6-15-2018

Discrete-Time Hybrid Control in Borel Spaces: Average Cost Optimality Criterion

Héctor Jasso-Fuentes
CINVESTAV-IPN, hjasso@math.cinvestav.mx

José-Luis Menaldi
Wayne State University, menaldi@wayne.edu

Tomás Prieto-Rumeau
National University of Distance Education, tprieto@ccia.uned.es

Maurice Robin
Fondation Campus Paris-Saclay, maurice.robin@polytechnique.edu

Recommended Citation

H. Jasso-Fuentes, J.-L. Menaldi, T. Prieto-Rumeau, and M. Robin, Discrete-time hybrid control in Borel spaces: Average cost optimality criterion, *Journal of Mathematical Analysis and Applications*, **462** (2018), 1695-1713. doi: 10.1016/j.jmaa.2018.03.007
Available at: https://digitalcommons.wayne.edu/mathfrp/64

Discrete-time hybrid control in Borel spaces: average cost optimality criterion

Héctor Jasso-Fuentes[1], José-Luis Menaldi[2],

Tomás Prieto-Rumeau[3], and Maurice Robin[4]

February 12, 2018

Abstract

This paper addresses an optimal hybrid control problem in discrete-time with Borel state and action spaces. By hybrid we mean that the evolution of the state of the system may undergo deep changes according to structural modifications of the dynamic. Such modifications occur either by the position of the state or by means of the controller's actions. The optimality criterion is of a long-run ratio-average (or ratio-ergodic) type. We provide the existence of optimal average policies for this hybrid control problem by analyzing an associated dynamic programming equation. We also show that this problem can be translated into a standard (or non-hybrid) optimal control problem with cost constraints. Besides, we show that our model includes some special and important families of control problems, such as those with an impulsive or switching mode. Finally, to illustrate our results, we provide an example on a pollution-accumulation problem.

2010 Mathematics Subject Classification: 93E20, 34A38, 60J05.
Keywords and phrases: *Hybrid control systems, Markov decision processes, dynamic programming, ergodic (a.k.a. average) criterion.*

1 Introduction

Hybrid systems have become a common tool for the analysis and control of complex systems where both continuous components (analogue) and discrete components (digital) interact and the term "hybrid system" is used to represent a variety of cases covering many situations (see, for instance, Branicky [12], Goebel et al. [14] or Lygeros [21]).

In the hybrid system under consideration here, the evolution is given by a standard (or usual-type) sub-dynamic running under almost any situation, but, due to special events (with internal or exogenous causes), the first sub-dynamic is no longer valid and a special sub-dynamic (so-called impulse-type or event-driven-type) becomes active, overruling the standard evolution. Two usual-type controls are considered, one for each sub-dynamic, but there is

[1]Departmento de Matemáticas. CINVESTAV-IPN. A. Postal 14-740, México D.F, 07000, México. Email: hjasso@math.cinvestav.mx.

[2]Department of Mathematics, Wayne State University, Detroit, MI 48202, USA Email: menaldi@wayne.edu

[3]Department of Statistics and Operations Research. UNED, Madrid, Spain. Email: tprieto@ccia.uned.es

[4]Foundation Campus Paris-Saclay, 91190 Saint-Aubin, France. Email: maurice.robin@polytechnique.edu

another control which determines which sub-dynamic is active. This control is permitted only when the state of the system is located in some subset of the state space. Moreover, the state itself is a pair where the first component describes the standard evolution of the system (or fast-type variable) and the second component (or slow-type variable) records the structural changes, namely the activations of the special sub-dynamic. In this context, we consider the problem of finding a policy, containing a mix of standard actions and special actions, in order to minimize an infinite horizon average cost.

Hybrid control systems have been extensively studied for continuous time models; see for instance, Bensoussan and Menaldi [7, 8], Borkar et al. [11], Branicky et al. [13]. The discrete time case appears in many fewer references: Abate et al. [1, 2], Summers and Lygeros [31]. We note that many works are devoted to particular cases of hybrid control, e.g., impulsive control problems (Bensoussan [4], Bensoussan and Lions [5, 6], Menaldi [22], Robin [27, 28], Stettner [29, 30]), switching control problems (Bensoussan and Lions [6], Menaldi and Blankenship [10], Zhang et al. [35]), and standard control problems (Bensoussan [4], Hernández-Lerma [16], Hernández-Lerma and Lasserre [17, 18], Puterman [24], and the references therein). Several of these works include the case of the average cost criterion. It is also worth mentioning the references Tkachev and Abate [32] and Tkachev et al. [33], which provide insight on some interesting problems that could be studied under the perspective of hybrid control, as dealing with non-additive criteria involving reachability of subsets of the state space, or verification of specifications both on finite or infinite horizon problems.

It may be convenient to recall that, in a hybrid model in discrete time, there are transitions which do not result in an increase of a "unit of time" and, therefore, one of the key differences between conventional and hybrid models lies on the objective function, and not on the dynamic system itself (which may not be the case in a continuous-time situation). In the paper Jasso-Fuentes et al. [19], the problem of minimizing an infinite horizon discounted cost was addressed with a state dependent discount factor. For the average cost case, we consider an auxiliar ergodic dynamic programming equation (DPE) which is studied by using results on Markov decision processes (see Hernández-Lerma [16] and Hernández-Lerma and Lasserre [17]).

Our main results are the existence of an optimal average policy for the hybrid control problem, the characterization of the optimal average cost as a solution of the DPE, and the equivalence of our problem to a non-hybrid control problem with a cost constraint.

The structure of the paper is as follows: Section 2 introduces the dynamics of the system, the control policies we are going to deal with, and some preliminary assumptions. In Section 3 we define the ratio-average optimality criterion and the DPE, and we establish the existence and characterization of average optimal control policies by means of this DPE. A useful characterization that signalizes the accurate times between changes of sub-dynamics in terms of the so-named contact set is also provided. In Section 4, we present an illustrative example on pollution accumulation. Finally, Section 5 is devoted to analyze an alternative formulation as a non-hybrid control problem which takes the form of a linear programming problem over a set of measures and, to conclude, Section 6 provides a comparison of our hybrid model with impulse and switching control problems.

Notation and terminology.

- We recall that a Borel space is a measurable subset of a complete and separable metric space.

- Any metric space Z will be endowed with its Borel σ-algebra $\mathcal{B}(Z)$ and measurability (of sets and functions) will be always referred to the corresponding Borel σ-algebras.
- Given some metric space Z, the family measurable bounded functions on Z (that is, with $||u|| = \sup_{x\in Z} |u(z)| < \infty$) will be denoted by $\mathbb{B}(Z)$. If, in addition, u is nonnegative, we will write $u \in \mathbb{B}^+(X)$.
- Given a metric space Z, we define the Dirac measure $\delta_x(\cdot) : \mathcal{B}(Z) \mapsto \{0,1\}$ concentrated at $x \in Z$ by

$$\delta_x(C) := \begin{cases} 1 & \text{if } x \in C, \\ 0 & \text{otherwise.} \end{cases}$$

  Similarly, for every $C \in \mathcal{B}(Z)$, we define the indicator function $\mathbf{1}_C(\cdot) : Z \mapsto \{0,1\}$ as follows

$$\mathbf{1}_C(x) := \begin{cases} 1 & \text{if } x \in C, \\ 0 & \text{otherwise.} \end{cases}$$

- The total variation norm of a signed measure is denoted by $|| \cdot ||_{TV}$.

2 Model definition

The state and action spaces. The *state space* X of a discrete-time hybrid system is the product $X = X^f \times X^s$ of two Borel spaces, where the components $x^f \in X^f$ and $x^s \in X^s$ are called the fast (or continuous, or regular) and slow (or discrete, or impulsive) states, respectively. The *action space* A is a Borel space and it is the union of two disjoint closed subsets: $A = V^f \cup V^s$. The sets V^f and V^s are referred to as the fast and the slow action sets, respectively.

State-action pairs. The set of feasible state-action pairs is given by a measurable set $\mathbb{K} \subseteq X \times A$ with nonempty X-sections, which are denoted by $(x^f, x^s) \mapsto A(x^f, x^s) \subseteq A$ for each $(x^f, x^s) \in X$. We assume further the existence of two measurable sets

$$D^\wedge \subseteq D^\vee \subseteq X$$

such that

(a) $A(x^f, x^s) \cap V^f = \emptyset$ when $(x^f, x^s) \in D^\wedge$, meaning that when the state of the system is in $D^\wedge$, the controller must necessarily choose an action in V^s (a slow action); and

(b) $A(x^f, x^s) \cap V^s = \emptyset$ when $(x^f, x^s) \in X \setminus D^\vee$, meaning that when the state of the system is outside $D^\vee$, the controller must necessarily choose an action in V^f (a fast action).

We assume that $\mathbb{K}$ contains the graph of some measurable function from X to A. Hence, the family $\mathbb{F}$ of measurable functions $\mathtt{f} : X \to A$ such that $\mathtt{f}(x^f, x^s) \in A(x^f, x^s)$ for all $(x^f, x^s) \in X$ is nonempty.

Dynamic system. The dynamic is composed by two sub-dynamics: one sub-dynamic is of a standard type (so-called usual or traditional sub-dynamic), and it only affects the fast states $x^f \in X^f$ through the stochastic transition kernel

$$Q^f : \mathcal{B}(X^f) \times \big(\mathbb{K} \cap (X \times V^f)\big) \to [0,1],$$

while the other sub-dynamic is of a special type (so-called impulse-type or event-driven sub-dynamic) and it produces a transition of both components $(x^f, x^s) \in X$ following the stochastic kernel

$$Q^s : \mathcal{B}(X) \times \big(\mathbb{K} \cap (X \times V^s)\big) \to [0,1].$$

Summarizing, the whole dynamic is given by

$$\mathbf{Q}(dy^f \times dy^s | x^f, x^s, a) = \begin{cases} Q^f(dy^f | x^f, x^s, a)\delta_{x^s}(dy^s) & \text{if } a \in V^f, \\ Q^s(dy^f \times dy^s | x^f, x^s, a) & \text{if } a \in V^s. \end{cases} \tag{2.1}$$

Control policies. Define $H_0 = X$ and $H_k = \mathbb{K}^k \times X$ for $k \geq 1$, and let $H_\infty = \mathbb{K}^\infty$, all endowed with the corresponding product σ-algebras. The history up to step k is

$$h_k = \big(x_0^f, x_0^s, a_0, \cdots, x_{k-1}^f, x_{k-1}^s, a_{k-1}, x_k^f, x_k^s\big) \in H_k.$$

A control policy is a sequence $\{\nu_k\}_{k\geq 0}$ of transition probability measures on A given H_k such that $\nu_k\big(A(x_k^f, x_k^s)|h_k\big) = 1$ for all $h_k \in H_k$. In particular, we necessarily have

$$\nu_k\big(A(x_k^f, x_k^s) \cap V^s | h_k\big) = 1 \quad \text{if } (x_k^f, x_k^s) \in D^\wedge, \text{ and}$$

$$\nu_k\big(A(x_k^f, x_k^s) \cap V^f | h_k\big) = 1 \quad \text{if } (x_k^f, x_k^s) \in X \setminus D^\vee.$$

We denote by Π the set of admissible control policies.

By the Ionescu-Tulcea theorem, for any initial state $x = (x^f, x^s) \in X$ and any policy $\nu \in \Pi$ there exists a unique probability measure on H_∞, denoted by P_x^ν, which models the controlled dynamic system under ν. Its expectation operator is denoted by E_x^ν.

If there is some $\mathtt{f} \in \mathbb{F}$ such that the policy $\nu \in \Pi$ satisfies $\nu_k(\cdot|h_k) = \delta_{\mathtt{f}(x_k^f, x_k^s)}(\cdot)$ for any $h_k \in H_k$ and $k \geq 0$, then we say that ν is a *deterministic stationary* (a.k.a. *feedback*) policy. In what follows, we will therefore identify —without risk of confusion— the family of deterministic stationary policies with the class of functions $\mathbb{F}$. Let Φ be the family of stochastic kernels on A given X such that $\varphi(A(x)|x) = 1$ for all $x \in X$. We say that the policy $\nu \in \Pi$ is *randomized stationary* if there exists $\varphi \in \Phi$ such that $\nu_k(\cdot|h_k) = \varphi(\cdot|x_k^f, x_k^s)$ for all $k \geq 0$ and $h_k \in H_k$. We will identify the set of randomized stationary policies with Φ. Hence, we have $\mathbb{F} \subseteq \Phi \subseteq \Pi$.

Remark 2.1 *The dynamic system can be also formulated, in particular, as a system of difference equations where the stochastic kernel* $\mathbf{Q}$ *is modeled by means of two measurable functions* $F : X \times V^f \times S \to X^f$ *and* $G : X \times V^s \times S \to X$, *with* S *a Borel space, and also by a sequence* $\{w_k\}$ *of i.i.d. random variables on* S. *More specifically, for each* $k \geq 0$ *and each initial* *condition* $(x_0^f, x_0^s) = (x^f, x^s)$*, the state of the system follows the rule*

$$\underbrace{\left(x_{k+1}^f, x_{k+1}^s\right) = \left(F(x_k^f, x_k^s, a_k, w_k), x_k^s\right)}_{\textit{standard sub-dynamic}} \quad \textit{if} \;\; a \in V^f,$$

or

$$\underbrace{\left(x_{k+1}^f, x_{k+1}^s\right) = G\left(x_k^f, x_k^s, a_k, w_k\right)}_{\textit{special sub-dynamic}} \quad \textit{if} \;\; a \in V^s.$$

We will consider a measurable running cost function $\mathbf{c} : \mathbb{K} \to [0, \infty)$, which will be written as

$$\mathbf{c}(x^f, x^s, a) = c(x^f, x^s, a)\mathbf{1}_{V^f}(a) + \ell(x^f, x^s, a)\mathbf{1}_{V^s}(a),$$

with

$$c : \mathbb{K} \cap (X \times V^f) \to [0, \infty) \quad \text{and} \quad \ell : \mathbb{K} \cap (X \times V^s) \to [0, \infty)$$

interpreted as the running cost functions for the standard and the special sub-dynamics, respectively.

Our next assumption uses the following notation. Given a transition probability measure $Q(\cdot|\cdot)$ on X given X, define $Q^1 = Q$ and recursively for $q \geq 1$

$$Q^{q+1}(B|x) = \int_X Q(B|y)Q^q(dy|x) \quad \text{for } B \in \mathcal{B}(X) \text{ and } x \in X, \tag{2.2}$$

which are the successive compositions of Q with itself. We will also use the following notation. Given $\mathtt{f} \in \mathbb{F}$, the kernel on X given X defined by $\mathbf{Q}(B|x, \mathtt{f}(x))$ for $x \in X$ and $B \in \mathcal{B}(X)$ (recall (2.1)) will be denoted by $\mathbf{Q}(\cdot|\cdot, \mathtt{f})$. In particular, since $\mathbf{Q}(\cdot|\cdot, \mathtt{f})$ is itself kernel on X given X, we can define $\mathbf{Q}^q(\cdot|\cdot, \mathtt{f})$ according to (2.2).

Assumption 2.2 *There exists a control policy* $\mathtt{f} \in \mathbb{F}$*, and constants* $q \in \mathbb{N}$*,* $\epsilon > 0$ *such that*

$$\mathbf{Q}^q(D|x^f, x^s, \mathtt{f}) \leq 1 - \epsilon \quad \textit{for all } (x^f, x^s) \in X,$$

where $D = \{(x^f, x^s) \in X : \mathtt{f}(x^f, x^s) \in V^s\}$*.*

This assumption means that there exists a deterministic stationary policy in $\mathbb{F}$ such that the probability of choosing an action in V^f after q transitions is bounded away from zero uniformly in the initial state.

Assumption 2.3 *For each* $(x^f, x^s) \in X$ *we have:*

(i) *The action set* $A(x^f, x^s)$ *is compact.*

(ii) *Given a bounded and measurable function* $u : X \to \mathbb{R}$*, the function*

$$a \mapsto \int_X u(y^f, y^s)\mathbf{Q}(dy^f \times dy^s | x^f, x^s, a)$$

is continuous in $a \in A(x^f, x^s)$*.*

(iii) *The cost function* $\mathbf{c}$ *is in* $\mathbb{B}^+(\mathbb{K})$ *and the function* $a \mapsto \mathbf{c}(x^f, x^s, a)$ *is continuous on* $A(x^f, x^s)$ *for each fixed* $(x^f, x^s) \in X$*.*

(iv) *There exists a constant* $\ell_0 > 0$ *with* $\ell(x^f, x^s, a) \geq \ell_0$ *for all* $(x^f, x^s, a) \in \mathbb{K} \cap (X \times V^s)$*.*

System-time component. We will also consider a sequence $\{\mathbf{t}_k : i \geq 0\}$ of measurable functions on $\mathbb{K}^\infty$ taking values in $\mathbb{N}$, that will represent the number of times that, previous to k, an action in V^f has been taken. At this point, we will use the notation

$$\omega = \big(x_0^f, x_0^s, a_0, \ldots, x_k^f, x_k^s, a_k, \ldots\big)$$

for an element of $H_\infty = \mathbb{K}^\infty$. Given arbitrary $\omega \in H_\infty$, we put $\mathbf{t}_0(\omega) = 0$ and, for each $k \geq 1$, we let

$$\mathbf{t}_k(\omega) = \sum_{j=0}^{k-1} \mathbf{1}_{V^f}(a_j). \tag{2.3}$$

We assume that when the standard sub-dynamic is used (that is, an action in V^f is taken) then the system-time component increases by one; in other words, a time unit passes. On the contrary, when the special sub-dynamic is used (that is, an action in V^s is taken) then the system-time does not change, and this is interpreted as an *instantaneous* transition. Therefore, $\mathbf{t}_k$ represents the system-time of the hybrid control model when the controller has taken k actions; in particular, among these k actions, $\mathbf{t}_k$ of them are in V^f, while $k - \mathbf{t}_k$ of them lie in V^s. In this scenario, the variable k can be regarded as an evolution-time, in the sense that it only counts each time an action is applied, and therefore, an "evolution" of the process —representing the dynamic system— takes place.

3 The ratio-average optimality criterion

We will consider the following ratio-average optimality criterion. For an initial state $(x^f, x^s) \in X$ and a control policy $\nu \in \Pi$, let

$$J(x^f, x^s, \nu) = \limsup_{n\to\infty} \frac{E^\nu_{x^f,x^s}\big[\sum_{k=0}^{n-1} \mathbf{c}(x_k^f, x_k^s, a_k)\big]}{E^\nu_{x^f,x^s}[\mathtt{t}_n]}, \tag{3.1}$$

with $\mathtt{t}_n$ as in (2.3). This criterion is interpreted as follows: for every $n \geq 1$ we compute the expected total cost of the n actions taken by the controller, and we divide by the expected system-time of this n-th action for the hybrid model (recall the previous discussion). Such ratio-average criteria, in with both the numerator and the denominator are derived from random variables, are very usual in, e.g., semi-Markov control models (see, for instance, Luque-Vásquez and Hernández-Lerma [20] or Wei and Guo [34]). The optimal ratio-average cost function is then defined as

$$J(x^f, x^s) = \inf_{\nu\in\Pi} J(x^f, x^s, \nu) \quad \text{for } (x^f, x^s) \in X, \tag{3.2}$$

and we will say that $\nu^* \in \Pi$ is ratio-average optimal when $J(x^f, x^s, \nu^*) = J(x^f, x^s)$ for all $(x^f, x^s) \in X$.

Remark 3.1 *Under Assumption 2.3, if the policy $\nu \in \Pi$ and $x \in X$ are such that*

$$\liminf_{n\to\infty} \frac{E^\nu_x[\mathtt{t}_n]}{n} = 0 \tag{3.3}$$

then $J(x,\nu)=\infty$. To see this, observe that

$$\frac{E^{\nu}_{x^f,x^s}\big[\sum_{k=0}^{n-1}\mathbf{c}(x^f_k,x^s_k,a_k)\big]}{E^{\nu}_{x^f,x^s}[\mathtt{t}_n]} \geq \frac{E^{\nu}_{x^f,x^s}\big[\ell_0\cdot(n-\mathtt{t}_n)\big]}{E^{\nu}_{x^f,x^s}[\mathtt{t}_n]} \geq \ell_0\frac{1-\frac{1}{n}E^{\nu}_x[\mathtt{t}_n]}{\frac{1}{n}E^{\nu}_x[\mathtt{t}_n]}.$$

Through a subsequence n' such that $\frac{1}{n'}E^{\nu}_x[\mathtt{t}_{n'}]\to 0$ we have

$$\frac{E^{\nu}_{x^f,x^s}\big[\sum_{k=0}^{n'-1}\mathbf{c}(x^f_k,x^s_k,a_k)\big]}{E^{\nu}_{x^f,x^s}[\mathtt{t}_{n'}]}\to\infty$$

and, by (3.1), we obtain that $J(x,\nu)$ is infinite.

The Remark 3.1 means that the cost is finite only if the usual dynamic is active often enough, as expressed in condition (3.3), where the key assumption $\ell(\cdot)\geq\ell_0>0$ (see Assumption 2.3(iv)) plays an important role. As a consequence of Remark 3.1, we will focus only on policies ν such that the lim inf in (3.3) is strictly positive. In fact, under Assumption 2.2, our next result shows that the class of such ν is nonempty.

Proposition 3.2 *If $\mathtt{f}\in\mathbb{F}$ is the control policy defined in Assumption 2.2 then*

$$\liminf_{n\to\infty}\frac{E^{\mathtt{f}}_x[\mathtt{t}_n]}{n}>0 \quad \textit{for all } x=(x^f,x^s)\in X.$$

Moreover, the criterion (3.1) is uniformly bounded if Assumption 2.3 is also satisfied.

Proof. Choose $n>q$ and any $x\in X$. By definition

$$E^{\mathtt{f}}_x[\mathtt{t}_n]=\sum_{k=0}^{n-1}P^{\mathtt{f}}_x\{\mathtt{f}(x_k)\in V^f\}=\sum_{k=0}^{n-1}P^{\mathtt{f}}_x\{x_k\in D^c\}.$$

If $n-1=p\cdot q+r$ for some integers $p\geq 1$ and $0\leq r<q$, then

$$E^{\mathtt{f}}_x[\mathtt{t}_n]\geq\sum_{j=1}^{p}P^{\mathtt{f}}_x\{x_{jq}\in D^c\}.$$

By Assumption 2.2, each $P^{\mathtt{f}}_x\{x_{jq}\in D^c\}$ is larger than or equal to ϵ, and so

$$E^{\mathtt{f}}_x[\mathtt{t}_n]\geq p\epsilon\geq\frac{n-1-q}{q}\cdot\epsilon.$$

We conclude that

$$\liminf_{n\to\infty}\frac{E^{\mathtt{f}}_x[\mathtt{t}_n]}{n}\geq\epsilon/q. \tag{3.4}$$

Finally, since $||\mathbf{c}||<\infty$ we conclude that

$$J(x,\mathtt{f})\leq\frac{q||\mathbf{c}||}{\epsilon},$$

and so $J(x,\mathtt{f})$ is finite for every $x\in X$, and $J(x)\leq\frac{q||\mathbf{c}||}{\epsilon}$ for all $x\in X$. □

Dynamic programming equations. For the ratio-average optimality criterion, it seems that there is not a "usual" average-cost dynamic programming optimality equation. Moreover, the vanishing discount approach technique neither leads to a limit average-cost equation adapted to the ratio-average optimality criterion, because of the high dependence of the discount factor on the state-action —see Jasso-Fuentes et al. [19].

To overcome this situation, we will introduce a parametric family of average-cost dynamic programming optimality equations (depending on some $\lambda \geq 0$) and we will prove that there exists some $\lambda^* \geq 0$ which provides a solution to the ratio-average cost hybrid control problem. Later, interpreting the parameter $\lambda \geq 0$ as a Langrange multiplier, we will show that the hybrid control problem is equivalent to a non-hybrid constrained control problem.

Fix a parameter $\lambda \geq 0$ and define the function $\mathbf{c}_\lambda : \mathbb{K} \to \mathbb{R}$ as $\mathbf{c}_\lambda(x,a) := \mathbf{c}(x,a) - \lambda \mathbf{1}_{V^f}(a)$. Given a function $u \in \mathbb{B}(X)$ we define the dynamic programming operator $\mathcal{T}_\lambda u$ on X as follows:

$$\mathcal{T}_\lambda u(x) = \inf_{a \in A(x)} \left\{ \mathbf{c}_\lambda(x,a) + \int_X u(y) \mathbf{Q}(dy \,|\, x, a) \right\} \quad \text{for } x \in X.$$

Taking into account the nature of the hybrid control model, we can define two associated operators $\mathcal{M}u$ and $\mathcal{H}_\lambda u$ as

$$\mathcal{M}u(x^f, x^s) = \inf_{a \in A(x^f, x^s) \cap V^s} \left\{ \ell(x^f, x^s, a) + \int_X u(y^f, y^s) Q^s(dy^f \times dy^s \,|\, x^f, x^s, a) \right\} \tag{3.5}$$

defined on $D^\vee$, and

$$\mathcal{H}_\lambda u(x^f, x^s) = \inf_{a \in A(x^f, x^s) \cap V^f} \left\{ c(x^f, x^s, a) - \lambda + \int_{X^f} u(y^f, x^s) Q^f(dy^f | x^f, x^s, a) \right\} \tag{3.6}$$

defined on $X \setminus D^\wedge$. Therefore, the dynamic programming operator $\mathcal{T}_\lambda$ can be written, for $u \in \mathbb{B}(X)$ and $(x^f, x^s) \in X$, as

$$\mathcal{T}_\lambda u(x^f, x^s) = \begin{cases} \mathcal{M}u(x^f, x^s), & \text{if } (x^f, x^s) \in D^\wedge, \\ \min\left\{ \mathcal{M}u(x^f, x^s), \mathcal{H}_\lambda u(x^f, x^s) \right\}, & \text{if } (x^f, x^s) \in D^\vee \setminus D^\wedge, \\ \mathcal{H}_\lambda u(x^f, x^s), & \text{if } (x^f, x^s) \in X \setminus D^\vee. \end{cases} \tag{3.7}$$

A pair $(\rho, u) \in \mathbb{R} \times \mathbb{B}(X)$ is said to be a solution of the *dynamic programming equation* (DPE) when

$$\rho + u(x^f, x^s) = \mathcal{T}_\lambda u(x^f, x^s) \quad \text{for all } (x^f, x^s) \in X, \tag{3.8}$$

which will be written, in short, as $\rho + u = \mathcal{T}_\lambda u$.

The following result summarizes some properties of the operators $\mathcal{M}$ and $\mathcal{H}_\lambda$. The proof follows from Proposition D.5 in Hernández-Lerma and Lasserre [17].

Lemma 3.3 *Under Assumption 2.3, for every $u \in \mathbb{B}(X)$ and $\lambda \geq 0$ the functions $\mathcal{M}u$, $\mathcal{H}_\lambda u$, and $\mathcal{T}_\lambda$ are measurable on their respective domains, and there exists $\mathtt{f}^* \in \mathbb{F}$ such that $\mathtt{f}^*(x)$ attains the minimum in the definition of $\mathcal{T}_\lambda u(x)$ for every $x \in X$; see (3.5)–(3.8).*

Finally we impose an ergodicity condition (cf. Assumption 3.1.4 in Hernández-Lerma [16]).

Assumption 3.4 *There exists* $0 < \beta < 1$ *such that*

$$\sup_{(x,a),(x',a')\in\mathbb{K}} ||\mathbf{Q}(\cdot|x,a) - \mathbf{Q}(\cdot|x',a')||_{TV} \leq 2\beta,$$

where the norm $\|\cdot\|_{TV}$ *was defined at the end of Section 1.*

Our next result ensures the existence of a solution to the DPE and provides a characterization of the constant ρ for each fixed λ.

Lemma 3.5 *Suppose that Assumptions 2.2, 2.3, and 3.4 are satisfied. Then the following assertions hold:*

(i) *For each* $\lambda \geq 0$*, there exists a pair* $(\rho_\lambda, u_\lambda) \in \mathbb{R} \times \mathbb{B}(X)$ *that satisfies the DPE* (3.8)*; i.e.,* $\rho_\lambda + u_\lambda = \mathcal{T}_\lambda u_\lambda$*. Furthermore, the constant* ρ_λ *satisfies:*

$$\rho_\lambda = \inf_{\nu\in\Pi} \limsup_{n\to\infty} \frac{1}{n} E_x^\nu \Big[\sum_{k=0}^{n-1} [\mathbf{c}(x_k, a_k) - \lambda \mathbf{1}_{V^f}(a_k)] \Big].$$

(ii) *There exists* $\lambda^* \geq 0$ *such that* $\rho_{\lambda^*} = 0$*, for which the DPE* (3.8) *becomes*

$$u_{\lambda^*} = \mathcal{T}_{\lambda^*} u_{\lambda^*}. \tag{3.9}$$

Proof. (i). This part is a direct consequence of Theorem 2.2 and Corollary 3.6 in Hernández-Lerma [16].

(ii). First of all, we are going to show that $\lambda \mapsto \rho_\lambda$ is a concave function. Choose λ_1, λ_2 in $[0,\infty)$ and $0 \leq \gamma \leq 1$. Using part (i), we have for every $(x,a) \in \mathbb{K}$

$$\begin{aligned} \rho_{\lambda_1} + u_{\lambda_1}(x) &\leq \mathbf{c}(x,a) - \lambda_1 \mathbf{1}_{V^f}(a) + \int_X u_{\lambda_1}(y) \mathbf{Q}(dy|x,a) \\ \rho_{\lambda_2} + u_{\lambda_2}(x) &\leq \mathbf{c}(x,a) - \lambda_2 \mathbf{1}_{V^f}(a) + \int_X u_{\lambda_2}(y) \mathbf{Q}(dy|x,a) \end{aligned}$$

Letting $u^* = \gamma u_{\lambda_1} + (1-\gamma) u_{\lambda_2} \in \mathbb{B}(X)$, this implies that

$$\begin{aligned} \gamma\rho_{\lambda_1} + (1-\gamma)\rho_{\lambda_2} + u^*(x) &\leq \mathbf{c}(x,a) - (\gamma\lambda_1 + (1-\gamma)\lambda_2) \mathbf{1}_{V^f}(a) + \int_X u^*(y) \mathbf{Q}(dy|x,a) \\ &= \mathbf{c}_{\gamma\lambda_1+(1-\gamma)\lambda_2}(x,a) + \int_X u^*(y) \mathbf{Q}(dy|x,a). \end{aligned}$$

Since this hold for every $(x,a) \in \mathbb{K}$, we deduce

$$\gamma\rho_{\lambda_1} + (1-\gamma)\rho_{\lambda_2} + u^* \leq \mathcal{T}_{\gamma\lambda_1+(1-\gamma)\lambda_2} u^*. \tag{3.10}$$

Now we are going to iterate this inequality and use a standard dynamic programming argument (as in, e.g., the proof of Theorem 2.2.a in Hernández-Lerma [16]). First we will use (3.10) to show that for any $\nu \in \Pi$, $x \in X$, and $n \geq 1$ we have

$$n\cdot(\gamma\rho_{\lambda_1} + (1-\gamma)\rho_{\lambda_2}) + u^*(x) \leq E_x^\nu \Big[\sum_{k=0}^{n-1} [\mathbf{c}(x_k, a_k) - (\gamma\lambda_1 + (1-\gamma)\lambda_2) \mathbf{1}_{V^f}(a_k)] + u^*(x_n) \Big], \tag{3.11}$$

and we will do it by induction on n. Clearly, (3.11) holds for $n=1$ as a consequence of (3.10) because

$$\begin{aligned} E_x^\nu[\mathbf{c}(x_0,a_0)-(\gamma\lambda_1+(1-\gamma)\lambda_2)\mathbf{1}_{V^f}(a_0)+u^*(x_1)] &\geq \mathcal{T}_{\gamma\lambda_1+(1-\gamma)\lambda_2}u^*(x) \\ &\geq \gamma\rho_{\lambda_1}+(1-\gamma)\rho_{\lambda_2}+u^*(x). \end{aligned}$$

Suppose now that (3.11) holds for some $n\geq 1$. Observing that

$$\begin{aligned} &E_x^\nu[\mathbf{c}(x_n,a_n)-(\gamma\lambda_1+(1-\gamma)\lambda_2)\mathbf{1}_{V^f}(a_n)+u^*(x_{n+1})] \\ &\geq E_x^\nu[\mathcal{T}_{\gamma\lambda_1+(1-\gamma)\lambda_2}u^*(x_n)]\geq \gamma\rho_{\lambda_1}+(1-\gamma)\rho_{\lambda_2}+E_x^\nu[u^*(x_n)], \end{aligned}$$

and by the induction argument, it turns out that (3.11) is satisfied for $n+1$.

We proceed with the proof. In (3.11), divide by n and take the lim sup as $n\to\infty$. Recalling that the function u^* is bounded, this yields

$$\gamma\rho_{\lambda_1}+(1-\gamma)\rho_{\lambda_2}\leq \limsup_{n\to\infty}\frac{1}{n}E_x^\nu\Big[\sum_{k=0}^{n-1}[\mathbf{c}(x_k,a_k)-(\gamma\lambda_1+(1-\gamma)\lambda_2)\mathbf{1}_{V^f}(a_k)]\Big].$$

But since $\nu\in\Pi$ and $x\in X$ are arbitrary, Lemma 3.5(i) implies that

$$\gamma\rho_{\lambda_1}+(1-\gamma)\rho_{\lambda_2}\leq\rho_{\gamma\lambda_1+(1-\gamma)\lambda_2}.$$

This shows concavity of ρ_λ.

Note now that the cost function $\mathbf{c}_\lambda$ being a monotone nonincreasing function of $\lambda\geq 0$, this implies that the infimum ρ_λ of the long-run average costs (as in Lemma 3.5(i)) is also a monotone nonincreasing function of λ. Summarizing the function $\lambda\mapsto\rho_\lambda$ has been shown to be concave and monotone nonincreasing on $[0,\infty)$. In particular, it must necessarily be continuous on $[0,\infty)$.

On the other hand, for the policy $\mathtt{f}$ in Assumption 2.2 and for $\lambda>0$

$$\frac{1}{n}E_x^{\mathtt{f}}\Big[\sum_{k=0}^{n-1}[\mathbf{c}(x_k,a_k)-\lambda\mathbf{1}_{V^f}(a_k)]\Big] \;=\; \frac{1}{n}E_x^{\mathtt{f}}[\sum_{k=0}^{n-1}\mathbf{c}(x_k,a_k)]-\lambda\frac{1}{n}E_x^{\mathtt{f}}[\mathtt{t}_n].$$

Taking the lim sup as $n\to\infty$ and using part (i) of this lemma, we conclude

$$\begin{aligned} \rho_\lambda &\leq \limsup\frac{1}{n}E_x^{\mathtt{f}}\Big[\sum_{k=0}^{n-1}[\mathbf{c}(x_k,a_k)-\lambda\mathbf{1}_{V^f}(a_k)]\Big] \\ &\leq \limsup\frac{1}{n}E_x^{\mathtt{f}}[\sum_{k=0}^{n-1}\mathbf{c}(x_k,a_k)]-\lambda\liminf\frac{1}{n}E_x^{\mathtt{f}}[\mathtt{t}_n] \\ &\leq \limsup\frac{1}{n}E_x^{\mathtt{f}}[\sum_{k=0}^{n-1}\mathbf{c}(x_k,a_k)]-\lambda\frac{\epsilon}{q}, \end{aligned}$$

where we have used (3.4). Hence, $\rho_\lambda\leq||\mathbf{c}||-\lambda\frac{\epsilon}{q}$. Since the cost function $\mathbf{c}$ is nonnegative, we have that $\rho_0\geq 0$; together with the latter inequality and continuity of ρ_λ on $[0,\infty)$, this yields the existence of some $\lambda^*\geq 0$ such that ρ_{λ^*} vanishes. □

In Theorem 3.7(b) below we will be able to show that there is, in fact, a unique $\lambda^*\geq 0$ such that $\rho_{\lambda^*}=0$.

Definition 3.6 *Any policy* $\mathtt{f} \in \mathbb{F}$ *attaining the minimum in the DPE* (3.9) *for some* $\lambda^* \geq 0$ *with* $\rho_{\lambda^*} = 0$ *will be called* <u>canonical</u>.

We have arrived to our main optimality results for the ratio-average optimality criterion given in (3.1)–(3.2).

Theorem 3.7 *Suppose that Assumptions 2.2, 2.3, and 3.4 hold. Then the following statements hold true.*

(a) *Given* $\lambda^* \geq 0$ *as in Lemma 3.5(ii), for every* $\nu \in \Pi$ *and* $x \in X$ *we have* $J(x,\nu) \geq \lambda^*$.

(b) *Every canonical policy is ratio-average optimal and the optimal ratio-average cost function* J *in* (3.2) *equals the constant* λ^*. *Hence, the solution* $\rho_\lambda = 0$ *for* $\lambda \geq 0$ *is unique.*

(c) *There exists a ratio-average optimal policy in the set of stationary policies* $\mathbb{F}$.

Proof. (a). The dynamic programming equation (3.8) for λ^* reads $u_{\lambda^*} = \mathcal{T}_{\lambda^*} u_{\lambda^*}$ because $\rho_{\lambda^*} = 0$ (see (3.9)). Given a policy $\nu \in \Pi$ and $x \in X$, we will suppose that $\liminf_{n\to\infty} E_x^\nu[\mathtt{t}_n]/n > 0$ (otherwise, by Remark 3.1, we trivially have $J(x,\nu) \geq \lambda^*$).

Now we proceed as in (3.11), but this time starting from the equality $u_{\lambda^*} = \mathcal{T}_{\lambda^*} u_{\lambda^*}$, to obtain that

$$0 \leq -u_{\lambda^*}(x) + E_x^\nu\left[\sum_{k=0}^{n-1} \mathbf{c}(x_k,a_k)\right] - \lambda^* E_x^\nu[\mathtt{t}_n] + E_x^\nu[u_{\lambda^*}(x_n)].$$

Hence, for large enough n,

$$\lambda^* \leq -\frac{1}{E_x^\nu[\mathtt{t}_n]} u_{\lambda^*}(x) + \frac{1}{E_x^\nu[\mathtt{t}_n]} E_x^\nu\left[\sum_{k=0}^{n-1} \mathbf{c}(x_k,a_k)\right] + \frac{1}{E_x^\nu[\mathtt{t}_n]} E_x^\nu[u_{\lambda^*}(x_n)]. \tag{3.12}$$

Since $\lim_{n\to\infty} E_x^\nu[\mathtt{t}_n] = \infty$, taking the limsup as $n \to \infty$ in (3.12) we derive $\lambda^* \leq J(x,\nu)$, which proves (a).

(b). Consider a canonical policy $\mathtt{f}^* \in \mathbb{F}$. Then, by definition, it attains the minimum in the equation $u_{\lambda^*} = \mathcal{T}_{\lambda^*} u_{\lambda^*}$. The average cost of $\mathtt{f}^*$ for the cost function $\mathbf{c}_{\lambda^*}$ is zero (recall that $\rho_{\lambda^*} = 0$), and so for any $x \in X$,

$$\limsup_{n\to\infty} \frac{E_x^{\mathtt{f}^*}\left[\sum_{k=0}^{n-1} \mathbf{c}(x_k,a_k) - \lambda^* \mathtt{t}_n\right]}{n} = 0.$$

But,

$$\frac{E_x^{\mathtt{f}^*}\left[\sum_{k=0}^{n-1} \mathbf{c}(x_k,a_k) - \lambda^* \mathtt{t}_n\right]}{n} \geq \frac{E_x^{\mathtt{f}^*}\left[\ell_0 \cdot (n - \mathtt{t}_n) - \lambda^* \mathtt{t}_n\right]}{n} = \ell_0 - (\ell_0 + \lambda^*) \frac{E_x^{\mathtt{f}^*}[\mathtt{t}_n]}{n}.$$

Taking the lim sup in the above inequality yields

$$0 \geq \ell_0 - (\ell_0 + \lambda^*) \liminf_{n\to\infty} \frac{E_x^{\mathtt{f}^*}[\mathtt{t}_n]}{n}.$$

This shows that it is not possible to have $\liminf_{n\to\infty} E_x^{\mathtt{f}^*}[\mathtt{t}_n]/n = 0$. Now, using the fact that $\liminf_{n\to\infty} E_x^{\mathtt{f}^*}[\mathtt{t}_n]/n > 0$, and arguing as in part (a) of this theorem, we obtain

$$\lambda^* = -\frac{1}{E_x^{\mathtt{f}^*}[\mathtt{t}_n]} u_{\lambda^*}(x) + \frac{1}{E_x^{\mathtt{f}^*}[\mathtt{t}_n]} E_x^{\mathtt{f}^*}\left[\sum_{k=0}^{n-1} \mathbf{c}(x_k,a_k)\right] + \frac{1}{E_x^{\mathtt{f}^*}[\mathtt{t}_n]} E_x^{\mathtt{f}^*}[u_{\lambda^*}(x_n)]$$

and then, proceeding as in (3.12), we can conclude that $\lambda^* = J(x, \mathtt{f}^*)$ for all $x \in X$. This last fact together with part (a) yields (b).

(c). This claim is a direct consequence of Lemma 3.3 and part (b) of this theorem. □

Theorem 3.7 gives sufficient conditions ensuring that the optimal ratio-average cost is the solution of a dynamic programming equation, and shows the existence of an optimal deterministic stationary policy from the fixed point equation $u^* = \mathcal{T}_{\lambda^*} u^*$. With the notation introduced in (3.5) and (3.6), this equation reads (cf. (3.7))

$$u^*(x^f, x^s) = \begin{cases} \mathcal{M}u^*(x^f, x^s), & \text{if } (x^f, x^s) \in D^\wedge, \\ \min\left\{\mathcal{M}u^*(x^f, x^s), \mathcal{H}_{\lambda^*}u^*(x^f, x^s)\right\}, & \text{if } (x^f, x^s) \in D^\vee \setminus D^\wedge, \\ \mathcal{H}_{\lambda^*}u^*(x^f, x^s), & \text{if } (x^f, x^s) \in X \setminus D^\vee. \end{cases}$$

The following is a direct consequence of both Theorem 3.7 and the definition of a canonical policy.

Corollary 3.8 *Let $\mathtt{f} \in \mathbb{F}$ be a canonical policy. On the set $D^\vee \setminus D^\wedge$, the following holds:*

(a). *If $\mathcal{M}u^*(x) > \mathcal{H}_{\lambda^*}u^*(x) = u^*(x)$, then the optimal action $\mathtt{f}(x)$ is in V^f.*

(b). *If $\mathcal{H}_{\lambda^*}u^*(x) > \mathcal{M}u^*(x) = u^*(x)$, then the optimal action $\mathtt{f}(x)$ is in V^s.*

Some comments of the previous results are the following:

Remark 3.9 **(i)** *The case $\mathcal{M}u^*(x) = \mathcal{H}_{\lambda^*}u^*(x) = u^*(x)$ is a little particular. In this scenario, there is an optimal action in V^f and there is also an optimal action in V^s, so that the decisor can run either the usual or the special sub-dynamic.*

(ii) *Obviously, when the state is in $D^\wedge$ or in $X \setminus D^\vee$, then optimal actions are necessarily in V^s or in V^f, respectively.*

Contact set and continuation region. We introduce the set D^* defined as:

$$D^* = \left\{x \in D^\vee : u^*(x) = \mathcal{M}u^*(x)\right\},$$

with $x := (x^f, x^s)$. This set is so-named *contact set* and it can be regarded as an *optimal* region, in the sense that outside it, the optimal choice is to apply a fast action in V^f (and so the standard sub-dynamic is activated) and once the state of the system reaches D^*, an optimal rule is to apply an action in V^s (and thus the special sub-dynamic is turned on). This fact can be summarized as follows.

Proposition 3.10 *For any canonical policy, an optimal rule outside D^* must necessarily be in V^f, whereas inside D^* the optimal rule can be taken in V^s.*

Proof. Note that the three possibilities (a), (b) and (i) described in both Corollary 3.8 and Remark 3.9 are mutually exclusive. As a consequence, the result follows because (b) and (i) correspond to the case when $x \in D^*$, while (a) holds when $x \notin D^*$. □

Note that Proposition 3.10 does not give any topological property of the contact set D^*. To provide more regularity on this set, we impose some additional conditions.

Proposition 3.11 *Suppose that Assumptions 2.2, 2.3, and 3.4 hold where the Assumptions 2.3(i)–(iii) are replaced with:*

(i)' *The multifunction $x \mapsto A(x)$ is continuous and compact-valued.*

(ii)' *For any $u \in \mathbb{B}(X)$, the mapping $(x,a) \mapsto \int_X u(y)\mathbf{Q}(dy|x,a)$ is continuous on $\mathbb{K}$.*

(iii)' *The cost function is in $\mathbb{B}^+(\mathbb{K})$ and it is continuous on $\mathbb{K}$.*

Under these conditions, the contact set D^* is closed.

Proof. Use Proposition D.3 in Hernández-Lerma [16] to prove that u^* and $\mathcal{M}u^*$ are continuous and so D^* is the inverse image of $\{0\}$ for the continuous function $u^* - \mathcal{M}u^*$. □

Observe that, under the conditions of this proposition, the sets defining (a) and (b) in Corollary 3.8 above are open, while the set given in the Remark 3.9(i) is closed.

Time-interface set. The time-interface set consists of real positive numbers $\{\tau_n\}_{n\geq 0}$ that signalize the times when the special sub-dynamic must be activated in an *optimal* way. This sequence can be obtained in terms of the contact set D^* as follows: Let

$$\omega = \big(x_0, a_0, \cdots, x_k, a_k, \cdots\big)$$

be an element of $H_\infty = \mathbb{K}^\infty$, with $x_k := (x_k^f, x_k^s)$, we recursively define:

$$\tau_i \equiv \tau(\omega, i) = \inf\{k \geq \tau(\omega, i-1) + 1 : J^*(x_k) = \mathcal{M}J^*(x_k)\} \quad \text{for all } i \geq 0,$$

with $\tau(\omega, -1) = -1$, which is regarded as a stopping time relative to the history, i.e., $\tau(\omega, i)$ is a random variable with values in $\{0, 1, 2, \cdots, \infty\}$.

4 Example: a pollution accumulation problem

Suppose that an economy consumes a specific good or product and that, as a byproduct of this consumption, it generates pollution. We assume that the pollution stock x_k is gradually degraded and is represented by the system

$$x_{k+1}^f = p(a_k) - g(x_k^s)x_k^f + \xi_k \qquad k \geq 0,$$

where the variable x^f denotes the stock of pollution with values in $X^f = [0, \infty)$, and x^s stands for the levels (or modes) of environmental contingency decided by the government, with values in $X^s = M = \{1, 2, \cdots, l\}$. Also, the quantity $a_k \geq 0$ denotes the consumption rate at time k with range in $[0, \gamma(x^f, x^s)]$, where $\gamma(x^f, x^s) > 0$ is a constant usually imposed by international protocols, and $p(a_k)$ is the amount of pollution derived to consume the quantity a_k. For each $i \in M$, there is a decay rate of pollution associated to level i that is represented by the function $g(i) \geq 0$. Finally, the sequence $\{\xi_k\}$ is a sequence of i.i.d. random variables that measure external events that are not predicted in the model.

The government is capable to change among the different modes of environmental contingency: such changes are instantaneous in time, and they produce a cost of switching between modes i to j, denoted by $\ell(i, j) > 0$. In addition, the new mode j can be restricted according to the value of the stock x^f; i.e., the action j must be restricted to the set $A(x^f, i) \cap X^s$ (in this example $V^s = X^s$).

As a response to the pollution levels, the economy has a disutility $D(x, i)$ and the objective is to find a consumption-mode policy $\{a_k\}$ with values in $[0, \gamma(x^f, x^s)]$ or in $A(x^f, x^s) \cap V^s$ that minimizes the long-run ratio-average expected cost (3.1) for $\mathbf{c} = D + \ell$.

The whole dynamic can be rewritten as follows:

$$\underbrace{\left(x^f_{k+1}, x^s_{k+1}\right) = \left(p(a_k) - g(x^s_k)x^f_k + \xi_k, x^s_k\right)}_{\text{standard sub-dynamic}} \quad \text{if} \;\; a_k \in [0, \gamma(x^f, x^s)],$$

or

$$\underbrace{\left(x^f_{k+1}, x^s_{k+1}\right) = \left(x^f_k, a_k\right)}_{\text{special sub-dynamic}} \quad \text{if} \;\; a_k \in X^s = M, \qquad k \geq 0.$$

This problem has been studied for many different classes of dynamics (discrete-time and continuous-time models) when a utility is also considered in the model (see, for instance, Kawaguchi and Morimoto [15] and their references). However, in all these models, the environmental contingency decided by the government has not been considered. The use of average payoff criteria is useful when we are taking into account future generations. This consideration may have ethic and economic impacts.

Finally, the sets $D^\wedge$ and $D^\vee$ can be regarded in this context as follows: $D^\wedge = [K_1, \infty) \times M$ and $D^\vee = [K_2, \infty) \times M$, with $K_2 \leq K_1$. In this case, if the mode is $i \in M$ at time k, and the stock of pollution x^f_k is in $[0, K_2)$ (low level of pollution), the rule suggests to keep consuming the quantity a_k and to keep the same mode i, but once the state $x^f_k \in [K_1, \infty)$ (high level of pollution), the rule is to change immediately the mode from i to some other state $j \in M$, $j \neq i$. Such a change also depends on the set $A(x^f, x^s)$ whose role is not allowing situations to change into worse modes that can increase the pollution; whereas on $[K_1, K_2] \times M$, the government has the option either to keep consuming or to change the environmental contingency mode, this choice may be highly dependent on which of the two actions produce a lower cost.

The advantage of this new model is that it considers situations when the pollution can be classified by levels and, according to each level, both the decay rate and the level of consumption change; for instance, the higher the level of pollution, the stronger the restriction for the consumption, a rather realistic situation.

5 Alternative non-hybrid formulation

Throughout this section, we will restrict ourselves to the class stationary policies Φ (see Section 2 for a detailed description of this class of policies).

By our ergodicity Assumption 3.4, we know that for each $\varphi \in \Phi$ the state process $\{(x^f_k, x^s_k)\}_{k\in\mathbb{N}}$ is an ergodic Markov chain with invariant probability measure μ_φ on $\mathcal{B}(X)$ which satisfies

$$\mu_\varphi(B) = \int_X \int_{A(x)} \mathbf{Q}(B|x, a)\varphi(da|x)\mu_\varphi(dx) \quad \text{for all } B \in \mathcal{B}(X).$$

Conversely, if a probability measure μ on $X \times A$ satisfies $\mu(\mathbb{K}) = 1$ and

$$\mu(B \times A) = \int_{\mathbb{K}} \mathbf{Q}(B|x, a)\mu(dx, da) \quad \text{for all } B \in \mathcal{B}(X), \tag{5.1}$$

then its marginal on X, i.e.,

$$\hat{\mu}(B) = \mu(B \times A) \quad \text{for } B \in \mathcal{B}(X)$$

is an invariant probability measure of the state process $\{(x_k^f, x_k^s)\}_{k\in\mathbb{N}}$ under some $\varphi \in \Phi$ (for further details, see Hernández-Lerma and Lasserre [17], Section 6.4). We will denote by Δ the (convex) set of probability measures on $\mathcal{B}(\mathbb{K})$ satisfying (5.1).

On the other hand, given any function $v \in \mathbb{B}^+(\mathbb{K})$, interpreted as a cost function, the long-run expected average cost of a policy $\varphi \in \Phi$ is constant (it does not depend on the initial state) and equals

$$\int_X \int_{A(x)} v(x,a)\varphi(da|x)\mu_\varphi(dx).$$

Therefore, the control problem of minimizing the long-run expected average cost for v is equivalent to the linear programming problem given by

$$\text{minimize} \int_{\mathbb{K}} v(x,a)\mu(dx,da) \quad \text{subject to} \quad \mu \in \Delta.$$

In addition, we can incorporate a constraint function $v_1 \in \mathbb{B}^+(X)$ and consider the constrained control problem to minimize the long-run average cost of v subject that the long-run average cost of v_1 is larger than or equal to some constant $\theta \in [0,1]$. This constrained problem is equivalent to the linear programming problem

$$\text{minimize} \int_{\mathbb{K}} v(x,a)\mu(dx,da) \quad \text{subject to} \quad \mu \in \Delta \quad \text{and} \quad \int_{\mathbb{K}} v_1(x,a)\mu(dx,da) \geq \theta.$$

In our context, under the additional condition that Proposition 3.11(ii)' holds (which ensures that Δ is compact with the weak topology), we will consider $v = \mathbf{c}$, the running cost function, and $v_1(x,a) = \mathbf{1}_{V^f}(a)$ for $(x,a) \in \mathbb{K}$. For $0 \leq \theta \leq 1$ define

$$V(\theta) = \inf_{\mu} \Big\{ \int_{\mathbb{K}} \mathbf{c}(x,a)\mu(dx,da) \ : \ \mu \in \Delta \quad \text{and} \quad \int_{\mathbb{K}} \mathbf{1}_{V^f}(a)\mu(dx,da) \geq \theta \Big\}. \tag{5.2}$$

Observe that the restriction $\int_{\mathbb{K}} \mathbf{1}_{V^f}(a)\mu(dx,da) \geq \theta$ can be interpreted as follows: the long-run expected proportion of actions taken in V^f is larger than or equal to θ:

$$\lim_{n\to\infty} \frac{1}{n} E_x^\varphi \Big[\sum_{k=0}^{n-1} \mathbf{1}_{V^f}(a_k) \Big] \geq \theta.$$

Note also that the set $\{(x,a) \in \mathbb{K} : a \in V^f\}$ is closed because the sets V^f and V^s are disjoint and closed. Then, by the Portmanteau theorem, it follows that the set of μ with $\int_{\mathbb{K}} \mathbf{1}_{V^f}(a)\mu(dx,da) \geq \theta$ is closed. In particular, the infimum in (5.2) is attained because it is the minimum of a continuous function on a compact space.

Proposition 5.1 *Suppose that Assumptions 2.2, 2.3, and 3.4 hold, and that the conditions of Proposition 3.11 are also satisfied.*

(i) *The function V is monotone nondecreasing, convex and continuous on $[0,1]$.*

(ii) *For some $0 \leq \theta^* \leq 1$ and some subderivative d_{θ^*} of V at θ^* we have $V(\theta^*) = \theta^* d_{\theta^*}$.*

Proof. (i). The fact that V is monotone nondecreasing is straightforward because the larger θ the smaller the feasible region.

Given $\theta_1, \theta_2 \in [0,1]$ and $0 \le \gamma \le 1$, for some $\mu_1, \mu_2 \in \Delta$ we have

$$V(\theta_1) = \int \mathbf{c} d\mu_1 \quad \text{and} \quad \int \mathbf{1}_{V^f}(a) d\mu_1 \ge \theta_1$$

and

$$V(\theta_2) = \int \mathbf{c} d\mu_2 \quad \text{and} \quad \int \mathbf{1}_{V^f}(a) d\mu_2 \ge \theta_2.$$

The measure $\mu = \gamma\mu_1 + (1-\gamma)\mu_2 \in \Delta$ verifies $\int v_1 d\mu \ge \gamma\theta_1 + (1-\gamma)\theta_2$, and so

$$V(\gamma\theta_1 + (1-\gamma)\theta_2) \le \int \mathbf{c} d\mu = \gamma V(\theta_1) + (1-\gamma)V(\theta_2),$$

which proves the convexity of V.

This last property makes the function V to be continuous on $(0,1)$. In addition, since V is monotone nondecreasing, it must be necessarily continuous at 0. Let us now prove continuity of V at 1. It is easy to see (again by the Portmanteau theorem) that if $\theta_n \uparrow 1$ and μ_n attains the infimum in (5.2) for θ_n, then the limit though some subsequence $\mu_{n'} \to \mu$ satisfies the constraint $\int_{\mathbb{K}} \mathbf{1}_{V^f}(a)\mu(dx,da) \ge 1$ with $V(1) \le \int_{\mathbb{K}} \mathbf{c} d\mu = V(1^-)$. By monotonicity of V this shows that, necessarily, $V(1^-) = V(1)$, thus proving continuity of V.

(ii). If $V(0) = 0$ then $\theta^* = 0$ satisfies the conditions given in (ii). So, suppose that $V(0) > 0$. It is easy to see that the set:

$$\{\alpha > 0 : V(\theta) = \alpha\theta \text{ for some } 0 \le \theta \le 1\}$$

is of the form $[\alpha_0, \infty)$ for some $\alpha_0 > 0$. We can deduce that, for $\theta^* > 0$ such that $V(\theta^*) = \alpha_0\theta^*$, the epigraph of V is contained in the half-space $\{y \ge \alpha_0 x\}$ and so α_0 is necessarily a subderivative of V at θ^*. This completes the proof. □

Proposition 5.2 *Assume the conditions of Proposition 5.1 are valid. Then, for every* $0 \le \theta \le 1$, *there exists some* $d_\theta \ge 0$ *and* $u \in \mathbb{B}(X)$ *such that*

$$V(\theta) + u(x) = \inf_{a \in A(x)} \Big\{ \mathbf{c}(x,a) - d_\theta(\mathbf{1}_{V^f}(a) - \theta) + \int_X u(y)\mathbf{Q}(dy|x,a) \Big\}$$

for all $x \in X$.

Proof. Given $\mu \in \Delta$ observe that $(\int \mathbf{1}_{V^f}(a)d\mu, \int \mathbf{c} d\mu) \in [0,1] \times [0,\infty)$ belongs to the epigraph of the function V. Let $d_\theta \ge 0$ be any subdifferential of V at θ. The epigraph is contained in the half-space

$$\{(x,y) \in \mathbb{R}^2 : y - V(\theta) \ge d_\theta(x - \theta)\}.$$

In particular,

$$V(\theta) \le \int \mathbf{c} d\mu - d_\theta\Big(\int \mathbf{1}_{V^f}(a) d\mu - \theta\Big),$$

and so

$$V(\theta) \le \inf_{\mu \in \Delta} \int \Big(\mathbf{c} - d_\theta(\mathbf{1}_{V^f}(a) - \theta)\Big) d\mu.$$

On the other hand, if μ^* solves (5.2) then

$$\inf_{\mu\in\Delta}\int\Big(\mathbf{c}-d_\theta(\mathbf{1}_{V^f}(a)-\theta)\Big)d\mu \le \int\Big(\mathbf{c}-d_\theta(\mathbf{1}_{V^f}(a)-\theta)\Big)d\mu^* \le \int \mathbf{c}d\mu^* = V(\theta).$$

Consequently, $V(\theta)$ equals the (unconstrained) minimum long-run average cost for the cost function $\mathbf{c}-d_\theta(\mathbf{1}_{V^f}(a)-\theta)$. Hence, the result follows because

$$V(\theta)+u(x) = \inf_{a\in A(x)}\Big\{\mathbf{c}(x,a)-d_\theta(\mathbf{1}_{V^f}(a)-\theta)+\int_X u(y)\mathbf{Q}(dy|x,a)\Big\}, \qquad (5.3)$$

for $x\in X$, is just the dynamic programming optimality equation of this (unconstrained) average cost problem. □

Given the hybrid control problem with the ratio-average optimality criterion, we can consider an associated control problem with the usual average cost optimality criterion

$$\tilde{J}(x^f,x^s,\nu) = \limsup_{n\to\infty}\frac{1}{n}E^{\nu}_{x^f,x^s}\big[\sum_{k=0}^{n-1}\mathbf{c}(x^f_k,x^s_k,a_k)\big].$$

This control problem consists in assuming that a "system-time unit" passes for whatever action is taken, either in V^f or V^s, and then the denominator n corresponds to the number of actions taken. We will refer to this problem as the associated non-hybrid control model.

Rewrite the equation (5.3) as

$$V(\theta)-d_\theta\theta+u(x) = \inf_{a\in A(x)}\Big\{\mathbf{c}(x,a)-d_\theta\mathbf{1}_{V^f}(a)+\int_X u(y)\mathbf{Q}(dy|x,a)\Big\} \quad \text{for } x\in X.$$

Let $0\le\theta^*\le 1$ be as in Proposition 5.1(ii); that is, such that $V(\theta^*)-d_{\theta^*}\theta^*=0$. We can proceed as in Theorem 3.7 to show that $d_{\theta^*}=\lambda^*$, the optimal ratio-average cost of the hybrid control problem.

This shows that the smallest slope $d\ge 0$ for which the straight line $\theta\mapsto d\theta$ intersects (tangently) the epigraph of V is precisely the optimal ratio-average cost of the problem $d=\lambda^*$. Knowledge of the function V, and hence of the abscissa θ^* at which tangency occurs, would yield knowledge of λ^*.

We can conclude the following: *the hybrid control problem with ratio-average cost is equivalent to a non-hybrid control problem with "usual" average cost, with the restriction that the long-run expected proportion of actions taken in V^f is bounded by below by some θ^*.*

6 Comparison with impulse and switching control

A discrete-time standard control model represents a dynamic system, where at each unit of time $\{0,1,2,\ldots\}$, a control (or action) a_k is chosen and the (stochastic) transition rule (influenced by the control) is activated to transform the state x_k into x_{k+1}; for each transition a cost c_k is assigned and this process continues until a final time (possible infinite). Thus a total cost is associated with each control policy (i.e., with each sequence of controls chosen), and the objective is to optimize the total cost with respect to all possible (admissible) policies. If the final time is unbounded then some assumptions on the cost c_k are necessary (e.g., discounted or long-run average cost), to obtain a bounded total cost. The action of stopping the transitions (i.e., the evolution of the system) can be accomplished either by selecting controls with zero-cost or by setting a specific model referred to as stopping time problem.

In a hybrid model, due to the instantaneous transitions, two counters of time are convenient, one referred to as the *internal-time* (of the system) $\mathbf{t}_k$ or the "natural" or "actual" time (which counts only the standard —non instantaneous— transitions) and another one, the *evolution-time* k (or transition time, which counts all transitions, including all the instantaneous ones). The cost c_k associated with an instantaneous transition cannot be discounted and should be eventually positive (so that infinite cycles with zero-cost are discarded); this assumption gives a distinct characteristic to the hybrid model, and the stopping action needs a specific setting (which is not considered in this model, for simplicity). Modeling this hybrid control introduces a non-standard control (the action of choosing, or not, an instantaneous control), which is expressed by using a control either in V^f or in V^s.

Purely impulse or switching control problems are perhaps the simplest classes of hybrid control problems: they correspond to the case where $D^\wedge = \emptyset$ and $D^\vee = X$, i.e., the controller is allowed to choose a control either in V^f or in V^s at every state of the system; and (due to the "purely" qualification) the state-action V^f has only one element (i.e., only one possible choice of standard or regular control). Moreover, the internal-time $\mathbf{t}_k$ is used implicitly, and only the evolution-time k is necessary. Impulse (or impulsive) control and switching control problems are not different at an abstract level. However, in the simplest case of impulse control (e.g., a typical model of inventory problems; see Bensoussan [4]) the component x^s is not distinguished (or single-out), the state-action $V^s \subset X$, and the impulse control produces an instantaneous change in x^f; while in the simplest case of switching control (e.g., a typical model of power generation as in Blankenship and Menaldi [10]) the component x^s has a finite range, the state-action V^s has only a finite number of elements so that $X^s = V^s$, and the switching control produces an instantaneous change in the x^s component, but (usually) not a change in the x^f component. It is then clear that a simplified notation results in a better description of these situations.

This means that, as in Remark 2.1, the dynamic system can be also formulated in an equivalent way by means of two measurable functions (of a particular form!) $F : X \times V^f \times S \to X^f$ and $G : X \times V^s \times S \to X$, with S a Borel space, and $V^f = \{v\}$ with only one element. Thus, in the simplest case, for an impulse control model with state $x = (x^f, x^s)$ (the grouping of the components in x^f and x^s is not needed),

$$
\begin{aligned}
&\underbrace{x_{k+1} = F(x_k, v, w_k)}_{\text{standard sub-dynamic}} \qquad \text{if } \ a_k \in V^f, \\
\text{or} \qquad & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (6.4) \\
&\underbrace{x_{k+1} = a_k}_{\text{special sub-dynamic}} \qquad\qquad \text{if } \ a_k \in V^s \subset X,
\end{aligned}
$$

and for a switching control model with state $x = (x^f, x^s)$, and V^s having a finite number of elements,

$$
\begin{aligned}
&\underbrace{\left(x^f_{k+1}, x^s_{k+1}\right) = \left(F(x^f_k, x^s_k, v, w_k), x^s_k\right)}_{\text{standard sub-dynamic}} \qquad \text{if } \; a_k \in V^f, \\
\text{or} \qquad & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (6.5) \\
&\underbrace{(x^f_{k+1}, x^s_{k+1}) = (x^f_k, a_k)}_{\text{special sub-dynamic}} \qquad\qquad\qquad \text{if } \; a_k \in V^s = X^s,
\end{aligned}
$$

where $\{w_k\}$ is a sequence of i.i.d. random variables on S. Note that in both cases there is an "implicit" action/control of choosing an element in either V^f or V^s, which in this model means the decision of switching (or making an impulse) at any time.

In Stettner [29] there are general conditions under which the long-run average cost problem is solved for the impulse control in discrete-time; and in Perthame et al. [23], the specific case of switching for reflected diffusion processes is studied.

Acknowledgment

This research was funded in part by CONACyT fellowship 266252 *"Estancias sabáticas en el extranjero 2015''* , by CONACyT grant 238045 and by grant MTM2016-75497-P from the Spanish Ministerio de Economía y Competitividad.

References

[1] Abate, A., Amin, S., Prandini, M., Lygeros, J., Sastry, S.: (2006). Reachability Analysis for Controlled Discrete-Time Stochastic Hybrid Systems. In A. Hespanha, J., & Tiwari A. (Eds.), *Lecture notes in computer science,* **3927,** pp. 49–63. Springer-Verlag Berlin Heidelberg.

[2] Abate, A., Prandini, M., Lygeros, J., Sastry, S.: (2008). Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems. *Automatica*, **44**, 2724-2734.

[3] Aliprantis, C. D, Border, K.C.: *Infinite dimensional analysis.* Springer-Verlag, New York, 2006.

[4] Bensoussan, A.: *Dynamic programming and inventory control.* IOS Press, Amsterdam, 2011.

[5] Bensoussan A. and Lions, J.L.: *Applications of Variational Inequalities in Stochastic Control*, North-Holland, Amsterdam, 1982. (pages are referred to the French edition)

[6] Bensoussan, A.; Lions, J.L.: *Impulse Control and Quasi-Variational Inequalities.* Gauthier-Villars, Paris, 1984. (pages are referred to the French edition)

[7] Bensoussan, A., Menaldi J.L.: (1997) Hybrid control and dynamic programming. *Dynamics and continuous, discrete and impulsive systems* **3**, pp. 395–442.

[8] Bensoussan, A., Menaldi J.L.: (2000) Stochastic hybrid control *J. Math. Anal. Appl.* **249**, pp. 261–288.

[9] Bertsekas, D.P., Shreve, S.E.: *Stochastic Optimal Control: The Discrete-Time Case.*, Athena Scientific, Belmont, M.A, 1996.

[10] Blankenship, G.L.; Menaldi J.L.: (1984) Optimal stochastic scheduling of power generation systems with scheduling delays and large cost differentials, *SIAM J.Control Optim.*, **22**, pp. 121–132.

[11] Borkar, V.S. and Ghosh, M.K. and Sahay, P.: (1999) Optimal control of a stochastic hybrid system with discounted cost *J. Optim. Theory Appl.* **101**, pp. 557–580.

[12] Branicky, M. S.: Studies in hybrid systems: Modeling, analysis, and control, Thesis MIT, LIDS-TH-2304 (1995), Cambridge, USA.

[13] Branicky, M.S; Borkar, V.S; Mitter, S.K.: (1998) A unified framework of hybrid control: model and optimal control theory *IEEE Trans. Autom. Control* **43**, pp. 31–45.

[14] Goebel, R., Sanfelice, R.G., Teel, A.R.: *Hybrid Dynamical Systems: Modeling, Stability, and Robustness.* Princeton University Press, Princeton, N.J., 2012.

[15] Kawaguchi, K., Morimoto, H. (2007): Long-run average welfare in a pollution accumulation model, *J. Econom. Dyn. Control* **31**, pp. 703–720.

[16] Hernández-Lerma, O.: *Adaptive Markov control processes.* Springer-Verlag, New York, 1989.

[17] Hernández-Lerma, O., Lasserre, J.B.: *Discrete-time Markov control processes: Basic optimality criteria.* Springer-Verlag, New York, 1996.

[18] Hernández-Lerma, O., Lasserre, J.B.: *Further topics on discrete-time Markov control processes.* Springer-Verlag, New York, 1999.

[19] Jasso-Fuentes, H; Menaldi, J.L., Prieto-Rumeau, T.: (2017) Discrete-time hybrid control in Borel spaces. Submitted.

[20] Luque-Vásquez, F., Hernández-Lerma, O.: (1999) Semi-Markov control models with average costs, *Appl. Math. (Warsaw)*, **26**, p. 315–331.

[21] Lygeros, J.: (2005). An Overview of Hybrid Systems Control. In A. Hristu-Varsakelis, D., & Levine W. S.. (Eds.), *Handbook of Networked and Embedded Control Systems,* **IV,** pp. 519–537. Birkhäuser Boston.

[22] Menaldi J.L.: (1987) Optimal impulse control problems for degenerate diffusions with jumps, *Acta Appl. Math.* **8**, pp. 165–198.

[23] Menaldi, J.L. and Perthame, B. and Robin, M.: (1990) Ergodic problem for optimal stochastic switching *J. Math. Anal. Appl.* **147**, pp. 512–530.

[24] Puterman, M.L.: *Markov decision processes: discrete stochastic dynamic programming.* John Wiley & Sons Inc, New York, 1994.

[25] Quadrat, J.P.: (1980) Existence de solution et algorithme de resolutions numériques de problémes stochastiques degenerées ou non, *SIAM J. Control and Optim.*, **18**, 199–226.

[26] Rieder. U.: (1978) Measurable selection theorems for optimization problems *Manuscripta Math.* **24**, pp. 115–131.

[27] Robin, M.: (1976) Controle impulsionnel avec retard pour les processus de diffusion, C.R.A.S, **282** 463–466.

[28] Robin, M.: *Controle impulsionnel des processus de Markov*, Thesis INRIA, TE-035, Paris, France, 1977.

[29] Stettner, L.: (1986) Discrete Time Adaptive Impulsive Control Theory, *Stoch. Processes and their Appl*, **23**, 177–197.

[30] Stettner, L.: (1989) On some stopping and impulsive control problems with a general discount rate criterion, *Probab. Math. Statist.*, **10**, 223–245.

[31] Summers S., Lygeros, J.: (2010) Verification of discrete time stochastic hybrid systems: A stochastic reach-avoid decision problem. *Automatica*, **46**, 1951–1961.

[32] Tkachev, I. and Abate, A. (2014). Characterization and computation of infinite horizon specifications over Markov processes. *Theoretical Computer Science*, **515**, 1–18.

[33] Tkachev, I., Mereacre, A., Katoen, J.P., and Abate, A. (2017). Quantitative model-checking of controlled discrete-time Markov processes. *Information and Computation*, **253**, 1–35.

[34] Wei, Q., Guo, X.P.: (2012) New average optimality conditions for semi-Markov decision processes in Borel spaces. *Journal of Optimization Theory and Applications*, **153**, 709–732

[35] Zhang, W., Jianghai, H., Abate, A.: (2012) Infinite-Horizon Switched LQR Problems in Discrete Time: A Suboptimal Algorithm With Performance Analysis, *IEEE Transactions on Automatic Control*, **57**, 1815–1821.