Journal of Mathematical Systems, Estimation, and Control
Vol. 4, No. 1, 1994, pp. 1–19

Conditions for Average Optimality in Markov Control Processes with Unbounded Costs and Controls*

Raúl Montes–de–Oca† Onésimo Hernández–Lerma†

Abstract

This paper extends and establishes (necessity and/or sufficiency) relations between recent conditions for the existence of average cost optimal stationary policies for Markov control processes with *Borel* state space and possibly *unbounded* costs and *non-compact* control constraint sets. All of these conditions are variants of the so-called "vanishing discount factor" approach.

Key words: (discrete time) Markov control processes, average cost, discounted cost, stochastic dynamic programming, stationary optimal policies

AMS Subject Classifications: 93E20, 90C40

1 Introduction

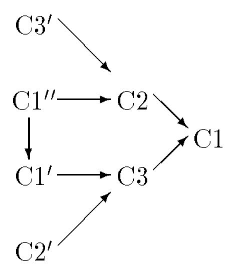
This paper deals with discrete-time Markov control processes (MCPs) with *Borel* state and control spaces, allowing *unbounded* one-stage costs and *non-compact* control constraint sets. The main objective is to extend to the latter context, and to establish—necessity and/or sufficiency—relations between recent conditions for the existence of *average cost* (AC) optimal stationary policies. All of these conditions are variants of the so-called vanishing-discount-factor approach, which even though it goes back to the 1960s—see e.g. [1,4,5,6,11,12,13] for earlier references—it has been used in the case of *unbounded* costs only in the last few years. Actually, as can be seen in the above references, most of the literature on the AC problem is

*November 17, 1992; received in final form March 6, 1993. Summary appeared in Volume 4, Number 1, 1994.

†This research was partially supported by the Consejo Nacional de Ciencia y Technologia (CONACYT), México, under grant 1132-E9206.

concentrated on MCPs for which (i) the state space is a *denumerable* set, and/or (ii) the control constraint sets are *compact*, and/or (iii) the one-stage costs are *bounded*, conditions that exclude of course important control problems, such as the "linear regulator" (or linear-quadratic systems; see Example 2.5), for which *none* of (i), (ii), (iii) holds. These conditions are *not* required in this paper; we propose instead a setting—see Assumptions 2.3 and 2.3′—that includes virtually all the AC control problems that appear in applications.

The weakest average-optimality condition we consider here—see C1 in §3—is an extension to Borel spaces of a condition used by Sennott [13] for MCPs with a *denumerable* state space and *finite* control sets. In §3 we show that C1, together with appropriate (semi-) continuity and (inf-) compactness assumptions, implies the existence of AC-optimal policies. In §4 we consider two conditions, C2 and C3 (previously used in [6,12] and [5,11] respectively), each of which is shown to be *sufficient* for C1. In our general *Borel* space context we have been unable to show that C1 implies either C2 or C3 (see Remark 6.1). However, in §5 it is shown that a suitable extension C1′ (resp. C1″) of C1 does imply C3 (resp. C2). Also in §5 we show that a strengthened version C2′ of C2 (resp. C3′ of C3) implies C3 (resp. C2). Schematically, in addition to obvious relations (such as C1″ $\Longrightarrow$ C1′ $\Longrightarrow$ C1, C2′ $\Longrightarrow$ C2, and C3′ $\Longrightarrow$ C3) we show that



This diagram summarizes our main results, which—as already mentioned—are developed in §§3, 4, 5. In §2 we introduce the Markov control processes we are concerned with.

2 Markov Control Processes

Notation. Given a Borel space X (i.e., a Borel subset of a complete and separable metric space) its Borel sigma-algebra is denoted by $\mathcal{B}(X)$. "Measurable" always means "Borel measurable". $M(X)_+$ denotes the set of real-valued nonnegative measurable functions on X, and $L(X)_+$ is the subset of lower semi-continuous (l.s.c.) functions. Recall that if X and Y are Borel spaces, then a stochastic kernel on X given Y is a function, say

$P(\cdot|\cdot)$, such that $P(\cdot|y)$ is a probability measure on X for every $y \in Y$, and $P(B|\cdot)$ is a measurable function on Y for every $B \in \mathcal{B}(X)$.

Markov control models. The basic–discrete-time, time-homogeneous–Markov control model (X, A, Q, c) consists of the state space X, the control (or action) set A, the transition law Q, and the cost-per-stage function c. Both X and A are assumed to be Borel spaces. To each state $x \in X$ we associate a nonempty measurable subset $A(x)$ of A, whose elements are the admissible control actions when the system is in state x. The set

$$\mathbb{K} := \{(x, a) |\, x \in X,\ \ a \in A(x)\}$$

of admissible state-action pairs is assumed to be a measurable subset of the Cartesian product of X and A. The transition law $Q(B|x, a)$, where $B \in \mathcal{B}(X)$ and $(x, a) \in \mathbb{K}$ is a stochastic kernel on X given $\mathbb{K}$. The cost-per-stage c is a nonnegative measurable function on $\mathbb{K}$, i.e. $c \in M(\mathbb{K})_+$.

We assume throughout that $\mathbb{K}$ contains the graph of a measurable map from X to A, which guarantees that the set, Δ, of policies (defined below) is nonempty.

Let $\mathbb{F}$ be the set of all measurable functions $f : X \to A$ such that $f(x) \in A(x)$ for all $x \in X$.

For each $t = 0, 1, \ldots$, define the space of admissible histories up to time t by $H_0 := X$ and

$$H_t := \mathbb{K}^t \times X = \mathbb{K} \times H_{t-1},\ \ t = 1, 2, \ldots$$

An element of H_t is a vector, or history, h_t of the form

$$h_t := (x_0, a_0, \ldots, x_{t-1}, a_{t-1}, x_t)$$

where $(x_n, a_n) \in \mathbb{K}$ for $n = 0, 1, \ldots, t-1$, and $x_t \in X$.

Definition 2.1 (a) A control policy is a sequence $\delta = \{\delta_t\}$ such that, for each $t = 0, 1, \ldots, \delta_t$ is a stochastic kernel on A given H_t, and which satisfies the constraint $\delta_t(A(x_t)|h_t) = 1$ for all $h_t \in H_t$. The set of all policies is denoted by Δ. (b) A control policy $\delta = \{\delta_t\}$ is said to be a *stationary policy* if there exists a function $f \in \mathbb{F}$ such that $\delta_t(\cdot|h_t)$ is concentrated at $f(x_t)$ for all $t = 0, 1, \ldots$; in this case, we identify δ with $f \in \mathbb{F}$, and refer to $\mathbb{F}$ as the set of stationary policies.

Performance criteria. We shall denote by P_x^δ the induced probability measure when using the policy δ, given the initial state $x_0 = x$ (see e.g. Hinderer [9] page 80 for the construction of P_x^δ). the corresponding expectation operator is denoted by E_x^δ.

For any policy $\delta \in \Delta$ and initial state $x \in X$, let

$$J(\delta, x) := \limsup_n (n+1)^{-1} \sum_{t=0}^{n} E_x^\delta [c(x_t, a_t)] \tag{2.1}$$

be the long-run expected average cost, and

$$V_\alpha(\delta, x) := E_x^\delta \left[\sum_{t=0}^{+\infty} \alpha^t c(x_t, a_t) \right] \tag{2.2}$$

the α-discounted expected total cost, where $\alpha \in (0,1)$ is the so-called discount factor. The functions

$$J(x) := \inf_\delta J(\delta, x), \text{ and } V_\alpha(x) := \inf_\delta V_\alpha(\delta, x), \; x \in X, \tag{2.3}$$

are the optimal average cost and the optimal α-discounted cost, respectively, when the initial state is x. A policy $\delta \in \Delta$ is said to be *average cost optimal* (or AC-optimal) if $J(x) = J(\delta, x)$ for all $x \in X$, and similarly for the α-discounted case.

To state one of our main hypotheses (Assumption 2.3(a) below) we require the following definition.

Definition 2.2. A real-valued function v on $\mathbb{K}$ is said to be *inf-compact* on $\mathbb{K}$ if the set $\{a \in A(x) | v(x,a) \leq r\}$ is compact for every $x \in X$ and $r \in \mathbb{R}$.

To guarantee the existence of "measurable minimizers", we need appropriate (semi-) continuity and compactness conditions on the components of the control model (X, A, Q, c). Here we will make the following assumptions.

Assumption 2.3

(a) $c(x,a)$ is lower semicontinuous (l.s.c.) and inf-compact on $\mathbb{K}$;

(b) The transition law Q is strongly continuous, i.e. the function $v'(x,a) := \int v(y) Q(dy|x,a)$ is a bounded and continuous function in $(x,a) \in \mathbb{K}$ for each measurable and bounded function v on X;

(c) The multifunction $x \to A(x)$ is upper semicontinuous, i.e. for any open set $G \subset A$, the set $\{x | A(x) \subset G\}$ is open in X.

Assumption 2.3' (a) Assumptions 2.3(a) and 2.3(c) hold;
(b) the transition law Q is weakly continuous, i.e. the function v' in Assumption 2.3(b) is continuous and bounded on $\mathbb{K}$ whenever v is continuous and bounded on X.

Remark 2.4 Obviously, Assumption 2.3 implies Assumption 2.3$'$. Assumption 2.3$'$(b) is easily seen to be equivalent to: v' is l.s.c. and bounded from below on $\mathbb{K}$ whenever v is l.s.c. and bounded from below on X.

The following example illustrates Assumption 2.3.

Example 2.5 Additive-noise systems with quadratic costs. Consider the stochastic system

$$x_{t+1} = F(x_t, a_t) + \varepsilon_t, \ t = 0, 1, \ldots \tag{2.4}$$

with quadratic cost

$$c(x, a) = x'qx + a'ra \qquad (\text{“prime” denotes transpose})$$

where, say, $X = \mathbb{R}^n$, $A(\cdot) = A = \mathbb{R}^m$, and $F : X \times A \to X$ is a continuous function; q and r are given symmetric, positive definite matrices. This example includes of course the linear-quadratic (or LQ) case in which F is of the form $F(x, a) = \gamma x + \beta a$, where γ and β are given matrices. We suppose that the random disturbances ε_t are i.i.d. (independent and identically distributed) random variables with a continuous density. In this example, Assumptions 2.3(a) and 2.3(c) trivially hold, and Assumption 2.3(b) is a consequence of Scheffé's Theorem (see, e.g., Hernández–Lerma [4] p. 125).

To conclude this section we state four lemmas that summarize important facts to be used in later sections. Each of these lemmas is provided with references for its proof.

Lemma 2.6 ([5] Lemma 2.3). *If Assumption 2.3 holds, and $u \in M(X)_+$, then the function $\overline{u}$ defined as*

$$\overline{u}(x) := \inf_{a \in A(x)} \left[c(x, a) + \int u(y) Q(dy|x, a) \right], \ x \in X \tag{2.5}$$

is measurable, and there exists $f \in \mathbb{F}$ such that

$$\overline{u}(x) = c(x, f(x)) + \int u(y) Q(dy|x, f(x)) \quad \forall x. \tag{2.6}$$

Lemma 2.6' ([7] Lemma 2.7(b)). *If Assumption 2.3' holds, and $u \in L(X)_+$, then the function $\overline{u}$ defined in Lemma 2.6 is l.s.c., and there exists a stationary policy $f \in \mathbb{F}$ such that (2.6) holds.*

Lemma 2.7 (cf. [1,6,11]). *For any policy* $\delta \in \Delta$,

$$\limsup_{\alpha \uparrow 1} (1-\alpha) V_\alpha(\delta, x) \le J(\delta, x), \quad x \in X.$$

Lemma 2.8 ([7] Theorem 4.2) *Suppose that Assumption 2.3 (or 2.3') holds, and let* $\alpha \in (0,1)$ *be an arbitrary but fixed discount factor. If*

$$V_\alpha(x) < +\infty \text{ for every } x \in X,$$

then V_α *is the (pointwise) minimal function in* $L(X)_+$ *that satisfies the Dynamic Programming equation, i.e.*

$$V_\alpha(x) = \min_{a \in A(x)} \left[c(x,a) + \alpha \int V_\alpha(y) Q(dy|x,a) \right], \;\; x \in X. \tag{2.7}$$

Moreover, there exists a stationary policy $f_\alpha \in \mathbb{F}$ *such that* $f_\alpha(x) \in A(x)$ *minimizes the r.h.s. (right-hand side) of (2.7) for all* $x \in X$*, i.e.*

$$V_\alpha(x) = c(x, f_\alpha(x)) + \alpha \int V_\alpha(y) Q(dy|x, f_\alpha(x)) \quad \forall\, x \in X \tag{2.8}$$

and f_α *is* α*-discount optimal.*

3 The Optimality Condition C1

In this section we give conditions that ensure the existence of AC-optimal policies.

Let $V_\alpha(\cdot)$ be the optimal α-discounted cost and let $\overline{x} \in X$ be an arbitrary, but fixed state. Define

$$h_\alpha(x) := V_\alpha(x) - V_\alpha(\overline{x}), \;\; x \in X, \;\; \alpha \in (0,1). \tag{3.1}$$

Condition 1 (C1). There exist nonnegative constants N and M, a nonnegative (not necessarily measurable) function b on X, and $\alpha_0 \in (0,1)$ such that

(a) $V_\alpha(x) < +\infty$ for every $x \in X$ and $\alpha \in (0,1)$;

(b) $(1-\alpha) V_\alpha(\overline{x}) \le M \qquad \forall \alpha \in [\alpha_0, 1)$;

(c) $-N \le h_\alpha(x) \le b(x)$, for every $x \in X$ and $\alpha \in [\alpha_0, 1)$.

It is convenient to explicitly include the requirement (a) in $C1$, as done by Sennott [13], but in fact, as noted by O. Vega–Amaya (private communication), it is redundant, i.e. (a) can be deduced from (b) and (c).

Example 3.1. For the LQ system in Example 2.5, with $n = m = 1$, suppose $\gamma \cdot \beta \neq 0$, both q and r are positive, and the disturbances ε_t have zero mean and finite variance $\sigma^2 > 0$. Under these hypotheses we have the following well-known results (see e.g. [2] or [5]):

$$V_\alpha(x) = k(\alpha)x^2 + \frac{k(\alpha)\alpha\sigma^2}{1-\alpha} \quad \forall x \in \mathbb{R},\ \alpha \in (0,1), \tag{3.2}$$

where $k(\alpha)$ is the unique positive solution to the equation

$$k = \left[1 - \alpha k\beta^2(r + \alpha k\beta^2)^{-1}\right] \alpha k\gamma^2 + q, \tag{3.3}$$

and $k(\alpha) \to k^*$ as $\alpha \uparrow 1$, where k^* is the unique positive solution of the equation obtained from (3.3) when $\alpha \uparrow 1$.

$C1$(a) evidently holds. Now, taking $\overline{x} = 0$ in (3.2), we obtain

$$(1-\alpha)V_\alpha(0) = k(\alpha)\alpha\sigma^2 \qquad \forall \alpha \in (0,1), \tag{3.4}$$

and

$$h_\alpha(x) = k(\alpha)x^2 \quad \forall x \in X,\ \alpha \in (0,1). \tag{3.5}$$

Let $\varepsilon > 0$ be arbitrary but fixed. Choose $\alpha_0 \in (0,1)$ such that

$$k(\alpha) \leq k^* + \varepsilon \qquad \forall \alpha \in [\alpha_0, 1)$$

Hence, defining $M := (k^* + \varepsilon)\sigma^2$ and $b(x) := (k^* + \varepsilon)x^2$ we conclude that for all $\alpha \in [\alpha_0, 1)$ and $x \in X$,

$$(1-\alpha)V_\alpha(0) \leq M, \text{ and } 0 \leq h_\alpha(x) \leq b(x).$$

Therefore $C1$(b) and $C1$(c) hold.

Lemma 3.2 *Under $C1$, there exists a constant $j^* \geq 0$ and a sequence of discount factors $\alpha_n \uparrow 1$ such that*

$$\lim_{n\to+\infty} (1-\alpha_n)V_{\alpha_n}(x) = j^* \quad \forall x \in X. \tag{3.6}$$

Proof: Let $\overline{x}$ be the fixed state in (3.1), and let j^* be a limit point of $(1-\alpha)V_\alpha(\overline{x})$ as $\alpha \uparrow 1$. Let $\alpha_n \uparrow 1$ be such that

$$\lim_{n\to+\infty} (1-\alpha_n)V_{\alpha_n}(\overline{x}) = j^*. \tag{3.7}$$

Now, for every $x \in X$, (3.1) together with $C1$ and (3.7) yields

$$\begin{aligned} |(1-\alpha_n)V_{\alpha_n}(x) - j^*| &\leq (1-\alpha_n)|h_{\alpha_n}(x)| \\ &\quad + |(1-\alpha_n)V_{\alpha_n}(\overline{x}) - j^*| \\ &\leq (1-\alpha_n)\max\{N, b(x)\} + |(1-\alpha_n)V_{\alpha_n}(\overline{x}) - j^*| \\ &\to 0 \text{ as } n \to +\infty \end{aligned}$$

This proves the lemma. □

We present next the main result in this section.

Theorem 3.3 *Suppose that Assumption 2.3 and C1 hold. Let j^* be the constant obtained in Lemma 3.2. Then:*

(a) There exists a measurable function h on X such that $-N \le h(\cdot) \le b(\cdot)$ and

$$j^* + h(x) \ge \min_{a \in A(x)} \left[c(x,a) + \int h(y) Q(dy|x,a) \right] \quad \forall x; \tag{3.8}$$

(b) There exists a stationary policy $f^* \in \mathbb{F}$ such that

$$j^* + h(x) \ge c(x, f^*(x)) + \int h(y) Q(dy|x, f^*(x)) \quad \forall x; \tag{3.9}$$

(c) f^* is AC-optimal, and $J(f^*, x) = j^* \quad \forall x$.

Proof: (a) Let $\{\alpha_n\}$, $\alpha_n \uparrow 1$, be as obtained in Lemma 3.2, and let h_α be as in (3.1). Define $h : X \to \mathbb{R}$ as

$$\begin{aligned} h(x) &:= \liminf_{n \to +\infty} h_{\alpha_n}(x) \\ &= \lim_{n \to +\infty} H_n(x), \ x \in X, \end{aligned} \tag{3.10}$$

where $H_n(x) := \inf_{k \ge n} h_{\alpha_k}(x) \uparrow h(x)$ as $n \to +\infty$. Let $x \in X$ be an arbitrary state. Then, by (3.1) and Lemma 2.8, for each n there exists $a_n \in A(x)$ such that

$$\begin{aligned} (1-\alpha_n) V_{\alpha_n}(\overline{x}) \ &+ \ h_{\alpha_n}(x) \\ &= \ c(x, a_n) + \alpha_n \int h_{\alpha_n}(y) Q(dy|x, a_n) \ \forall n. \end{aligned} \tag{3.11}$$

On the other hand, from (3.7) and (3.10), for any given $\varepsilon > 0$ there exists an integer $n(\varepsilon)$ and a subsequence $\{\alpha_{n_i}\}$ of $\{\alpha_n\}$ such that

$$j^* + h(x) + \varepsilon \ge (1 - \alpha_{n_i}) V_{\alpha_{n_i}}(\overline{x}) + h_{\alpha_{n_i}}(x) \ \forall n_i \ge n(\varepsilon) \tag{3.12}$$

Thus, combining (3.11) and (3.12) we get

$$j^* + h(x) + \varepsilon$$

$$\begin{aligned} &\geq \quad c(x,a_{n_i}) + \alpha_{n_i}\int h_{\alpha_{n_i}}(y)Q(dy|x,a_{n_i}) \\ &\geq \quad c(x,a_{n_i}) + \alpha_{n_i}\int H_{n_i}(y)Q(dy|x,a_{n_i}) \;\; \forall\, n_i \geq n(\varepsilon) \qquad (3.13)\end{aligned}$$

Adding $\alpha_{n_i} \cdot N$ on both sides of (3.13), and using that $\alpha_{n_i} < 1$, we have

$$j^* + (h(x) + N) + \varepsilon$$

$$\geq c(x,a_{n_i}) + \alpha_{n_i}\int (H_{n_i}(y) + N)Q(dy|x,a_{n_i}) \;\; \forall n_i \geq n(\varepsilon). \qquad (3.14)$$

Now, for $i = 1, 2, \ldots$, define $D_i(x)$ as the set of all $a \in A(x)$ for which (3.14) holds, i.e.

$$\begin{aligned} D_i(x) := \{a \in A(x) | c(x,a) \quad &+ \quad \alpha_{n_i}\int (H_{n_i}(y) + N)Q(dy|x,a) \\ &\leq \quad j^* + (h(x) + N) + \varepsilon\}\end{aligned}$$

Let $r := j^* + (h(x) + N) + \varepsilon$, and $A_r(x) := \{a \in A(x) | c(x,a) \leq r\}$. Note that, from Assumption 2.3 (a)–(b), together with $H_{n_i}(\cdot) + N \geq 0$, we deduce that $D_i(x)$ is a closed subset of $A_r(x)$, which is compact. Hence, $D_i(x)$ is compact. Moreover, $a_{n_i} \in D_i(x)$, and since $\alpha_{n_i} \cdot (H_{n_i}(\cdot) + N) \uparrow (h(\cdot) + N)$, the $D_i(x)$ form a non-increasing sequence of nonempty compact sets converging to the nonempty compact set

$$\begin{aligned} D(x) := \{a \in A(x) | c(x,a) \quad &+ \quad \int (h(y) + N)Q(dy|x,a) \\ &\leq \quad j^* + (h(x) + N) + \varepsilon\}.\end{aligned}$$

Consequently, there exists a subsequence of $\{n_i\}$—which we denote by $\{n_i\}$ again to avoid complicating the notation—and a control $a_x \in D(x)$ such that $a_{n_i} \to a_x$ as $i \to \infty$. Now, let $L > n(\varepsilon)$ be an arbitrary integer. From (3.14), we get

$$j^* + (h(x) + N) + \varepsilon$$

$$\geq c(x,a_{n_i}) + \alpha_{n_i}\int (H_L(y) + N)Q(dy|x,a_{n_i}) \; \forall n_i > L. \qquad (3.15)$$

Thus, letting $i \to +\infty$, Assumption 2.3(a)–(b) yields

$$j^* + (h(x) + N) + \varepsilon \geq c(x,a_x) + \int (H_L(y) + N)Q(dy|x,a_x)$$

which in turn, letting $L \to +\infty$, yields

$$j^* + (h(x) + N) + \varepsilon \geq c(x,a_x) + \int (h(y) + N)Q(dy|x,a_x).$$

Hence,

$$\begin{aligned} j^* + h(x) + \varepsilon &\geq c(x, a_x) + \int h(y) Q(dy|x, a_x) \\ &\geq \min_{a \in A(x)} \left[c(x, a) + \int h(y) Q(dy|x, a) \right], \end{aligned}$$

and, since ε was arbitrary, we conclude (3.8).

(b) Taking $u(\cdot) := h(\cdot) + N \geq 0$ in Lemma 2.6, we guarantee the existence of $f^* \in \mathbb{F}$ such that

$$\min_{a \in A(x)} \left[c(x, a) + \int (h(y) + N) Q(dy|x, a) \right]$$

$$= c(x, f^*(x)) + \int (h(y) + N) Q(dy|x, f^*(x)) \ \forall x.$$

Finally, in the latter equality drop the constant N and then use (3.8) to obtain (3.9).

(c) This is a well-known consequence of (3.9) and $h(\cdot) \geq -N$; see e.g. [5,6,11,12]. This completes the proof of Theorem 3.3. □

In the next two sections we present sufficient conditions for $C1$, to complete the diagram in Section 1.

4 Sufficient Conditions for $C1$

We will now introduce two conditions, $C2$ and $C3$, each of which implies $C1$, which yields the right-hand half of the diagram in the Introduction.

Condition 2 ($C2$). [6,12]. There exists a constant $N \geq 0$, a function $b \in M(X)_+$, a number $\alpha_0 \in (0, 1)$, and a stationary policy $\widehat{f} \in \mathbb{F}$ such that

(a) $V_\alpha(x) < +\infty$ for every $x \in X$ and $\alpha \in (0, 1)$;

(b) $h_\alpha(x) \geq -N$ for every $x \in X$ and $\alpha \in [\alpha_0, 1)$;

(c) $h_\alpha(x) \leq b(x)$, and $\int b(y) Q(dy|x, \widehat{f}(x)) < +\infty$ for every $x \in X$ and $\alpha \in [\alpha_0, 1)$.

Define $m_\alpha := \inf_x V_\alpha(x)$, and $g_\alpha(x) := V_\alpha(x) - m_\alpha$ for $x \in X$ and $\alpha \in (0, 1)$.

Condition 3 ($C3$). [5,11]. (a) There is a policy $\widehat{\delta}$ and an initial state $\widehat{x}$ such that $J(\widehat{\delta}, \widehat{x}) < +\infty$;
(b) There exists $\beta \in [0,1)$ such that $\sup_{\beta<\alpha<1} g_\alpha(x) < +\infty$ for every $x \in X$.

For an extension of C3 to semi-Markov models see Vega–Amaya [14].

Theorem 4.1. *Under Assumption 2.3 (or 2.3′), each of the conditions $C2$ and $C3$ implies $C1$.*

Proof: Suppose that $C2$ holds. Evidently, $C1$(a) and $C1$(c) hold. Now, from (2.7) and (3.1) we have, for any $\alpha \in (0,1)$,

$$\begin{aligned}
(1-\alpha)V_\alpha(\overline{x}) &= V_\alpha(\overline{x}) - \alpha V_\alpha(\overline{x}) \\
&\leq c(\overline{x}, \widehat{f}(\overline{x})) + \alpha \int V_\alpha(y) Q(dy|\overline{x}, \widehat{f}(\overline{x})) - \alpha V_\alpha(\overline{x}) \\
&= c(\overline{x}, \widehat{f}(\overline{x})) + \alpha \int h_\alpha(y) Q(dy|\overline{x}, \widehat{f}(\overline{x})) \\
&\leq c(\overline{x}, \widehat{f}(\overline{x})) + \int b(y) Q(dy|\overline{x}, \widehat{f}(\overline{x})) \quad [\text{by } C2(\text{c})].
\end{aligned}$$

This yields $C1$(b) with $M := c(\overline{x}, \widehat{f}(\overline{x})) + \int b(y)Q(dy|\overline{x}, \widehat{f}(\overline{x}))$. Therefore $C1$ holds.

Suppose now that $C3$ holds, and let $\overline{x} := \widehat{x}$. Using $C3$(a) and Lemma 2.7 it is easily shown that ([5,11])

$$0 \leq g^L \leq g^U \leq j < +\infty, \tag{4.1}$$

where

$$\begin{aligned}
j &:= \inf_{x\in X} \inf_{\delta\in\Delta} J(\delta, x), \\
g^L &:= \liminf_{\alpha\uparrow 1} (1-\alpha) m_\alpha, \\
g^U &:= \limsup_{\alpha\uparrow 1} (1-\alpha) m_\alpha.
\end{aligned}$$

Next, we show that $m_\alpha < +\infty \ \forall \alpha \in (0,1)$. Indeed, if there exists $\alpha_1 \in (0,1)$ such that $m_{\alpha_1} = +\infty$, then $m_\alpha = +\infty \ \forall \alpha \geq \alpha_1$, since the mapping $\alpha \to m_\alpha$, $\alpha \in (0,1)$, is nondecreasing (recall that $c \geq 0$). Hence,

$$\limsup_{\alpha\uparrow 1} (1-\alpha) m_\alpha = +\infty,$$

which contradicts (4.1). Therefore, $m_\alpha < +\infty \ \forall \alpha \in (0,1)$. In turn, the latter yields

$$V_\alpha(x) = (V_\alpha(x) - m_\alpha) + m_\alpha$$

$$\begin{aligned} &= g_\alpha(x) + m_\alpha \\ &\leq \sup_{\beta<\alpha<1} g_\alpha(x) + m_\alpha \\ &< +\infty \ [\text{by } C3(\text{b})], \end{aligned}$$

i.e. $V_\alpha(x) < +\infty$ for every $x \in X$ and $\alpha \in (\beta, 1)$. Hence, since $V_\alpha(x)$ is increasing in α, we have $V_\alpha(x) < \infty \ \forall \ \alpha \in (0,1)$, so that $C1$(a) holds.

Notice that from $C3$(a) and Lemma 2.7, we also obtain

$$\limsup_{\alpha\uparrow 1}(1-\alpha)V_\alpha(\overline{x}) \leq J(\overline{x}) < +\infty.$$

Therefore, there exist constants $M = M(\overline{x}) > 0$ and $\mu = \mu(\overline{x})$ such that

$$0 \leq (1-\alpha)V_\alpha(\overline{x}) \leq M \qquad \forall \alpha \in [\mu, 1).$$

This yields $C1$(b) with $\alpha_0 := \mu$. Finally, defining

$$N := \sup_{\beta<\alpha<1} g_\alpha(\overline{x}) \ \text{ and } \ b(x) := \sup_{\beta<\alpha<1} g_\alpha(x) \ \ x \in X,$$

and using that $h_\alpha(\cdot) \leq g_\alpha(\cdot)$, we obtain

$$-N \leq h_\alpha(x) \leq b(x) \ \ \forall x \in X \ \text{ and } \ \alpha \in (\beta, 1).$$

Hence, $C1$(c) holds for arbitrary $\alpha_0 \in (\beta, 1)$; thus $C1$(c) and (b) both hold with $\alpha_0 := \max(\mu, \beta)$. This completes the proof that $C3$ implies $C1$. □

5 Further Results

In this section we consider two special results. The first one is related to the case in which Assumption 2.3(b) in Theorem 3.3 is replaced by the weaker Assumption 2.3′(b). In the second result, we give conditions to establish additional relations between $C1$, $C2$ and $C3$, which yield the left-hand half of the diagram in the Introduction; the right-hand half was already obtained in Theorem 4.1.

First, let us consider the following theorem:

Theorem 5.1 *We suppose Assumption 2.3′ and $C1$. Let $j^* \geq 0$ and $\{\alpha_n\}$, $\alpha_n \uparrow 1$, be as obtained in Lemma 3.2. Then, the conclusions of Theorem 3.3 hold if in addition one of the following conditions (i), (ii) is satisfied.*
(i) $H_n(\cdot) := \inf_{m\geq n} h_{\alpha_m}(\cdot)$ is l.s.c.;
(ii) X is a convex subset of a normed and locally compact space, and $h_{\alpha_n}(\cdot)$ is a continuous and convex function.

Example 5.2. Both conditions (i) and (ii) in Theorem 5.1 are satisfied by the LQ system in Example 3.1. In fact, from (3.5),

$$h_{\alpha_n}(x) = k(\alpha_n)x^2,\ n = 1, 2, \ldots, x \in X,$$

so that, evidently, condition (ii) holds. Furthermore,

$$H_n(x) = (\inf_{m \geq n} k(\alpha_m))x^2,\ n = 1, 2, \ldots, x \in X,$$

which trivially is l.s.c. Hence, condition (i) is satisfied.

Remark 5.3. (a) If $h_{\alpha_n}(\cdot)$ in (ii) is convex, then it is continuous if X is an open set. Sufficient conditions for $V_\alpha(\cdot)$—hence $h_\alpha(\cdot)$—to be convex are given e.g. in [8] and references therein.

(b) Condition (ii) in Theorem 5.1 has been used by Fernández-Gaucherand, Marcus and Arapostathis [3] to obtain a solution of the average cost optimality equation, which results when equality holds in (3.8).

Proof of Theorem 5.1. In both cases (i) and (ii), it suffices to prove that there is a measurable function h on X such that j^* and h satisfy the inequality (3.8), for then the conclusion (b) in Theorem 3.3 follows from Lemma 2.6′, and conclusion (c) is a consequence of (3.9).

Case (i). Exactly as in the proof of (a) in Theorem 3.3, we can obtain the inequality (3.15), i.e.

$$j^* + (h(x) + N) + \varepsilon$$

$$\geq c(x, a_{n_i}) + \alpha_{n_i} \int (H_L(y) + N) Q(dy|x, a_{n_i})\ \forall n_i > L,$$

where $h(\cdot)$ and $H_L(\cdot)$ are as in (3.10), and $a_{n_i} \to a_x \in A(x)$ as $i \to +\infty$. Now, letting $i \to +\infty$ in the latter inequality, Assumption 2.3′ (see the Remark 2.4) and condition (i) yields

$$j^* + (h(x) + N) + \varepsilon$$

$$\geq c(x, a_x) + \int (H_L(y) + N) Q(dy|x, a_x)$$

Now we conclude (3.8) as in the proof of Theorem 3.3(a).

Case (ii). Under condition (ii), Fernández-Gaucherand, Marcus and Arapostathis [3] have shown the existence of a subsequence of $\{\alpha_n\}$, which for simplicity we denote again as α_n, with $\alpha_n \uparrow 1$, and a continuous function $h : X \to \mathbb{R}$ such that

$$h_{\alpha_n}(x) \to h(x) \qquad \forall x, \tag{5.1}$$

the convergence being uniform on compact subsets of X. On the other hand, since X is separable and locally compact, there exists a sequence of open sets G_l, $l = 1, 2, \ldots$ which have a compact closure $\overline{G}_l$ and $G_l \uparrow X$ (see, e.g. Royden (1988) page 169, problem 48). Let $\overline{x}$ be the given state in (3.1), and let $x \in X$ be an arbitrary, fixed, state. Choose G_l such that $x \in G_l$ and notice that $x \in G_L \ \forall \ L \geq l$. Let $\varepsilon > 0$ be given. By (3.6) and (5.1) on $\overline{G}_l$, there exists $T > 0$ such that

$$(1-\alpha_n)V_{\alpha_n}(\overline{x}) \leq j^* + \varepsilon \quad \forall n \geq T, \text{ and} \tag{5.2a}$$

$$h(y) - \varepsilon \leq h_{\alpha_n}(y) \leq h(y) + \varepsilon \quad \forall n \geq T, \ y \in \overline{G}_l. \tag{5.2b}$$

Since $x \in G_l$, (5.2a) and (5.2b) yield

$$j^* + h(x) + 2\varepsilon \geq (1-\alpha_n)V_{\alpha_n}(\overline{x}) + h_{\alpha_n}(x) \ \forall n \geq T. \tag{5.3}$$

But, from (2.8), for each n there exists $a_n \in A(x)$ such that

$$(1-\alpha_n)V_{\alpha_n}(\overline{x}) + h_{\alpha_n}(x)$$

$$= c(x, a_n) + \alpha_n \int h_{\alpha_n}(y) Q(dy|x, a_n) \ \forall n \geq T. \tag{5.4}$$

Hence, from (5.3) and (5.4):

$$j^* + h(x) + 2\varepsilon$$

$$\geq c(x, a_n) + \alpha_n \int h_{\alpha_n}(y) Q(dy|x, a_n) \ \forall n \geq T.$$

Adding $\alpha_n \cdot N$ on both sides of the latter inequality and using that $\alpha_n < 1$, we have

$$j^* + (h(x) + N) + 2\varepsilon$$

$$\geq c(x, a_n) + \alpha_n \int (h_{\alpha_n}(y) + N) Q(dy|x, a_n) \ \forall n \geq T. \tag{5.5}$$

Thus, from (5.2b) and using that $h(\cdot) + N \geq 0$, we get

$$j^* + (h(x) + N) + 3\varepsilon$$

$$\geq c(x, a_n) + \alpha_n \int I_{G_l}(y)(h(y) + N) Q(dy|x, a_n) \ \forall n \geq T, \tag{5.6}$$

where I_{G_l} stands for the indicator function of G_l. Now, let

$$B_n(x) := \{a \in A(x) | c(x,a) + \alpha_n \int I_{G_l}(y)(h(y)+N)Q(dy|x,a) \leq j^* + (h(x)+N) + 3\varepsilon\},$$

$n = 1, 2, \ldots$. As in the proof of (a) in Theorem 3.3, we obtain that for each n, $B_n(x)$ is nonempty and compact. Moreover, $B_n(x) \downarrow B(x)$, where

$$B(x) := \{a \in A(x) | c(x,a) + \int I_{G_l}(y)(h(y)+N)Q(dy|x,a) \leq j^* + (h(x)+N) + 3\varepsilon\},$$

and $B(x)$ is nonempty. Hence, there exists a subsequence $\{a_{n_i}\}$ of $\{a_n\}$ and $a_x \in B(x)$ such that $a_{n_i} \to a_x$ as $i \to +\infty$. Consider (5.6) with $n = n_i$. Letting $i \to +\infty$, using that $I_{G_l}(\cdot)(h(\cdot)+N)$ is l.s.c., together with Assumption 2.3$'$ we conclude

$$j^* + (h(x)+N) + 3\varepsilon \geq c(x,a_x) + \int I_{G_l}(y)(h(y)+N)Q(dy|x,a_x). \tag{5.7}$$

Now, letting successively $l \to +\infty$ and $\varepsilon \to 0$ in (5.7), we get

$$j^* + (h(x)+N) \geq c(x,a_x) + \int (h(y)+N)Q(dy|x,a_x).$$

From this inequality we may obtain (3.8) for the arbitrarily chosen $x \in G_l$ as in the proof of Theorem 3.3(a). Since $G_l \uparrow X$, this proves the case (ii). □

We now turn our attention to the left-hand half on the diagram in §1, which requires strengthened versions of $C1$, $C2$ and $C3$. First consider the following conditions (B) and (C), in which $\overline{x} \in X$ is the fixed state in (3.1).

Condition (B). There exist $T > 0$ and $\alpha_0 \in (0,1)$ such that $V_\alpha(\overline{x}) - m_\alpha \leq T$ $\forall \alpha \in (\alpha_0, 1)$. (Recall that $m_\alpha := \inf_x V_\alpha(x)$.)

Condition (C). There exist a number $\theta \in (0,1)$, a nonnegative measurable function φ on X, and a θ-discount optimal policy $f_\theta \in \mathbb{F}$ such that, for all $x \in X$, $\sup_{\theta < \alpha < 1} h_\alpha(x) \leq h_\theta(x) + \varphi(x)$, and $\int \varphi(y)Q(dy|x, f_\theta(x)) < +\infty$.

Now, consider:

Condition 1' $(C1')$: $C1$ and (B) hold.

Condition 1" $(C1'')$: $C1$ and both (B) and (C) hold.

Condition 2' $(C2')$: $C2$ and both (B) and (C) hold.

Condition 3' $(C3')$: $C3$ and (C) hold.

Example 5.4. The LQ system in Example 3.1 satisfies (B) and (C). Indeed, (B) trivially holds, since taking $\overline{x} = 0$, we have $V_\alpha(0) = m_\alpha = \frac{k(\alpha)\alpha\sigma^2}{1-\alpha}$; therefore, $V_\alpha(0) - m_\alpha = 0$. Now, to verify (C), let α_0 be as in Example 3.1, so that

$$k(\alpha) \leq k^* + \varepsilon \qquad \forall \alpha \in [\alpha_0, 1);$$

hence, $\sup_{\alpha_0 < \alpha < 1} k(\alpha) < +\infty$. Choose $\theta \in (\alpha_0, 1)$ such that

$$\sup_{\alpha_0 < \alpha < 1} k(\alpha) \leq k(\theta) + 1.$$

Therefore,

$$\sup_{\theta < \alpha < 1} k(\alpha) \leq k(\theta) + 1.$$

Multiply by x^2 on both sides of the latter inequality to obtain

$$\sup_{\theta < \alpha < 1} (k(\alpha)x^2) \leq k(\theta)x^2 + x^2;$$

i.e., from (3.5), and letting $\varphi(x) := x^2$,

$$\sup_{\theta < \alpha < 1} h_\alpha(x) \leq h_\theta(x) + \varphi(x).$$

Finally, since $E(\varepsilon_t) = 0$ and $E(\varepsilon_t^2) = \sigma^2 < +\infty$,

$$\int y^2 Q(dy|x, f_\theta) = (\gamma x + \beta f_\theta(x))^2 + \sigma^2,$$

which is finite for every $x \in X$. Hence, (C) holds.

Theorem 5.5. *Suppose Assumption 2.3. Then:*

(a) $C1'$ *implies* $C3$;

(b) $C2'$ *implies* $C3$;

(c) $C3'$ *implies* $C2$;

(d) $C1''$ *implies* $C2$.

Proof: (a) Under Assumption 2.3 and $C1'$, Theorem 3.3 holds; hence, there exists an AC-optimal policy f^*. Then, taking $\widehat{\delta} = f^*$, we obtain $C3$(a). Now, from $C1$ and (B)

$$\begin{aligned} g_\alpha(x) &:= V_\alpha(x) - m_\alpha \\ &= (V_\alpha(x) - V_\alpha(\overline{x})) + (V_\alpha(\overline{x}) - m_\alpha) \\ &\leq h_\alpha(x) + T \ \text{[by (B)]} \\ &\leq b(x) + T \end{aligned}$$

$\forall \alpha \in (\alpha_0, 1)$. This implies $C3$(b), which completes the proof of part (a).

(b) We omit this proof since it is very similar to the proof of (a).

(c) Suppose that $C3'$ holds. As in the proof of Theorem 4.1, we have that $C2$(a) and $C2$(b) hold. Now, notice that by (C) and Lemma 2.8, the function $\sup_{\theta<\alpha<1} (h_\alpha(x)+N)$ is finite and l.s.c.; in particular, it is measurable. Furthermore, by (C) again,

$$\int \sup_{\theta<\alpha<1} (h_\alpha(y) + N) Q(dy|x, f_\theta(x))$$

$$\leq \int (h_\theta(y) + N) Q(dy|x, f_\theta(x)) + \int \varphi(y) Q(dy|x, f_\theta(x)) \ \forall x. \qquad (5.8)$$

Since the r.h.s. of (5.8) is finite for every x, defining $b(x) := \sup_{\theta<\alpha<1} (h_\alpha(x)+N)$ and $\widehat{f} := f_\theta$, we obtain $C2$(c) for arbitrary $\alpha_0 \in (\theta, 1)$. This completes the proof of part (c).

(d) If $C1''$ holds, then, by (a), $C3'$ holds. Thus, by (c), $C2$ follows.

This completes the proof of the Theorem 5.5. □

6 Concluding Remarks

We have presented in this paper several conditions that ensure the existence of AC-optimal stationary policies, the basic idea being to show that there is a constant j^* and a function $h(\cdot)$ that satisfy the "optimality inequality" (3.8). On the other hand, as already noted in Remark 5.3(b), under suitable assumptions it is possible to obtain the "optimality *equation*", in which equality holds in (3.8). This is an important fact, because it allows to use the value iteration (or successive approximations) algorithm [1,2,4] to approximate the optimal AC cost. Thus it would be of interest to verify if one can obtain the optimality equation under $C1$, $C2$ or $C3$. Research in this direction is in progress.

Remark 6.1 After submitting this paper, we could prove that, under Assumption 2.3, $C1$ implies $C3$, which combined with Theorem 4.1 ($C3 \Rightarrow C1$) yields that $C1$ and $C3$ are in fact *equivalent*. This has also been noted (without proof) by Sennott [13] for MCPs with countable state space and finite control sets.

References

[1] A. Arapostathis, V. Borkar, E. Fernández–Gaucherand, M.K. Ghosh and S.I. Marcus. Discrete-time controlled Markov process with average cost criterion: A survey, *SIAM J. Control Optim.* **31** (1993), 282–344.

[2] D.P. Bertsekas. *Dynamic Programming: Deterministic and Stochastic Models.* Englewood Cliffs: Prentice–Hall, 1987.

[3] E. Fernández–Gaucherand, S.I. Marcus and A. Arapostathis. Convex stochastic control problems. SIE Working Paper #92–006, Systems and Industrial Eng. Dept., University of Arizona, February 1992.

[4] O. Hernández–Lerma. *Adaptive Markov Control Processes.* New York: Springer–Verlag, 1989.

[5] O. Hernández–Lerma. Average optimality in dynamic programming on

Borel spaces—unbounded costs and controls, *Syst. Control Lett.* **17** (1991), 237–242.

[6] O. Hernández–Lerma and J.B. Lasserre. Average cost optimal policies for Markov control processes with Borel state space and unbounded costs, *Syst. Control Lett.* **15** (1990), 349–356.

[7] O. Hernández–Lerma and M. Muñoz de Ozak. Discrete-time Markov control processes with discounted costs: optimality criteria, *Kybernetika* (Prague) **28** (1992), 191–212.

[8] O. Hernández–Lerma and W.J. Runggaldier. Monotone approximations for convex stochastic control problems, *J. Math. Syst., Estimation and Control* (to appear).

[9] K. Hinderer. *Foundations of Non-Stationary Dynamic Programming with Discrete Time Parameter.* Lecture Notes Oper. Res. **33**, New York: Springer–Verlag, 1970.

[10] H.L. Royden. Real Analysis. New York: Macmillan, 1988.

[11] M. Schäl. Average optimality in dynamic programming with general state space, *Math. Oper. Res.* **18** (1993), 163–172.

[12] L.I. Sennott. Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs, *Oper. Res.* **37** (1989), 626–633.

[13] L.I. Sennott. The average cost optimality equation and critical number policies, *Prob. Eng. Inform. Sci.* **7** (1993), 47–67.

[14] O. Vega–Amaya. Average optimality in semi-Markov control models on Borel spaces: unbounded costs and controls, *Boletín Soc. Mat. Mexicana* (to appear).

Departamento de Matemáticas, Universidad Autónoma Metropolitana, Unidad Iztapalapa, A. Postal 55534, México D.F. 09340, MEXICO.

Departamento de Matemáticas, CINVESTAV–IPN, A. Postal 14–740, México D.F. 07000, MEXICO.

Communicated by Anders Lindquist