X})$ and $\varphi \in \mathcal{P}^i(\mathbf{A}_\Delta^i | \mathbf{X}_\Delta)$ satisfying respectively Eqs. (24) and (25). Consider an arbitrary set $\Gamma \in \mathcal{B}(\mathbf{X})$. For notational convenience, let us denote $\eta^i(\Gamma \times \mathbf{A}_\Delta^i)$ by $\widehat{\eta}^i(\Gamma)$ and $\eta^g(\Gamma \times \mathbf{A}^g)$ by $\widehat{\eta}^g(\Gamma)$. Observe that $Q_\Delta^\varphi(\Gamma|\{\Delta\}) = 0$. By using the definitions of $\eta^g$ and $\eta^i$ and Eq. (16) which is satisfied by $\eta$, it is easy to see that

$$\widehat{\eta}^g(\Gamma) = \eta^i(\Gamma \times \{\Delta\}) - \frac{1}{\alpha}\int_{\mathbf{X} \times \mathbf{A}^g} I_\Gamma(x)\overline{q}^\pi(\mathbf{X}|x, a)\widehat{\eta}^g(dx), \tag{28}$$

$$\widehat{\eta}^i(\Gamma) = \delta_{x_0}(\Gamma) + \frac{1}{\alpha}\widehat{\eta}^g \overline{q}^\pi(\Gamma) + \int_{\mathbf{X}} Q_\Delta^\varphi(\Gamma|z)\widehat{\eta}^i(dz) \tag{29}$$

for any $\Gamma \in \mathcal{B}(\mathbf{X})$ by recalling that $q(\Gamma|x, a) = \overline{q}(\Gamma|x, a) - I_\Gamma(x)\overline{q}(\mathbf{X}|x, a)$. Consequently,

$$\widehat{\eta}^i(\Gamma) = \delta_{x_0}(\Gamma) + \frac{1}{\alpha}\widehat{\eta}^g \overline{q}^\pi(\Gamma) + \widehat{\eta}^i I_X Q_\Delta^\varphi(\Gamma). \tag{30}$$

Moreover, $I_X Q_\Delta^\varphi I_X = Q_\Delta^\varphi I_X$ and therefore, by iterating Eq. (30) we have

$$\int_{\mathbf{X}} \sum_{k=0}^n (Q_\Delta^\varphi)^k(\Gamma|x)\left[\delta_{x_0}(dx) + \frac{1}{\alpha}\widehat{\eta}^g \overline{q}^\pi(dx)\right] \leq \widehat{\eta}^i(\Gamma) \tag{31}$$

for any $n \in \mathbb{N}$. Observe that $\sum_{k=0}^\infty (Q_\Delta^\varphi)^k(\Gamma|x) = \widetilde{P}^\varphi(\Gamma \times \mathbf{A}_\Delta^i|x)$. For notational convenience, let us denote by $\rho$ the measure $\left[\delta_{x_0} + \frac{1}{\alpha}\widehat{\eta}^g \overline{q}^\pi\right]\widetilde{P}^\varphi \in \mathcal{M}(\mathbf{X}_\Delta \times \mathbf{A}_\Delta^i)$ which is concentrated on $\mathbb{K}_\Delta^i$. Applying the monotone convergence Theorem and taking the limit as $n$ tends to infinity in Eq. (31), it follows that

$$\rho(\Gamma \times \mathbf{A}_\Delta^i) \leq \widehat{\eta}^i(\Gamma). \tag{32}$$

By using the fact that $\widetilde{P}^\varphi(dz, db|x) = \varphi(db|z)\widetilde{P}^\varphi(dz \times \mathbf{A}_\Delta^i|x)$ for any $x \in \mathbf{X}$, it follows that

$$\rho(dz, db) = \varphi(db|z)\rho(dz, \mathbf{A}_\Delta^i). \tag{33}$$

Combining Eqs. (25) and (32)–(33), it implies

$$\rho \leq \eta^i. \tag{34}$$

Clearly, the measure $\rho \in \mathcal{M}_f(\mathbf{X} \times \mathbf{A}_\Delta^i)$ and satisfies

$$\rho(\Gamma \times \mathbf{A}_\Delta^i) = \delta_{x_0}(\Gamma) + \frac{1}{\alpha}\widehat{\eta}^g\overline{q}^\pi(\Gamma) + \int_\mathbf{X} Q_\Delta^\varphi(\Gamma|z)\rho(dz, \mathbf{A}_\Delta^i)$$

and so, recalling (33)

$$\rho(\Gamma \times \mathbf{A}_\Delta^i) = \delta_{x_0}(\Gamma) + \frac{1}{\alpha}\widehat{\eta}^g\overline{q}^\pi(\Gamma) + \int_{\mathbf{X} \times \mathbf{A}^i} Q_\Delta(\Gamma|z, b)\rho(dz, db).$$

Since $Q_\Delta(\mathbf{X}|z, b) = I_{\mathbb{K}^i}(z, b)$ and $\rho(\mathbb{K}^i) = \rho(\mathbf{X} \times \mathbf{A}^i)$, we obtain from the previous equation

$$\rho(\mathbf{X} \times \{\Delta\}) + \rho(\mathbf{X} \times \mathbf{A}^i) = \delta_{x_0}(\mathbf{X}) + \frac{1}{\alpha}\widehat{\eta}^g\overline{q}^\pi(\mathbf{X}) + \rho(\mathbf{X} \times \mathbf{A}^i)$$

showing

$$\rho(\mathbf{X} \times \{\Delta\}) = \delta_{x_0}(\mathbf{X}) + \frac{1}{\alpha}\widehat{\eta}^g\overline{q}^\pi(\mathbf{X}).$$

Moreover, from the definition of $\rho$ we obtain that

$$\int_\mathbf{X} \big[\widetilde{P}^\varphi(\mathbf{X} \times \{\Delta\}|x) - 1\big]\delta_{x_0}(dx) + \int_\mathbf{X} \big[\widetilde{P}^\varphi(\mathbf{X} \times \{\Delta\}|x) - 1\big]\widehat{\eta}^g\overline{q}^\pi(dx) = 0.$$

Lemma B.1 yields that $\widetilde{P}^\varphi(\mathbf{X} \times \{\Delta\}|x) \leq 1$ for any $x \in \mathbf{X}$. Consequently,

$$\big[\delta_{x_0} + \widehat{\eta}^g\overline{q}^\pi\big](\mathcal{S}_\varphi^c) = 0, \tag{35}$$

where the set $\mathcal{S}_\varphi$ has been introduced in (17). Now, combining Eqs. (29) and (34), it follows

$$\widehat{\eta}^g(\Gamma) \geq \widetilde{P}^\varphi(\Gamma \times \{\Delta\}|x_0) + \frac{1}{\alpha}\widehat{\eta}^g p_\varphi^\pi(\Gamma),$$

where

$$p_\varphi^\pi(\Gamma|x) = \overline{q}^\pi \widetilde{P}^\varphi(\Gamma \times \{\Delta\}|x) - I_\Gamma(x)\overline{q}^\pi(\mathbf{X}|x). \tag{36}$$

Now, from Eq. (35) we have $\widetilde{P}^\varphi(\mathbf{X} \times \{\Delta\}|x_0) = 1$ and

$$\widehat{\eta}^g p_\varphi^\pi(\mathbf{X}) = \int_\mathbf{X} \widetilde{P}^\varphi(\mathbf{X} \times \{\Delta\}|x)\widehat{\eta}^g\overline{q}^\pi(dx) - \widehat{\eta}^g\overline{q}^\pi(\mathbf{X}) = 0.$$

Therefore, since $\widehat{\eta}^g(\mathbf{X}) = 1$ it follows that the positive measure $\gamma$ defined on $\mathbf{X}$ by

$$\gamma = \widehat{\eta}^g - \widetilde{P}^\varphi(\cdot, \{\Delta\}|x_0) - \frac{1}{\alpha}\widehat{\eta}^g p_\varphi^\pi,$$

satisfies $\gamma(\mathbf{X}) = 0$ and so

$$\widehat{\eta}^g(\Gamma) = \widetilde{P}^\varphi(\Gamma \times \{\Delta\}|x_0) + \frac{1}{\alpha}\widehat{\eta}^g p_\varphi^\pi(\Gamma). \tag{37}$$

However, from the definition of $R^\varphi$ in Eq. (18) we have

$$\widetilde{R}^\varphi(dz, db|x) = \widetilde{P}^\varphi(dz, db|x)I_{\mathcal{S}_\varphi}(x) + \delta_{(x,\Delta)}(dz, db)I_{\mathcal{S}_\varphi^c}(x),$$

on $\mathbf{X} \times \mathbf{A}_\Delta^i$. Now, the previous equation and (35) yield

$$\left[\delta_{x_0} + \frac{1}{\alpha}\widehat{\eta}^g\overline{q}^\pi\right]\widetilde{R}^\varphi = \left[\delta_{x_0} + \frac{1}{\alpha}\widehat{\eta}^g\overline{q}^\pi\right]\widetilde{P}^\varphi. \tag{38}$$

Combining Eqs. (36)–(38), we obtain that

$$\widehat{\eta}^g(\Gamma) = \widetilde{R}^\varphi(\Gamma \times \{\Delta\}|x_0) + \frac{1}{\alpha}\widehat{\eta}^g r_\varphi^\pi(\Gamma),$$

where $r_\varphi^\pi$ has been defined in Eq. (21). Clearly, $r_\varphi^\pi$ is a transition rate on $\mathbf{X}$ given $\mathbf{X}$ and so applying the uniqueness result of item $d$) in Theorem 3.2 in [24] we have from (20) that $\widehat{\eta}_{u^{\pi,\varphi}}^g = \widehat{\eta}^g$. From the definitions of $u^{\pi,\varphi}$ and $\eta_u^g$ it is easy to show that $\eta_{u^{\pi,\varphi}}^g(dx, da) = \widehat{\eta}_{u^{\pi,\varphi}}^g(dx)\pi(da|x)$. Therefore,

$$\eta_{u^{\pi,\varphi}}^g(dx, da) = \widehat{\eta}_{u^{\pi,\varphi}}^g(dx)\pi(da|x) = \widehat{\eta}^g(dx)\pi(da|x) = \eta^g(dx, da), \tag{39}$$

showing the first part of the result.

Now, combining the previous equation, (34) and (38) we get that

$$\left[\delta_{x_0} + \frac{1}{\alpha}\widehat{\eta}^g\overline{q}^\pi\right]\widetilde{R}^\varphi = \left[\delta_{x_0} + \frac{1}{\alpha}\widehat{\eta}^g\overline{q}^\pi\right]\widetilde{P}^\varphi = \rho \leq \eta^i,$$

and so, by using Eqs. (19) and (39), we have that $\eta_{u^{\pi,\varphi}}^i = \left[\delta_{x_0} + \frac{1}{\alpha}\widehat{\eta}_{u^{\pi,\varphi}}^g\overline{q}^\pi\right]\widetilde{P}^\varphi \leq \eta^i$ giving the last assertions. $\qquad\qquad\square$

Finally, the following corollary shows that, although $\eta^i$ and $\eta^i_{u^{\pi,\varphi}}$ are not equal, there exists a subset of $\mathbf{X} \times \mathbf{A}^i$ on which they coincide.

**Corollary 4.10** *Under the conditions of Theorem 4.9, there exists a set $D \in \mathcal{B}(\mathbf{X})$ such that for any $z \in D$, $Q^\varphi_\Delta(D|z) = 1$; $\widehat{\eta}^g_{u^{\pi,\varphi}}(D) = \widehat{\eta}^i_{u^{\pi,\varphi}}(D) = 0$, and $\eta^i(\Gamma) = \eta^i_{u^{\pi,\varphi}}(\Gamma)$ for any $\Gamma \in \mathcal{B}\big((\mathbf{X} \setminus D) \times \mathbf{A}^i_\Delta\big)$.*

*Proof* From Proposition 4.4 the pair of measures $(\eta^g_{u^{\pi,\varphi}}, \eta^i_{u^{\pi,\varphi}})$ satisfies

$$\widehat{\eta}^i_{u^{\pi,\varphi}}(\Gamma) = \delta_{x_0}(\Gamma) + \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} \overline{q}(\Gamma|x,a)\eta^g_{u^{\pi,\varphi}}(dx,da) + \int_{\mathbf{X}} Q^\varphi_\Delta(\Gamma|z)\widehat{\eta}^i_{u^{\pi,\varphi}}(dz),$$
(40)

since for $u^{\pi,\varphi}$ we have $\psi_n(\cdot|h_n) = \delta_{+\infty}(\cdot)$ for any $h_n = (y_0, \theta_1, \dots \theta_n, y_n) \in \mathbf{H}_n$. Moreover, according to Eq. (29), we have

$$\widehat{\eta}^i(\Gamma) = \delta_{x_0}(\Gamma) + \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} \overline{q}(\Gamma|x,a)\eta^g(dx,da) + \int_{\mathbf{X}} Q^\varphi_\Delta(\Gamma|z)\widehat{\eta}^i(dz)$$

for any $\Gamma \in \mathcal{B}(\mathbf{X})$. Therefore, the measure $\gamma$ defined on $\mathbf{X}$ by $\gamma = \widehat{\eta}^i - \widehat{\eta}^i_{u^{\pi,\varphi}}$ satisfies the following equation

$$\gamma(\Gamma) = \int_{\mathbf{X}} Q^\varphi_\Delta(\Gamma|z)\gamma(dz) \text{ for any } \Gamma \in \mathcal{B}(\mathbf{X})$$
(41)

since $\eta^g_{u^{\pi,\varphi}} = \eta^g$. Define the sequence of sets $(\mathbf{X}_n)_{n \in \mathbb{N}}$ by $\mathbf{X}_{n+1} = \{z \in \mathbf{X}_n : Q^\varphi_\Delta(\mathbf{X}_n|z) = 1\}$, for $n \in \mathbb{N}^*$ and $\mathbf{X}_0 = \mathbf{X}$. This sequence satisfies $\gamma(\mathbf{X}_n \setminus \mathbf{X}_{n+1}) = 0$, $Q^\varphi_\Delta(\mathbf{X}_n|z) = 1$, for any $z \in \mathbf{X}_{n+1}$ and

$$\gamma(\mathbf{X}_n) = \int_{\mathbf{X}_n} Q^\varphi_\Delta(\mathbf{X}_n|z)\gamma(dz),$$
(42)

for any $n \in \mathbb{N}$. Indeed, Eq. (42) clearly holds for $n = 0$ by using (41). Therefore, we have $\gamma(z \in \mathbf{X}_0 : Q^\varphi_\Delta(\mathbf{X}_0|z) < 1) = 0$ implying that $\gamma(\mathbf{X}_0 \setminus \mathbf{X}_1) = 0$. Moreover, by definition of $\mathbf{X}_1$ it is straightforward to see that $Q^\varphi_\Delta(\mathbf{X}_0|z) = 1$ for any $z \in \mathbf{X}_1$. Suppose the decreasing family of sets $(\mathbf{X}_j)_{j \in \mathbb{N}_n}$ satisfies the above equations. From Eq. (42) we have $\gamma(z \in \mathbf{X}_n : Q^\varphi_\Delta(\mathbf{X}_n|z) < 1) = 0$ showing that $\gamma(\mathbf{X}_n \setminus \mathbf{X}_{n+1}) = 0$. Then by using Eq. (41), we obtain $\gamma(\mathbf{X}_{n+1}) = \int_{\mathbf{X}_{n+1}} Q^\varphi_\Delta(\mathbf{X}_{n+1}|z)\gamma(dz)$. Finally, by definition of $\mathbf{X}_{n+1}$, we have $Q^\varphi_\Delta(\mathbf{X}_n|z) = 1$ for any $z \in \mathbf{X}_{n+1}$. Let us introduce the set $D \subset \mathbf{X}$ defined by $D = \bigcap_{j=0}^{\infty} \mathbf{X}_j$. Then $\gamma(\mathbf{X} \setminus D) = \sum_{j=0}^{\infty} \gamma(\mathbf{X}_j \setminus \mathbf{X}_{j+1}) = 0$. Consequently, for any $\Gamma \in \mathcal{B}\big((\mathbf{X} \setminus D)\big)$ we have $\widehat{\eta}^i(\Gamma) = \widehat{\eta}^i_{u^{\pi,\varphi}}(\Gamma)$, so the measures $\eta^i$ and $\eta^i_{u^{\pi,\varphi}}$ coincide on $(\mathbf{X} \setminus D) \times \mathbf{A}^i_\Delta$ because $\eta^i(dx, da) = \widehat{\eta}^i(dx)\varphi(da|x)$ and $\eta^i_{u^{\pi,\varphi}}(dx, da) = \widehat{\eta}^i_{u^{\pi,\varphi}}(dx)\varphi(da|x)$ due to (27).

Now, observe that for any $z \in D$ and $j \in \mathbb{N}$, $Q_\Delta^\varphi(\mathbf{X}_j|z) = 1$ implying $Q_\Delta^\varphi(D|z) = 1$ and so, choosing $\Gamma = D$ in Eq. (40), we have

$$\widehat{\eta}_{u^{\pi,\varphi}}^i(D) = \delta_{x_0}(D) + \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} \overline{q}(D|x,a) \eta_{u^{\pi,\varphi}}^g(dx,da) + \widehat{\eta}_{u^{\pi,\varphi}}^i(D)$$
$$+ \int_{\mathbf{X} \setminus D} Q_\Delta^\varphi(D|z) \widehat{\eta}_{u^{\pi,\varphi}}^i(dz)$$

leading to $\delta_{x_0}(\Gamma) + \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} \overline{q}(\Gamma|x,a) \eta_{u^{\pi,\varphi}}^g(dx,da) + \int_{\mathbf{X} \setminus D} Q_\Delta^\varphi(\Gamma|z) \widehat{\eta}_{u^{\pi,\varphi}}^i(dz) =$ 0, for any $\Gamma \in \mathcal{B}(D)$. Consequently, $\delta_{x_0}(D) + \frac{1}{\alpha} \eta_{u^{\pi,\varphi}}^g \overline{q}(D) = 0$ and $\left[\delta_{x_0} + \frac{1}{\alpha} \eta_{u^{\pi,\varphi}}^g \overline{q}\right] \sum_{k \in \mathbb{N}} (Q_\Delta^\varphi)^k I_{\mathbf{X} \setminus D} Q_\Delta^\varphi(D) = 0$, where we have used (27) to get the last equation. Combining the two previous equations, it can be shown easily by induction that $\left[\delta_{x_0} + \frac{1}{\alpha} \eta_{u^{\pi,\varphi}}^g \overline{q}\right] (Q_\Delta^\varphi)^k(D) = 0$ for any $k \in \mathbb{N}$ and so, recalling (27), it follows that $\widehat{\eta}_{u^{\pi,\varphi}}^i(D) = 0$.

Now, from Eq. (40) and by using the fact that $\delta_{x_0}(D) = \widehat{\eta}_{u^{\pi,\varphi}}^i(D) = 0$, we obtain

$$\int_{\mathbf{X} \times \mathbf{A}^i} Q_\Delta(D|z,b) \eta_{u^{\pi,\varphi}}^i(dz,db) + \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} \overline{q}(D|x,a) \eta_{u^{\pi,\varphi}}^g(dx,da) = 0.$$

Therefore, Eq. (16) yields $\widehat{\eta}_{u^{\pi,\varphi}}^g(D \times \mathbf{A}^g) = \frac{1}{\alpha} \int_{D \times \mathbf{A}^g} q(\{x\}|x,a) \eta_{u^{\pi,\varphi}}^g(dx,da)$. Since $\eta_{u^{\pi,\varphi}}^g \geq 0$ and $q(\{x\}|x,a) \leq 0$, we conclude that $\widehat{\eta}_{u^{\pi,\varphi}}^g(D \times \mathbf{A}^g) = 0$. □

## 5 The LP Formulation

The main objective of this section is to show the existence of an optimal strategy for the constrained optimal control problem introduced in Definition 5.1. The idea to get this existence result can be decomposed into two steps. First, we introduce the (primal) linear program $\mathbb{PLP}$ associated with the optimization problem under consideration (see Definition 5.1) and show in Theorem 5.2 that there exists an optimal strategy for the constrained optimization control problem if and only if the linear program $\mathbb{PLP}$ is solvable. The second step consists in showing that the $\mathbb{PLP}$ is solvable (Theorem 5.5) and this is done by introducing an auxiliary linear program (whose properties are studied in Proposition 5.4) and by considering an additional set of hypotheses (see Assumption C ). Combining these two steps, it is straightforward to obtain the existence of an optimal randomized control strategy for the constrained optimal control problem (see Theorem 5.6). An easy consequence of this result is that the class of strategies introduced in Definition 4.7 is a sufficient set. Assumption C is a standard hypothesis in the literature on CTMDPs (see for example [24]) and mainly requires that the parameters of the system be lower semicontinuous and that the transition rate be weakly continuous. As an independent result, the dual linear program associated with the linear program PLP is briefly discussed at the end of this section.

**Definition 5.1** The constrained linear program, labeled $\mathbb{PLP}$, is defined as minimize $\eta(C_0)$ subject to $\eta \in \mathbb{L}$ where $\mathbb{L}$ is defined by the set of measures $\eta$ in $\mathcal{M}_f(\mathbf{X} \times \mathbf{A})$ which are admissible in the sense of Definition 4.5 and such that for any $j \in \mathbb{N}_p^*$, $\eta(C_j) \leq B_j$.

The nonnegative real number $\inf_{\eta \in \mathbb{L}} \eta(C_0)$ is called the value of the constrained linear program $\mathbb{PLP}$. Below, we say that $\mathbb{PLP}$ is solvable if there is $\eta^* \in \mathbb{L}$ such that $\eta^*(C_0) = \inf_{\eta \in \mathbb{L}} \eta(C_0)$.

**Theorem 5.2** *The values of the constrained control problem and the linear program* $\mathbb{PLP}$ *are equivalent:*

$$\inf_{\eta \in \mathbb{L}} \eta(C_0) = \inf_{u \in \mathcal{U}^f} \mathcal{V}_0(u, x_0).$$

*Moreover, assume the existence of* $\overline{u} \in \mathcal{U}^f$ *such that* $\mathcal{V}_0(\overline{u}, x_0) < \infty$. *Then the following assertions hold:*

  (i) *The measure* $\eta_{\overline{u}}$ *as defined in Eq.* (11) *for the strategy* $\overline{u}$ *belongs to* $\mathbb{L}$ *and* $\eta_{\overline{u}}(C_0) < \infty$.
 (ii) *The constrained optimal control problem as introduced in Definition* 3.1 *is solvable if and only if the linear program* $\mathbb{PLP}$ *is solvable.*
(iii) *If the constrained optimal control problem is solvable then there exists a randomized stationary optimal control strategy where the interventions only occur after the natural jumps and with a possible intervention at the initial moment.*

*Proof* if $\eta \in \mathbb{L}$ then it is admissible in the sense of Definition 4.5. From Theorem 4.9, the control strategy $u^{\pi,\varphi} \in \mathcal{U}$ where $\pi$ (respectively, $\varphi$) has been defined in Eq. (24) (respectively, (25)) satisfies $\eta_{u^{\pi,\varphi}}^g = \eta^g$ and $\eta_{u^{\pi,\varphi}}^i \leq \eta^i$ with $\eta^g$ (respectively, $\eta^i$) given in (22) (respectively, (23)). Therefore, we have for any $j \in \mathbb{N}_p^*$, $\mathcal{V}_j(u^{\pi,\varphi}, x_0) \leq \eta(C_j) \leq B_j$ and $\mathcal{V}_0(u^{\pi,\varphi}, x_0) \leq \eta(C_0)$. In particular, this first statement implies that on one hand $\mathbb{L}$ is empty if $\mathcal{U}^f$ is empty and on the other hand if the set $\mathcal{U}^f$ is not empty with $\mathcal{V}_0(u, x_0) = \infty$ for any $u \in \mathcal{U}^f$ then either $\mathbb{L}$ is empty or $\inf_{\eta \in \mathbb{L}} \eta(C_0) = \infty$ showing in any case $\inf_{\eta \in \mathbb{L}} \eta(C_0) = \inf_{u \in \mathcal{U}^f} \mathcal{V}_0(u, x_0) = \infty$.

Now, if $\mathcal{V}_0(u, x_0) < \infty$ for $u \in \mathcal{U}$ and $\mathcal{V}_j(u, x_0) \leq B_j$ for any $j \in \mathbb{N}_p^*$ then recalling Assumptions (A2), (A3) and B we have necessarily $\eta_u^i(\mathbf{X} \times \mathbf{A}^i) < \infty$. From Lemma A.1, it gives $\eta_u^i(\mathbf{X} \times \mathbf{A}_\Delta^i) < \infty$. Consequently, according to Theorem 4.6, for any admissible control strategy $u \in \mathcal{U}$ such that $\mathcal{V}_0(u, x_0) < \infty$ and $\mathcal{V}_j(u, x_0) \leq B_j$ for any $j \in \mathbb{N}_p^*$, there exists a finite measure $\eta_u \in \mathcal{M}_f(\mathbf{X} \times \mathbf{A})$ concentrated on $\mathbb{K}^g \cup \mathbb{K}^i$ satisfying Eq. (16) with $\eta_u(C_j) = \mathcal{V}_j(u, x_0)$ for any $j \in \mathbb{N}_p$ implying that $\eta_u \in \mathbb{L}$ and $\eta_u(C_0) < \infty$.

Combining these two statements, we obtain easily the results. $\square$

To study the solvability of the linear program $\mathbb{PLP}$, we need to introduce an auxiliary linear program. First, let us define $\mathbf{X}_\sigma$ by $\mathbf{X} \cup \{\sigma\}$ where $\sigma$ is an isolated point and the kernel $\widetilde{Q}$ on $\mathbf{X}_\sigma$ given $\mathbb{K}^g \cup \mathbb{K}^i \cup (\{\sigma\} \times \mathbf{A})$ by

$$\widetilde{Q}(\Gamma|x,a) = \left\{ \frac{1}{K+\alpha}\left[\overline{q}(\Gamma \cap \mathbf{X}|x,a) + \delta_x(\Gamma \cap \mathbf{X})(K - \overline{q}(\mathbf{X}|x,a))\right] \right.$$
$$\left. + \frac{\alpha}{K+\alpha}\delta_\sigma(\Gamma)\right\}I_{\mathbb{K}^g}(x,a) + Q(\Gamma \cap \mathbf{X}|x,a)I_{\mathbb{K}^i}(x,a) + \delta_\sigma(\Gamma)I_{\{\sigma\}}(x).$$

$$(43)$$

**Definition 5.3** The auxiliary linear program, labeled $\mathbb{LP}'$, is defined as minimize $\rho(C_0')$ subject to $\rho \in \mathbb{L}'$ where $\mathbb{L}'$ is defined by the set of measures $\rho$ in $\mathcal{M}(\mathbf{X}_\sigma \times \mathbf{A})$ concentrated on $\mathbb{K}^g \cup \mathbb{K}^i \cup (\{\sigma\} \times \mathbf{A})$ such that for any $\Gamma \in \mathcal{B}(\mathbf{X}_\sigma)$

$$\rho(\Gamma \times \mathbf{A}) = \delta_{x_0}(\Gamma) + \int_{\mathbf{X}_\sigma \times \mathbf{A}} \widetilde{Q}(\Gamma|z,b)\rho(dz,db), \tag{44}$$

$$\rho(C_j') \leq B_j, \text{ for any } j \in \mathbb{N}_p^*,$$

$$\rho(\mathbb{K}^g) \leq \frac{K+\alpha}{\alpha},$$

with $C_j'(x,a) = \frac{1}{K+\alpha}C_j^g(x,a)I_{\mathbb{K}^g}(x,a) + c_j^i(x,a)I_{\mathbb{K}^i}(x,a)$ for $j \in \mathbb{N}_p$.

**Proposition 5.4** *The following assertions hold:*

*i) If the measure $\eta$ belongs to $\mathbb{L}$ then the measure $\rho$, defined on $\mathbf{X}_\sigma \times \mathbf{A}$ by*

$$\rho(\Gamma) = \eta(\Gamma \cap \mathbb{K}^i) + \frac{K+\alpha}{\alpha}\eta(\Gamma \cap \mathbb{K}^g),$$

*for any $\Gamma \in \mathcal{B}(\mathbf{X} \times \mathbf{A})$ and $\rho(\{\sigma\} \times \Gamma) = +\infty$ for any $\Gamma \in \mathcal{B}(\mathbf{A})$, belongs to $\mathbb{L}'$. Moreover, $\rho(C_j') = \eta(C_j)$ for any $j \in \mathbb{N}_p$.*

*ii) If the measure $\rho \in \mathbb{L}'$ satisfies $\rho(C_0') < \infty$ then the measure $\eta$ defined on $\mathbf{X} \times \mathbf{A}$ by*

$$\eta(\Gamma) = \rho(\Gamma \cap \mathbb{K}^i) + \frac{\alpha}{K+\alpha}\rho(\Gamma \cap \mathbb{K}^g)$$

*belongs to $\mathbb{L}$. Moreover, $\eta(C_j) = \rho(C_j')$ for any $j \in \mathbb{N}_p$.*

*Proof* Regarding item *i*), it is clear that $\rho$ so defined is a positive measure on $\mathbf{X}_\sigma \times \mathbf{A}$ concentrated on $\mathbb{K}^g \cup \mathbb{K}^i \cup (\{\sigma\} \times \mathbf{A})$. Moreover, a straightforward calculus show that $\rho$ satisfies Eq. (44) with $\rho(C_j') = \eta(C_j)$, for any $j \in \mathbb{N}_p$, giving the first part of the result. For item *ii*), we have $\eta(\mathbb{K}^g) = \frac{\alpha}{K+\alpha}\rho(\mathbb{K}^g) \leq 1$. Moreover, combining assumptions (A2), (A3) and B it follows that $\underline{c}\eta(\mathbb{K}^i) \leq \sum_{j \in \mathbb{N}_p^*} B_j + \eta(C_0')$. Consequently, $\eta \in \mathcal{M}_f(\mathbf{X} \times \mathbf{A})$ and $\eta$ is clearly concentrated on $\mathbb{K}^g \cup \mathbb{K}^i$. Now, simple algebraic manipulations yield that the measure $\eta$ satisfies Eq. (16) with $\eta(C_j) = \rho(C_j')$, for any $j \in \mathbb{N}_p$ showing the last part of the result. □

**Assumption C** (C1) The transition rate $q$ is weakly continuous, that is, for any $h \in \mathbb{C}_b(\mathbf{X})$, $qh \in \mathbb{C}_b(\mathbb{K}^g)$.
(C2) If $\mathbb{K}^i \neq \emptyset$, the transition kernel $Q$ is weakly continuous, that is, $Qh \in \mathbb{C}_b(\mathbb{K}^i)$ for any $h \in \mathbb{C}_b(\mathbf{X})$.
(C3) For any $j \in \mathbb{N}_p$, the function $c_j^i$ is lower semicontinuous on $\mathbb{K}^i$.
(C4) For any $j \in \mathbb{N}_p$, the function $C_j^g$ is lower semicontinuous on $\mathbb{K}^g$.
(C5) For any $x \in \mathbf{X}$, $\mathbf{A}^g(x)$ and $\mathbf{A}^i(x)$ are compact sets.
(C6) The multifunction $\Psi : \mathbf{X} \to \mathbf{A}$ defined by $\Psi(x) = \mathbf{A}^g(x) \cup \mathbf{A}^i(x)$ is upper semicontinuous.

**Theorem 5.5** *Assume that there exists $\overline{\eta} \in \mathbb{L}$ such that $\overline{\eta}(C_0) < \infty$. Then the linear program $\mathbb{PLP}$ is solvable.*

*Proof* Introduce $\mathbb{K} = \mathbb{K}^g \cup \mathbb{K}^i \cup (\{\sigma\} \times \mathbf{A})$. According to Assumption C, it is easy to see that the stochastic kernel $\widetilde{Q}$ on $\mathbf{X}_\sigma$ given $\mathbb{K}$ is weakly continuous and the mappings $C_j'$ are nonnegative and lower semicontinuous on $\mathbb{K}$ for any $j \in \mathbb{N}_p$. Consequently, by using Theorem 7.2 in [5], the auxiliary linear program $\mathbb{LP}'$ is solvable because, according to item $i$) of Proposition 5.4, there is $\overline{\rho} \in \mathbb{L}'$ such that $\overline{\rho}(C_0') = \overline{\eta}(C_0) < \infty$. Observe that Theorem 7.2 in [5] is an extension of Theorem 4.1 in [7] to the case where the action space depends on the state variable. Now, from Proposition 5.4, it is easy to see that the auxiliary linear program $\mathbb{LP}'$ is solvable if and only if the linear program $\mathbb{LP}$ is solvable, since there exists $\overline{\rho} \in \mathbb{L}'$ such that $\overline{\rho}(C_0') < \infty$. Let $\rho^*$ be a measure in $\mathbb{L}'$ such that $\inf_{\rho \in \mathbb{L}'} \rho(C_0') = \rho^*(C_0')$. Then, $\rho^*(C_0') \leq \overline{\rho}(C_0') < \infty$. Consequently, item $ii$) of Proposition 5.4 implies that the measure $\eta^*$ defined on $\mathbf{X} \times \mathbf{A}$ by

$$\eta^*(\Gamma) = \rho^*(\Gamma \cap \mathbb{K}^i) + \frac{\alpha}{K + \alpha} \rho^*(\Gamma \cap \mathbb{K}^g)$$

belongs to $\mathbb{L}$ with $\eta^*(C_0) = \rho^*(C_0')$. Finally, it is easy to show that we have necessarily

$$\inf_{\eta \in \mathbb{L}} \eta(C_0) = \eta^*(C_0)$$

by using item $i$) of Proposition 5.4, giving the result.                              $\square$

If $\mathbb{L} \neq \emptyset$ and $\eta(C_0) = \infty$ for any $\eta \in \mathbb{L}$ then $\mathbb{PLP}$ is also solvable without Assumption C. The last following Theorem is the main result of this section. It establishes the existence of an optimal randomized control strategy and states that the class of strategies introduced in 4.7 is a sufficient set.

**Theorem 5.6** *Assume that there exists $\overline{u} \in \mathcal{U}^f$ such that $\mathcal{V}_0(\overline{u}, x_0) < \infty$. Then, there exists a randomized stationary optimal control strategy, for the constrained optimal control problem introduced in Definition 3.1, where the interventions only occur after the natural jumps and with a possible intervention at the initial moment, that is, there exist $\pi^* \in \mathcal{P}^g(\mathbf{A}^g|\mathbf{X})$ and $\varphi^* \in \mathcal{P}^i(\mathbf{A}_\Delta^i|\mathbf{X}_\Delta)$ such that the strategy $u^{\pi^*,\varphi^*}$ as introduced in Definition 4.7 belongs to $\mathcal{U}^f$ and satisfies*

$$\inf_{u \in \mathcal{U}^f} \mathcal{V}_0(u, x_0) = \mathcal{V}_0(u^{\pi^*,\varphi^*}, x_0) = \eta_{u^{\pi^*,\varphi^*}}(C_0) = \inf_{\eta \in \mathbb{L}} \eta(C_0). \qquad (45)$$

*As a consequence, the class of randomized stationary control strategy as introduced in Definiton 4.7 is a sufficient class of control strategy for the constrained control problem under consideration.*

*Proof* This result is a straightforward consequence of Theorems 5.2 and 5.5.    □

We now discuss briefly the dual program associated to the linear program $\mathbb{PLP}$ in the case of an unconstrained control problem, that is $p = 0$. Assume that $|C_0^g(x,a)| \leq K$ for any $(x,a) \in \mathbb{K}^g$. The dual linear program is defined as maximise $W(x_0)$ subject to $W \in \mathbb{L}^*$ where $\mathbb{L}^*$ is the set of functions belonging to $\mathbb{B}(\mathbf{X})$ such that

$$W(x) \leq C_0(x,a) + I_{\mathbb{K}^i}(x,a) \int_{\mathbf{X}} W(y) Q(dy|x,a) + \frac{1}{\alpha} I_{\mathbb{K}^g}(x,a) \int_{\mathbf{X}} W(y) q(dy|x,a).$$

It is easy to show that that for any $\eta \in \mathbb{L}$ and $W \in \mathbb{L}^*$, we have $\eta(C_0) \geq W(x_0)$ showing that $\inf_{\eta \in \mathbb{L}} \eta(C_0) \geq \sup_{W \in \mathbb{L}^*} W(x_0)$. Now, if Assumptions A and B hold then we have $\inf_{u \in \mathcal{U}} \mathcal{V}_0(u, x_0) = \inf_{\eta \in \mathbb{L}} \eta(C_0) \geq \sup_{W \in \mathbb{L}^*} W(x_0)$, from Theorem 5.2. Observe that the set $\mathbb{L}^*$ can be equivalently described by the set of functions belonging to $\mathbb{B}(\mathbf{X})$ and satisfying the following two inequalities

$$\begin{cases} \alpha W(x) \leq \displaystyle\inf_{a \in \mathbf{A}^g(x)} \left\{ C_0^g(x,a) + \int_{\mathbf{X}} W(y) q(dy|x,a) \right\}, \\ W(x) \leq \displaystyle\inf_{a \in \mathbf{A}^i(x)} \left\{ c_0^i(x,a) + \int_{\mathbf{X}} W(y) Q(dy|x,a) \right\}. \end{cases} \qquad (46)$$

Now, assume that the sets $\mathbf{A}^g$ and $\mathbf{A}^i$ are compact and the sets $\mathbb{K}^g$ and $\mathbb{K}^i$ are closed in $\mathbf{X} \times \mathbf{A}^g$ and $\mathbf{X} \times \mathbf{A}^i$ correspondingly and suppose that Assumptions (C1)–(C4) hold. In this context, according to item *c*) of Corollary 4.8 in [8], $\inf_{u \in \mathcal{U}} \mathcal{V}_0(u, x_0) = V(x_0)$ where $V$ is the unique bounded measurable solution to the Bellman equation

$$\inf_{a \in \mathbf{A}^g(x)} \left\{ -\alpha V(x) + C_0^g(x,a) + \int_{\mathbf{X}} V(y) q(dy|x,a) \right\}$$
$$\wedge \inf_{a \in \mathbf{A}^i(x)} \left\{ -V(x) + c_0^i(x,a) + \int_{\mathbf{X}} V(y) Q(dy|x,a) \right\} = 0.$$

Since $V$ satisfies the inequalities in (46), then

$$\inf_{u \in \mathcal{U}} \mathcal{V}_0(u, x_0) = \inf_{\eta \in \mathbb{L}} \eta(C_0) = \sup_{W \in \mathbb{L}^*} W(x_0) = V(x_0),$$

and there is no duality gap. Consequently, if there are no constraints then solving the dual linear program is equivalent to solving the Bellman equation.

## Appendix 1: Proofs of Propositions 4.3 and 4.4

The next result provides a sufficient condition in terms of the finiteness of the occupation measure to ensure that the process is not explosive.

**Lemma A.1** *For any* $u \in \mathcal{U}$, $\eta_u^i(\mathbf{X} \times \{\Delta\}) \leq 1 + \dfrac{1}{\alpha} \displaystyle\int_{\mathbb{K}^i} \overline{q}(\mathbf{X}|x, a)\eta_u^g(dx, da) + \eta_u^i(\mathbf{X} \times \mathbf{A}^i)$. *If* $\eta_u^i(\mathbf{X} \times \mathbf{A}_\Delta^i) < \infty$ *then* $\mu_u^i(\mathbf{Y}) < \infty$ *and* $\mathbb{P}_{x_0}^u(T_\infty < \infty) = 0$.

*Proof* Note that

$$
\eta_u^i(\mathbf{X} \times \{\Delta\}) = \widetilde{\mu}_u^i(\mathbf{X} \times \{\Delta\}) = \sum_{j=1}^{\infty} \mu_u^i(\{\mathbf{y} \in \mathbf{Y} : \mathbf{y}_j \in \mathbf{X} \times \{\Delta\}\})
$$

$$
= \sum_{j=1}^{\infty} \mu_u^i(\mathbf{Y}_{j-1}) = \mu_u^i(\mathbf{Y}) = \mathbb{E}_{x_0}^u \left[ \int_{]0,T_\infty[} e^{-\alpha s} \mu(ds, \mathbf{Y}) \right] + u_0(\mathbf{Y}|x_0).
$$

Since $\nu = \nu_0 + \nu_1$ is the predictable projection of $\mu$ and $\nu_1(ds, \cdot)$ is concentrated on $\mathbf{Y}^*$, we see that

$$
\eta_u^i(\mathbf{X} \times \{\Delta\}) = \mathbb{E}_{x_0}^u \left[ \int_{]0,T_\infty[} e^{-\alpha s} \int_{\mathbf{A}^g} \overline{q}(\mathbf{X}|\overline{x}(\xi_{s-}), a)\pi(da|s)ds \right]
$$

$$
+ \mathbb{E}_{x_0}^u \left[ \int_{]0,T_\infty[} e^{-\alpha s} \nu_1(ds, \mathbf{Y}) \right] + 1
$$

$$
= \frac{1}{\alpha} \int_{\mathbb{K}^i} \overline{q}(\mathbf{X}|x, a)\eta_u^g(dx, da) + \mathbb{E}_{x_0}^u \left[ \int_{]0,T_\infty[} e^{-\alpha s} \nu_1(ds, \mathbf{Y}^*) \right] + 1
$$

and

$$
\mathbb{E}_{x_0}^u \left[ \int_{]0,T_\infty[} e^{-\alpha s} \nu_1(ds, \mathbf{Y}^*) \right] \leq \mu_u^i(\mathbf{Y}^*) = \sum_{k=1}^{\infty} \mu_u^i(\mathbf{Y}_k).
$$

Finally, since $\{\mathbf{y} \in \mathbf{Y} : \mathbf{y}_j \in \mathbf{X} \times \mathbf{A}^i\} = \bigcup_{k=j}^{\infty} \mathbf{Y}_k$,

$$
\eta_u^i(\mathbf{X} \times \mathbf{A}^i) = \sum_{j=1}^{\infty} \mu_u^i(\{\mathbf{y} \in \mathbf{Y} : \mathbf{y}_j \in \mathbf{X} \times \mathbf{A}^i\}) = \sum_{j=1}^{\infty} j\mu_u^i(\mathbf{Y}_j) \geq \sum_{k=1}^{\infty} \mu_u^i(\mathbf{Y}_k),
$$

showing the first part of the result. To prove the last statement, observe first that for any $j \in \mathbb{N}^*$, we have

$$
\{\mathbf{y} \in \mathbf{Y} : \mathbf{y}_j \in \mathbf{X} \times \mathbf{A}_\Delta^i\} = \{\mathbf{y} \in \mathbf{Y} : \mathbf{y}_j \in \mathbf{X} \times \mathbf{A}^i\} \cup \{\mathbf{y} \in \mathbf{Y} : \mathbf{y}_j \in \mathbf{X} \times \{\Delta\}\}
$$

$$
= \bigcup_{k=j}^{\infty} \mathbf{Y}_k \cup \mathbf{Y}_{j-1}
$$

Consequently, $\eta_u^i(\mathbf{X} \times \mathbf{A}_\Delta^i) = \widetilde{\mu}_u^i(\mathbf{X} \times \mathbf{A}_\Delta^i) = \sum_{j \in \mathbb{N}}(j+1)\mu_u^i(\mathbf{Y}_j) \geq \mu_u^i(\mathbf{Y})$. Now, we have that $\mathbb{E}_{x_0}^u\left[\sum_{n=2}^\infty e^{-\alpha T_n} I_{\{T_n < T_\infty\}}\right] \leq \mu_u^i(\mathbf{Y}) < \infty$, showing the last part of the result. $\qquad\square$

*Proof of Proposition* 4.3 Consider $\Gamma \in \mathcal{B}(\mathbf{X})$. From Lemma A.1, $\mathbb{P}_{x_0}^u(T_\infty = +\infty) = 1$ and so, by using the product formula for functions of bounded variation

$$e^{-\alpha t} I_\Gamma(\overline{x}(\xi_t)) = I_\Gamma(\overline{x}(y_1)) - \int_0^t \alpha e^{-\alpha s} I_\Gamma(\overline{x}(\xi_s))ds$$
$$+ \int_{]0,t] \times \mathbf{Y}} e^{-\alpha s}\Big[I_\Gamma(\overline{x}(z)) - I_\Gamma(\overline{x}(\xi_{s-}))\Big]\mu(ds, dz).$$

Therefore, combining the bounded convergence Theorem and the fact that $\mu_u^i(\mathbf{Y}) < \infty$ (see Lemma A.1), we have

$$\eta_u^g(\Gamma \times \mathbf{A}^g) = \alpha\mathbb{E}_{x_0}^u\left[\int_0^\infty e^{-\alpha s} I_\Gamma(\overline{x}(\xi_s))ds\right]$$
$$= \mathbb{E}_{x_0}^u\left[\int_{\mathbf{Y}} I_\Gamma(\overline{x}(y))u_0(dy|x_0)\right]$$
$$+ \mathbb{E}_{x_0}^u\left[\int_{]0,\infty[\times \mathbf{Y}} e^{-\alpha s}\Big[I_\Gamma(\overline{x}(z)) - I_\Gamma(\overline{x}(\xi_{s-}))\Big]\mu(ds, dz)\right].$$

Recalling the definition $\mu_u^i$ (see Eq. 9) and the fact that $\nu$ is the predictable projection of $\mu$, we obtain by using Lemma 3.2

$$\eta_u^g(\Gamma \times \mathbf{A}^g) = \int_{\mathbf{Y}} I_\Gamma(\overline{x}(y))\mu_u^i(dy)$$
$$- \mathbb{E}_{x_0}^u\left[\int_0^\infty e^{-\alpha s}\int_{\mathbf{A}^g} I_\Gamma(\overline{x}(\xi_s))\overline{q}(\mathbf{X}|\overline{x}(\xi_s), a)\pi(da|s)ds\right]$$
$$- \mathbb{E}_{x_0}^u\left[\sum_{n \in \mathbb{N}^*}\int_{]T_n, T_{n+1}]} e^{-\alpha s} I_\Gamma(\overline{x}(\xi_{s-}))\frac{\psi_n(ds - T_n|H_n)}{\psi_n([s - T_n, +\infty]|H_n)}\right],$$

and so,

$$\eta_u^g(\Gamma \times \mathbf{A}^g) = \int_{\mathbf{Y}} I_\Gamma(\overline{x}(y))\mu_u^i(dy)$$
$$- \mathbb{E}_{x_0}^u\left[\int_{\mathbf{X} \times \mathbf{A}^g} I_\Gamma(x)\overline{q}(\mathbf{X}|x, a)\int_0^\infty e^{-\alpha s}\delta_{\overline{x}(\xi_s)}(dx)\pi(da|s)ds\right]$$
$$- \mathbb{E}_{x_0}^u\left[\sum_{n \in \mathbb{N}^*} I_\Gamma(\overline{x}(Y_n))\int_{]T_n, T_{n+1}]} e^{-\alpha s}\frac{\psi_n(ds - T_n|H_n)}{\psi_n([s - T_n, +\infty]|H_n)}\right].$$

By using Fubini's Theorem, we have that

$$
\mathbb{E}^u_{x_0}\left[\int_{\mathbf{X}\times\mathbf{A}^g} I_\Gamma(x)\overline{q}(\mathbf{X}|x,a)\int_0^\infty e^{-\alpha s}\delta_{\overline{x}(\xi_s)}(dx)\pi(da|s)ds\right]
$$
$$
= \frac{1}{\alpha}\int_{\mathbf{X}\times\mathbf{A}^g} I_\Gamma(x)\overline{q}(\mathbf{X}|x,a)\eta^g_u(dx,da).
$$

Moreover, observe that

$$
\mathbb{E}^u_{x_0}\left[\int_{]T_n,T_{n+1}]} e^{-\alpha s}\frac{\psi_n(ds-T_n|H_n)}{\psi_n([s-T_n,+\infty]|H_n)}\Big|\mathcal{F}_{T_n}\right] = e^{-\alpha T_n}\int_{]0,\infty[} e^{-\alpha s}\psi_n(ds|H_n).
$$

Combining the last three equations, it follows that

$$
\eta^g_u(\Gamma\times\mathbf{A}^g) = \int_{\mathbf{Y}} I_\Gamma(\overline{x}(y))\mu^i_u(dy) - \frac{1}{\alpha}\int_{\mathbf{X}\times\mathbf{A}^g} I_\Gamma(x)\overline{q}(\mathbf{X}|x,a)\eta^g_u(dx,da)
$$
$$
- \mathbb{E}^u_{x_0}\left[\sum_{n\in\mathbb{N}^*} e^{-\alpha T_n} I_\Gamma(\overline{x}(Y_n))\int_{]0,\infty[} e^{-\alpha s}\psi_n(ds|H_n)\right].
$$

Finally, remark that $I_\Gamma(\overline{x}(y)) = \sum_{j=1}^\infty I_{\Gamma\times\{\Delta\}}(y_j)$ for any $y = (y_1, y_2, \ldots, y_j, \ldots) \in \mathbf{Y}$. Therefore, $\int_{\mathbf{Y}} I_\Gamma(\overline{x}(y))\mu^i_u(dy) = \eta^i_u(\Gamma\times\{\Delta\})$ showing the result.    □

**Lemma A.2** *Consider a strategy* $u = (u_n)_{n\in\mathbb{N}} \in \mathcal{U}$ *fixed with* $u_n = (\psi_n, \pi_n, \gamma^0_n, \gamma^1_n)$ *for* $n \in \mathbb{N}^*$.

*Then, for any* $n \in \mathbb{N}^*$, $\Gamma \in \mathcal{B}(\mathbf{X}_\Delta)$, $t \in \mathbb{R}_+$, $x \in \mathbf{X}$ *and* $h_n \in \mathbf{H}_n$

$$
\widetilde{\gamma}^0_n(\Gamma\times\mathbf{A}^i_\Delta|h_n,t,x) = \delta_x(\Gamma) + \int_{\mathbf{X}_\Delta\times\mathbf{A}^i_\Delta} Q_\Delta(\Gamma|z,a)\widetilde{\gamma}^0_n(dz,da|h_n,t,x), \quad (47)
$$

$$
\widetilde{\gamma}^1_n(\Gamma\times\mathbf{A}^i_\Delta|h_n) = \delta_{\overline{x}(y_n)}(\Gamma) + \int_{\mathbf{X}_\Delta\times\mathbf{A}^i_\Delta} Q_\Delta(\Gamma|z,a)\widetilde{\gamma}^1_n(dz,da|h_n), \quad (48)
$$

*where* $h_n = (y_0, \theta_1, y_1, \ldots, \theta_n, y_n) \in \mathbf{H}_n$. *Similarly, for any* $\Gamma \in \mathcal{B}(\mathbf{X}_\Delta)$ *and* $x \in \mathbf{X}$

$$
\widetilde{u}_0(\Gamma\times\mathbf{A}^i_\Delta|x) = \delta_x(\Gamma) + \int_{\mathbf{X}_\Delta\times\mathbf{A}^i_\Delta} Q_\Delta(\Gamma|z,a)\widetilde{u}_0(dz,da|x). \quad (49)
$$

*Proof* This Lemma is a straightforward consequence of Lemma 9.4.3 in [14].    □

*Proof of Proposition* 4.4 By using the fact that $\nu$ is the predictable projection of $\mu$ and Lemma 3.2, we have

$$\eta_u^i(\Gamma \times \mathbf{A}_\Delta^i) = \widetilde{u}_0(\Gamma \times \mathbf{A}_\Delta^i | x_0)$$

$$+ \mathbb{E}_{x_0}^u \left[ \int_{]0,\infty[} e^{-\alpha s} \int_{\mathbf{A}^g} \int_\mathbf{X} \widetilde{\gamma}^0(\Gamma \times \mathbf{A}_\Delta^i | x, s) \overline{q}(dx | \overline{x}(\xi_{s-}), a) \pi(da|s) ds \right]$$

$$+ \mathbb{E}_{x_0}^u \left[ \sum_{n \in \mathbb{N}^*} \int_{]T_n, T_{n+1}]} e^{-\alpha s} \widetilde{\gamma}_n^1(\Gamma \times \mathbf{A}_\Delta^i | H_n) \frac{\psi_n(ds - T_n | H_n)}{\psi_n([s - T_n, +\infty[|H_n)} \right].$$

Now, by using Lemma A.2, it follows that

$$\eta_u^i(\Gamma \times \mathbf{A}_\Delta^i)$$

$$= \delta_{x_0}(\Gamma) + \int_{\mathbf{X}_\Delta \times \mathbf{A}_\Delta^i} Q_\Delta(\Gamma|z,b) \widetilde{u}_0(dz, db | x_0)$$

$$+ \mathbb{E}_{x_0}^u \left[ \int_{]0,\infty[} e^{-\alpha s} \int_{\mathbf{A}^g} \int_\mathbf{X} \delta_x(\Gamma) \overline{q}(dx | \overline{x}(\xi_{s-}), a) \pi(da|s) ds \right]$$

$$+ \mathbb{E}_{x_0}^u \left[ \int_{]0,\infty[} e^{-\alpha s} \int_{\mathbf{A}^g} \int_\mathbf{X} \int_{\mathbf{X}_\Delta \times \mathbf{A}_\Delta^i} Q_\Delta(\Gamma|z,b) \widetilde{\gamma}^0(dz, db | x, s) \overline{q}(dx | \overline{x}(\xi_{s-}), a) \pi(da|s) ds \right]$$

$$+ \mathbb{E}_{x_0}^u \left[ \sum_{n \in \mathbb{N}^*} \int_{]T_n, T_{n+1}]} e^{-\alpha s} I_\Gamma(\overline{x}(Y_n)) \frac{\psi_n(ds - T_n | H_n)}{\psi_n([s - T_n, +\infty[|H_n)} \right]$$

$$+ \mathbb{E}_{x_0}^u \left[ \sum_{n \in \mathbb{N}^*} \int_{]T_n, T_{n+1}]} e^{-\alpha s} \int_{\mathbf{X}_\Delta \times \mathbf{A}_\Delta^i} Q_\Delta(\Gamma|z,b) \widetilde{\gamma}_n^1(dz, db | H_n) \frac{\psi_n(ds - T_n | H_n)}{\psi_n([s - T_n, +\infty[|H_n)} \right].$$

Consequently

$$\eta_u^i(\Gamma \times \mathbf{A}_\Delta^i) = \delta_{x_0}(\Gamma) + \int_{\mathbf{X}_\Delta \times \mathbf{A}_\Delta^i} Q_\Delta(\Gamma|z,b) \eta_u^i(dz, db)$$

$$+ \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} \overline{q}(\Gamma \cap \mathbf{X} | x, a) \eta_u^g(dx, da)$$

$$+ \mathbb{E}_{x_0}^u \left[ \sum_{n \in \mathbb{N}^*} \int_{]T_n, T_{n+1}]} e^{-\alpha s} I_\Gamma(\overline{x}(Y_n)) \frac{\psi_n(ds - T_n | H_n)}{\psi_n([s - T_n, +\infty[|H_n)} \right]$$

showing the result.                                                                                          □


## Appendix 2: Proof of Proposition 4.8

This appendix is dedicated to the proof of Proposition 4.8. We first need to derive some technical results. In all this section, we consider $\pi \in \mathcal{P}^g(\mathbf{A}^g | \mathbf{X})$ and $\varphi \in \mathcal{P}^i(\mathbf{A}_\Delta^i | \mathbf{X}_\Delta)$ fixed. Let us introduce the stochastic kernel $G_{\pi,\varphi}$ on $\mathbb{R}_+^* \times \mathbf{Y}$ given $\mathbf{Y}$

$$G_{\pi,\varphi}(dt, dy|z) = \overline{q}^\pi R^\varphi(dy | \overline{x}(z)) e^{-t \overline{q}^\pi (\mathbf{X} | \overline{x}(z))} dt, \tag{50}$$

and the stochastic kernel $L_\pi$ on $\mathbf{X}$ given $\mathbf{Y}$

$$L_\pi(dx|y) = \frac{\delta_{\overline{x}(y)}(dx)}{\alpha + \overline{q}^\pi(\mathbf{X}|\overline{x}(y))}. \tag{51}$$

For notational convenience, we denote

$$H_{\pi,\varphi} = L_\pi \overline{q}^\pi R^\varphi. \tag{52}$$

**Lemma B.1** *Let $\gamma \in \mathcal{P}(\mathbf{Y})$. Then $\widetilde{\gamma}$ is supported on $\mathbb{K}^i_\Delta$ and $\widetilde{\gamma}(\mathbf{X} \times \{\Delta\}) = 1$. Consider $x \in \mathbf{X}$ and a randomized stationary policy $\varphi$ for the model $\mathcal{M}^i$ then $\widetilde{P}^\varphi(\mathbf{X} \times \{\Delta\}|x) \leq 1$. Moreover, $\widetilde{P}^\varphi(\mathbf{X} \times \{\Delta\}|x) = 1$ if and only if $P^\varphi(\mathbf{Y}|x) = 1$.*

*Proof* Let $j \in \mathbb{N}^*$. Observe that $\{\mathbf{y} \in \mathbf{Y} : y_j \in \mathbf{X} \times \{\Delta\}\} = \mathbf{Y}_{j-1}$ and the first assertion is clear. Regarding the second claim, we have $P^\varphi\left(\{\mathbf{y} \in (\mathbb{K}^i_\Delta)^\infty : y_j \in \mathbf{X} \times \{\Delta\}\}|x\right) = P^\varphi\left(\mathbf{Y}_{j-1}|x\right)$ for $x \in \mathbf{X}$ since $P^\varphi$ is the strategic measure for the model $\mathcal{M}^i$ generated by $\varphi$, showing the last part of the result.                    □

**Lemma B.2** *For any $\Upsilon \in \mathcal{B}(\mathbf{Y})$ and $n \in \mathbb{N}^*$, we have*

$$\mathbb{E}^{u^{\pi,\varphi}}_{x_0}\left[I_{\{T_n < \infty\}} e^{-\alpha T_n} \delta_{Y_n}(\Upsilon)\right] = R^\varphi H^{n-1}_{\pi,\varphi}(\Upsilon|x_0). \tag{53}$$

*Proof* Let us show the result by induction. Clearly, this equation holds for $n = 1$. Now, assume that Eq. (53) holds for $n$. Consider $\Upsilon \in \mathcal{B}(\mathbf{Y})$. Then,

$$\mathbb{E}^{u^{\pi,\varphi}}_{x_0}\left[I_{\{T_{n+1} < \infty\}} e^{-\alpha T_{n+1}} \delta_{Y_{n+1}}(\Upsilon)\right]$$

$$= \mathbb{E}^{u^{\pi,\varphi}}_{x_0}\left[I_{\{T_n < \infty\}} e^{-\alpha T_n} \mathbb{E}^{u^{\pi,\varphi}}_{x_0}\left[I_{\{\Theta_{n+1} < \infty\}} e^{-\alpha \Theta_{n+1}} \delta_{Y_{n+1}}(\Upsilon)|\mathcal{F}_{T_n}\right]\right]$$

$$= \mathbb{E}^{u^{\pi,\varphi}}_{x_0}\left[I_{\{T_n < \infty\}} e^{-\alpha T_n} \int_{\mathbb{R}^*_+} e^{-\alpha s} G_{\pi,\varphi}(ds, \Upsilon|Y_n)\right]$$

$$= \mathbb{E}^{u^{\pi,\varphi}}_{x_0}\left[I_{\{T_n < \infty\}} e^{-\alpha T_n} \int_{\mathbb{R}^*_+} e^{-\alpha s} \overline{q}^\pi R^\varphi(\Upsilon|\overline{x}(Y_n)) e^{-s\overline{q}^\pi(\mathbf{X}|\overline{x}(Y_n))} ds\right]$$

$$= \mathbb{E}^{u^{\pi,\varphi}}_{x_0}\left[I_{\{T_n < \infty\}} e^{-\alpha T_n} H_{\pi,\varphi}(\Upsilon|Y_n)\right]$$

$$= \int_{\mathbf{Y}} H_{\pi,\varphi}(\Upsilon|z) R^\varphi H^{n-1}_{\pi,\varphi}(dz|x_0),$$

showing the result.                                                           □

**Proposition B.3** *The following three equalities hold:*

$$\eta^i_{u^{\pi,\varphi}}(dx, da) = \widetilde{R}^\varphi(dx, da|x_0) + \sum_{n=1}^\infty R^\varphi H^{n-1}_{\pi,\varphi} \widetilde{H}_{\pi,\varphi}(dx, da|x_0), \tag{54}$$

$$\widehat{\eta}^g_{u^{\pi,\varphi}}(dx) = \alpha \sum_{n=1}^\infty R^\varphi H^{n-1}_{\pi,\varphi} L_\pi(dx|x_0), \tag{55}$$

*and*

$$\eta_{u^{\pi,\varphi}}^i(\Gamma \times \{\Delta\}) = \widehat{\eta}_{u^{\pi,\varphi}}^g(\Gamma) + \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} I_\Gamma(x) \overline{q}^\pi(\mathbf{X}|x) \widehat{\eta}_{u^{\pi,\varphi}}^g(dx), \qquad (56)$$

*Proof* From the definition of $\mu_{u^{\pi,\varphi}}^i$ (see Eq. 9) and Lemma B.2, we have

$$\mu_{u^{\pi,\varphi}}^i(dy) = R^\varphi(dy|x_0) + \sum_{n=2}^\infty \mathbb{E}_{x_0}^{u^{\pi,\varphi}}\left[ I_{\{T_n < \infty\}} e^{-\alpha T_n} \delta_{Y_n}(dy) \right]$$

$$= R^\varphi(dy|x_0) + \sum_{n=1}^\infty R^\varphi H_{\pi,\varphi}^n(dy|x_0).$$

Observe that $\widetilde{H}_{\pi,\varphi} = L_\pi \overline{q}^\pi \widetilde{R}^\varphi$. Since $\eta_{u^{\pi,\varphi}}^i(dx, da) = \widetilde{\mu}_{u^{\pi,\varphi}}^i(dx, da)$, we obtain easily Eq. (54).

Moreover, we have

$$\mathbb{E}_{x_0}^{u^{\pi,\varphi}}\left[ \int_0^{T_\infty} e^{-\alpha s} \delta_{\overline{x}(\xi_{s-})}(dx)ds \right]$$

$$= \sum_{n=1}^\infty \mathbb{E}_{x_0}^{u^{\pi,\varphi}}\left[ I_{\{T_n < \infty\}} \int_{]T_n, T_{n+1}]} e^{-\alpha s} \delta_{\overline{x}(\xi_{s-})}(dx)ds \right]$$

$$= \frac{1}{\alpha} \sum_{n=1}^\infty \mathbb{E}_{x_0}^{u^{\pi,\varphi}}\left[ I_{\{T_n < \infty\}} e^{-\alpha T_n} \delta_{\overline{x}(Y_n)}(dx)\left(1 - e^{-\alpha\Theta_{n+1}}\right) \right]$$

$$= \sum_{n=1}^\infty \mathbb{E}_{x_0}^{u^{\pi,\varphi}}\left[ I_{\{T_n < \infty\}} e^{-\alpha T_n} \frac{\delta_{\overline{x}(Y_n)}(dx)}{\alpha + \overline{q}^\pi(\mathbf{X}|\overline{x}(Y_n))} \right],$$

and so by using the definition of $L_\pi$ and Lemma B.2, we obtain (55).

Now, from Eq. (55) we get

$$\widehat{\eta}_{u^{\pi,\varphi}}^g(\Gamma) + \frac{1}{\alpha} \int_{\mathbf{X} \times \mathbf{A}^g} I_\Gamma(x) \overline{q}^\pi(\mathbf{X}|x) \widehat{\eta}_{u^{\pi,\varphi}}^g(dx)$$

$$= \sum_{n=1}^\infty \int_{\mathbf{Y}} \frac{\alpha I_\Gamma(\overline{x}(y))}{\alpha + \overline{q}^\pi(\mathbf{X}|\overline{x}(y))} R^\varphi H_{\pi,\varphi}^{n-1}(dy|x_0)$$

$$+ \sum_{n=1}^\infty \int_{\mathbf{Y}} \frac{I_\Gamma(\overline{x}(y)) \overline{q}^\pi(\mathbf{X}|\overline{x}(y))}{\alpha + \overline{q}^\pi(\mathbf{X}|\overline{x}(y))} R^\varphi H_{\pi,\varphi}^{n-1}(dy|x_0)$$

$$= \sum_{n=1}^\infty \int_{\mathbf{Y}} I_\Gamma(\overline{x}(y)) R^\varphi H_{\pi,\varphi}^{n-1}(dy|x_0) = \widetilde{R}^\varphi(\Gamma \times \{\Delta\}|x_0)$$

$$+ \sum_{n=1}^\infty R^\varphi H_{\pi,\varphi}^{n-1} \widetilde{H}_{\pi,\varphi}(\Gamma \times \{\Delta\}|x_0).$$

Recalling (54), we have Eq. (56), showing the result.                     □

*Proof of Proposition* 4.8 Observe that $\widetilde{H}_{\pi,\varphi} = L_\pi \overline{q}^\pi \widetilde{R}^\varphi$, and so

$$\eta_{u^\pi,\varphi}^i(dx, da) = \widetilde{R}^\varphi(dx, da|x_0) + \sum_{n=1}^\infty R^\varphi H_{\pi,\varphi}^{n-1} L_\pi \overline{q}^\pi \widetilde{R}^\varphi(dx, da|x_0),$$

and with (55) we get (19). The measure $\widehat{\eta}_{u^\pi,\varphi}^g$ is finite by definition and so, by using Propositions B.3 and Assumption (A1), we have that for any $\Gamma \in \mathcal{B}(\mathbf{X})$

$$\sum_{n=1}^\infty R^\varphi H_{\pi,\varphi}^{n-1} L_\pi(\Gamma|x_0) < \infty \tag{57}$$

$$\sum_{n=1}^\infty R^\varphi H_{\pi,\varphi}^{n-1} \widetilde{H}_{\pi,\varphi}(\Gamma \times \{\Delta\}|x_0) < \infty \tag{58}$$

Now, from Eq. (55) and the definition of $H_{\pi,\varphi}$ (see Eq. 52), we have that for any $\Gamma \in \mathcal{B}(\mathbf{X})$

$$\frac{1}{\alpha} \widehat{\eta}_{u^\pi,\varphi}^g \overline{q}^\pi \widetilde{R}^\varphi(\Gamma \times \{\Delta\}) = \sum_{n \in \mathbb{N}^*} R^\varphi H_{\pi,\varphi}^{n-1} L_\pi \overline{q}^\pi \widetilde{R}^\varphi(\Gamma \times \{\Delta\}|x_0)$$

$$= \sum_{n \in \mathbb{N}} R^\varphi H_{\pi,\varphi}^n \widetilde{H}_{\pi,\varphi}(\Gamma \times \{\Delta\}|x_0). \tag{59}$$

Moreover, observe that $\dfrac{\overline{q}^\pi(\mathbf{X}|\overline{x}(y))}{\alpha + \overline{q}^\pi(\mathbf{X}|\overline{x}(y))} = 1 - \dfrac{\alpha}{\alpha + \overline{q}^\pi(\mathbf{X}|\overline{x}(y))}$ and so, for $n \geq 2$

$$\int_{\mathbf{X}} I_\Gamma(x) \overline{q}^\pi(\mathbf{X}|x) R^\varphi H_{\pi,\varphi}^{n-1} L_\pi(dx|x_0)$$

$$= \int_{\mathbf{Y}} \frac{I_\Gamma(\overline{x}(y)) \overline{q}^\pi(\mathbf{X}|\overline{x}(y))}{\alpha + \overline{q}^\pi(\mathbf{X}|\overline{x}(y))} R^\varphi H_{\pi,\varphi}^{n-1}(dy|x_0)$$

$$= R^\varphi H_{\pi,\varphi}^{n-2} \widetilde{H}_{\pi,\varphi}(\Gamma \times \{\Delta\}|x_0) - \alpha R^\varphi H_{\pi,\varphi}^{n-1} L_\pi(\Gamma|x_0) \tag{60}$$

and

$$\int_{\mathbf{X}} I_\Gamma(x) \overline{q}^\pi(\mathbf{X}|x) R^\varphi L_\pi(dx|x_0) = \int_{\mathbf{Y}} \frac{I_\Gamma(\overline{x}(y)) \overline{q}^\pi(\mathbf{X}|\overline{x}(y))}{\alpha + \overline{q}^\pi(\mathbf{X}|\overline{x}(y))} R^\varphi(dy|x_0)$$

$$= \widetilde{R}^\varphi(\Gamma \times \{\Delta\}|x_0) - \alpha R^\varphi L_\pi(\Gamma|x_0). \tag{61}$$

Consequently, by using the expression of $\widehat{\eta}_{u^\pi,\varphi}^g$ in (55) and Eqs. (60)−(61)

$$\frac{1}{\alpha} \int_{\mathbf{X}} I_\Gamma(x) \overline{q}^\pi(\mathbf{X}|x) \widehat{\eta}_{u^\pi,\varphi}^g(dx|x_0) = \sum_{n=2}^\infty \int_{\mathbf{X}} I_\Gamma(x) \overline{q}^\pi(\mathbf{X}|x) R^\varphi H_{\pi,\varphi}^{n-1} L_\pi(dx|x_0)$$

$$+ \int_{\mathbf{X}} I_\Gamma(x) \overline{q}^\pi(\mathbf{X}|x) R^\varphi L_\pi(dx|x_0)$$

$$= \sum_{n \in \mathbb{N}} R^\varphi H^n_{\pi,\varphi} \widetilde{H}_{\pi,\varphi}(\Gamma \times \{\Delta\} | x_0) + \widetilde{R}^\varphi(\Gamma \times \{\Delta\} | x_0)$$

$$- \alpha \sum_{n \in \mathbb{N}} R^\varphi H^n_{\pi,\varphi} L_\pi(\Gamma | x_0). \tag{62}$$

Note that the above calculations are possible since the quantities $\sum_{n=1}^\infty R^\varphi H^{n-1}_{\pi,\varphi} L_\pi$ $(\Gamma | x_0)$ and $\sum_{n=1}^\infty R^\varphi H^{n-1}_{\pi,\varphi} \widetilde{H}_{\pi,\varphi}(\Gamma \times \{\Delta\} | x_0)$ are finite (see inequalities (58) and (58)). Recalling (55) and combining Eqs. (59) and (62) we obtain that

$$\frac{1}{\alpha} \widehat{\eta}^g_{u^{\pi,\varphi}} \overline{q}^\pi \widetilde{R}^\varphi(\Gamma \times \{\Delta\}) - \frac{1}{\alpha} \int_{\mathbf{X}} I_\Gamma(x) \overline{q}^\pi(\mathbf{X}|x) \widehat{\eta}^g_{u^{\pi,\varphi}}(dx|x_0)$$

$$= -\widetilde{R}^\varphi(\Gamma \times \{\Delta\} | x_0) + \widehat{\eta}^g_{u^{\pi,\varphi}}(\Gamma | x_0),$$

showing the result.      $\square$

## References

1. Arapostathis, A., Borkar, V.S., Ghosh, M.K.: Ergodic Control of Diffusion Processes. Encyclopedia of Mathematics and Its Applications, vol. 143. Cambridge University Press, Cambridge (2012)
2. Bhatt, Abhay G., Borkar, Vivek S.: Occupation measures for controlled Markov processes: characterization and optimality. Ann. Probab. **24**(3), 1531–1562 (1996)
3. Buckdahn, R., Goreac, D., Quincampoix, M.: Stochastic optimal control and linear programming approach. Appl. Math. Optim. **63**(2), 257–276 (2011)
4. Christensen, Sören: On the solution of general impulse control problems using superharmonic functions. Stoch. Process. Appl. **124**(1), 709–729 (2014)
5. Costa, O.L.V., Dufour, F.: A linear programming formulation for constrained discounted continuous control for piecewise deterministic Markov processes. J. Math. Anal. Appl. **424**(2), 892–914 (2015)
6. Davis, M.H.A.: Markov Models and Optimization. Monographs on Statistics and Applied Probability, vol. 49. Chapman & Hall, London (1993)
7. Dufour, F., Horiguchi, M., Piunovskiy, A.B.: The expected total cost criterion for Markov decision processes under constraints: a convex analytic approach. Adv. Appl. Probab. **44**(3), 774–793 (2012)
8. Dufour, F., Piunovskiy, A.B.: Impulsive control for continuous-time Markov decision processes. Adv. Appl. Probab. **47**(1), 106–127 (2015)
9. Guo, X., Hernández-Lerma, O.: Continuous-Time Markov Decision Processes: Theory and applications. Stochastic Modelling and Applied Probability, vol. 62. Springer, Berlin (2009)
10. Guo, X., Hernández-Lerma, O., Prieto-Rumeau, T.: A survey of recent results on continuous-time Markov decision processes. Top **14**(2), 177–261 (2006). With comments and a rejoinder by the authors
11. Guo, X., Piunovskiy, A.: Discounted continuous-time Markov decision processes with constraints: unbounded transition and loss rates. Math. Oper. Res. **36**(1), 105–132 (2011)
12. Helmes, K., Stockbridge, R.H.: Linear programming approach to the optimal stopping of singular stochastic processes. Stochastics **79**(3–4), 309–335 (2007)
13. Hernández-Lerma, O., Lasserre, J.B.: Discrete-Time Markov Control Processes. Applications of Mathematics, vol. 30. Springer, New York (1996)
14. Hernández-Lerma, O., Lasserre, J.B.: Further Topics on Discrete-Time Markov Control Processes. Applications of Mathematics, vol. 42. Springer, New York (1999)
15. Hordijk, A., van der Duyn Schouten, F.A.: Average optimal policies in Markov decision drift processes with applications to a queueing and a replacement model. Adv. Appl. Probab. **15**(2), 274–303 (1983)
16. Hordijk, A., Schouten, F.A.V.D.D.: Markov decision drift processes: conditions for optimality obtained by discretization. Math. Oper. Res. **10**(1), 160–173 (1985)
17. Jacod, J.: Multivariate point processes: predictable projection, Radon-Nikodým derivatives, representation of martingales. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, **31**, 235–253 (1974/75)

18. Jacod, J.: Calcul Stochastique et Problèmes de Martingales. Lecture Notes in Mathematics, vol. 714. Springer, Berlin (1979)
19. Kitaev, M.Y., Rykov, V.V.: Controlled Queueing Systems. CRC Press, Boca Raton (1995)
20. Kurtz, T.G., Stockbridge, R.H.: Existence of Markov controls and characterization of optimal Markov controls. SIAM J. Control Optim. **36**(2), 609–653 (1998)
21. Kushner, H.J., Dupuis, P.G.: Numerical Methods for Stochastic Control Problems in Continuous Time. Applications of Mathematics (New York), vol. 24. Springer, New York (1992)
22. Kushner, H.J., Martins, L.F.: Numerical methods for stochastic singular control problems. SIAM J. Control Optim. **29**(6), 1443–1475 (1991)
23. Last, G., Brandt, A.: Marked Point Processes on the Real Line. Probability and Its Applications (New York). Springer, New York (1995)
24. Piunovskiy, A., Zhang, Y.: Discounted continuous-time Markov decision processes with unbounded rates: the convex analytic approach. SIAM J. Control Optim. **49**(5), 2032–2061 (2011)
25. Piunovskiy, A.B.: Multicriteria impulsive control of jump Markov processes. Math. Methods Oper. Res. **60**(1), 125–144 (2004)
26. Prieto-Rumeau, T., Hernández-Lerma, O.: Ergodic control of continuous-time Markov chains with pathwise constraints. SIAM J. Control Optim. **47**(4), 1888–1908 (2008)
27. Prieto-Rumeau, T., Hernández-Lerma, O.: Selected Topics on Continuous-Time Controlled Markov Chains and Markov Games. ICP Advanced Texts in Mathematics, vol. 5. Imperial College Press, London (2012)
28. Stockbridge, R.H.: Time-average control of martingale problems: a linear programming formulation. Ann. Probab. **18**(1), 206–217 (1990)
29. Yushkevich, A.A.: Continuous time Markov decision processes with interventions. Stochastics **9**(4), 235–274 (1983)
30. Yushkevich, A.A.: Markov decision processes with both continuous and impulsive control. In: Stochastic optimization (Kiev, 1984). Lecture Notes in Control and Information Sciences, vol. 81, pp. 234–246. Springer, Berlin (1986)
31. Yushkevich, A.A.: Bellman inequalities in Markov decision deterministic drift processes. Stochastics **23**(1), 25–77 (1987)
32. Yushkevich, A.A.: Verification theorems for Markov decision processes with controllable deterministic drift, gradual and impulse controls. Teor. Veroyatnost. i Primenen. **34**(3), 528–551 (1989)