Boletín de la Sociedad Matemática Mexicana Vol. 39, 1994

CONDITIONS FOR AVERAGE OPTIMALITY IN MARKOV CONTROL PROCESSES ON BOREL SPACES*

BY R. MONTES-DE-OCA, J.A. MINJÁREZ SOSA AND O. HERNÁNDEZ-LERMA

1. Introduction

The analysis of average cost (AC), discrete-time Markov control processes (MCPs), which can be traced back to the late 1950s, has been mainly 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 cost function is bounded. However, there are important cases —e.g., the "linear regulator (or LQ) problem", to name the most conspicuous— in which none of the conditions (*i*), (*ii*), (*iii*) is satisfied. This is precisely the class of MCPs we are concerned with in this paper, namely, MCPs with *Borel* state space, and possibly *noncompact* control sets and *unbounded* costs. Our main objective is to make a "comparative analysis" of (i.e. to give counterexamples or establish implications, whenever they exist, between) conditions that ensure the existence of AC-optimal control policies.

Such an analysis was began in [18], where two of the authors compared three conditions, called (C1), (C2) and (C3) (see §3 below), previously studied in [23], [12,22] and [7,10,20] respectively. In this paper we present new relationships between (C1)–(C3) and consider two additional conditions, (C4) (from [16]) and (C5) (from [11]; see also [8,13,14,24]), with which we cover —to the best of our knowledge— all the currently known AC-optimality conditions for MCPs on Borel spaces, with unbounded one-stage costs. Related studies appear in [1, Section 6]; see also [5] for denumerable state MCPs.

The paper is organized as follows: In Section 2 we introduce general definitions and other preliminaries on MCPs. Section 3 contains the conditions (C1)–(C5), which are then compared in Sections 4 and 5.

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 $\mathcal{B}(X)$, and "measurable", for either sets or functions, means"Borel measurable". Let X and Y be Borel spaces. Then a *stochastic kernel* $Q(dx \mid y)$ on X given Y is a function such that $Q(\cdot \mid y)$ is a measure on X for each fixed $y \in Y$, and $Q(B \mid \cdot)$ is a measurable function on Y for each fixed $B \in \mathcal{B}(X)$.

*This research was partially supported by the Consejo Nacional de Ciencia y Tecnología (CONACYT) under grant 1332–E9206.

Markov control models. [1, 3, 9, 15] The basic—discrete time, time-homogeneous— Markov control model (X, A, Q, c) consists of the state space X, the control set A, the transition law Q and the one-stage cost 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 actions when the system is in the state x. The set

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

of admissible state-action pairs is assumed to be a measurable subset of the cartesian product of X and A. The transition law $Q(B \mid x, a)$, where $B \in \mathcal{B}(X)$ and $(x, a) \in K$, is a stochastic kernel on X given K. The one-stage cost c is a nonnegative measurable function on K.

We assume throughout the following that K contains the graph of a measurable map from X to A, which guarantees that the set of policies (see Definition 2.2) is nonempty.

Control policies. For each $t = 1, 2, \ldots$, let $H_t := K \times H_{t-1}$, with $H_0 := X$, be the space of admissible histories up to time t, i.e., vectors $h_t = (x_0, a_0, \ldots, x_{t-1}, a_{t-1}, x_t)$ where $(x_n, a_n) \in K$ for every $n = 0, 1, \ldots, t-1$, and $x_t \in X$.

DEFINITION (2.1). F denotes the set of all measurable functions $f: X \to A$ such that $f(x) \in A(x)$ for all $x \in X$.

DEFINITION (2.2). (a) A *control policy* is a sequence $\delta = \{\delta_t\}$ of stochastic kernels δ_t on A given H_t, $t = 0, 1, \ldots$ satisfying the constraint

$$\delta_t(A(x_t) \mid h_t) = 1 \qquad \forall h_t \in H_t, \quad t = 0, 1, \ldots$$

The set of all policies is denoted by Δ. A control policy $\delta = \{\delta_t\}$ is said to be a

(b) *stationary policy* if there exists $f \in F$ such that $\delta_t(\cdot \mid h_t)$ is concentrated at $f(x_t)$ for all $h_t \in H_t$ and $t = 0, 1, \ldots$.

As usual, we identify F with the set of all stationary policies. Thus, we may write $F \subset \Delta$.

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 [15] 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

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

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

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

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* (hereafter abbreviated *AC-optimal*) if $J(x) = J(\delta, x)$ for all $x \in X$, and similary for the α-discounted case.

To guarantee the existence of "measurable minimizers" we need appropriate (semi-) continuity and (inf-) compactness conditions on the components of the Markov control model, such as the following.

ASSUMPTION (2.3). *(a) c is lower semicontinuous (l.s.c), and inf-compact on K, i.e. the set*

$$A_r(x) := \{a \in A(x) \mid c(x,a) \leq r\}$$

is compact for every $r \in \Re$ and $x \in X$;

(b) Q is strongly continuous, i.e. the mapping $(x,a) \mapsto \int u(y) Q(dy \mid x,a)$ is continuous on K for every measurable and bounded function u on X;

3. Optimality conditions

In this section we state the AC-optimality conditions we wish to compare.

Let $V_\alpha(\cdot)$ be the optimal α-discounted cost (see (2.3)), and let $\bar{x} \in X$ be an arbitrary, but fixed state. Define

$$h_\alpha(x) := V_\alpha(x) - V_\alpha(\bar{x}), \quad x \in X, \quad \alpha \in (0,1).$$

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) $(1-\alpha)V_\alpha(\bar{x}) \leq M \quad \forall \alpha \in [\alpha_0, 1)$;
(c) $-N \leq h_\alpha(x) \leq b(x)$ for every $x \in X$ and $\alpha \in [\alpha_0, 1)$.

The next condition is a variant of (C1).

CONDITION 2 (C2). There exist a constant N$\geq$ 0, a nonnegative and measurable function b, a number $\alpha_0 \in (0,1)$ and a stationary policy $f \in F$ 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 $\int b(y)Q(dy \mid x, f(x)) < \infty$ for every $x \in X$ and $\alpha \in [\alpha_0, 1)$.

Both (C1) and (C2) were introduced by L. C. Sennott, in [23] and [22] respectively, for countable-state MCPs with finite action sets, and were extended to the Borel space case by Montes-de-Oca and Hernández-Lerma [18], and Hernández-Lerma and Lasserre [12]. In these references, it is shown that, together with Assumption 2.3 (which trivially holds in the setting of [22,23]) each of (C1) and (C2) ensures the existence of AC-optimal stationary policies.

Let us now define

$$m_\alpha := \inf_x V_\alpha(x), \quad \text{and} \quad g_\alpha(x) := V_\alpha(x) - m_\alpha, \qquad x \in X, \alpha \in (0,1).$$

CONDITION 3 (C3). (a) There is a policy $\hat{\delta}$ and an initial state $\hat{x}$ such that $J(\hat{\delta}, \hat{x}) < \infty$;

(b) There exists $\beta \in [0,1)$ such that $\sup_{\beta<\alpha<1} g_\alpha(x) < \infty$ for every $x \in X$.

Condition (C3) was used in [10] and is a slight modification of conditions used by Schäl [20]; see also Gatarek and Stettner [7]. Assumption 2.3 and (C3) guarantee the existence of AC-optimal stationary policies [10]. In [18] it is shown that (C3) implies (C1), and it was announced without proof that the converse is also true; a proof is provided in Theorem 4.1 below. The fact that (C1) and (C3) are equivalent, was first noted (again, without proof) by Sennott [23] for MCPs with denumerable state space and finite action sets.

For each $\alpha \in (0,1)$, let δ_α be a given policy, and define

$$H_\alpha(T,x) := J(\delta_\alpha, x) - \inf_{t \geq T} (t+1)^{-1} E_x^{\delta_\alpha} \left[\sum_{k=0}^{t} c(x_k, a_k) \right] \tag{3.1}$$

where $T \geq 0$ is an integer and $x \in X$.

CONDITION 4 (C4). There exist a sequence of discount factors $\alpha_n \uparrow 1$, and policies δ_{α_n} and δ such that δ_{α_n} is α_n-discount optimal for each n, with a finite-valued discounted cost $V_{\alpha_n}(\cdot)$, and

(a) $\limsup_{n\to\infty} J(\delta_{\alpha_n}, x) \geq J(\delta, x)$ for every $x \in X$
(b) $\lim_{T\to\infty} \sup_n H_{\alpha_n}(T,x) = 0$ for every $x \in X$.

Condition (C4) is a discrete-time version of a condition used by Hordijk and Van der Duyn Schouten [16] to prove the existence of AC-optimal policies for Markov decision drift processes with Borel state and action spaces. For completeness, we include here the proof of this fact for discrete-time MCPs. We shall use the following notation: $\forall n = 0, 1, \ldots, \delta \in \Delta, x \in X$ let,

$$J_n(\delta, x) := \sum_{t=0}^{n} E_x^\delta c(x_t, a_t).$$

Note that, from (2.1),

$$J(\delta, x) = \limsup_{n\to\infty} (n+1)^{-1} J_n(\delta, x).$$

THEOREM (3.1). *Suppose that condition (C4) holds. Then there exists an AC-optimal policy.*

Proof. Let α_n, δ_{α_n} and δ be as in (C4). For each n, $x \in X$ and $T > 0$, and using the assumption that c is nonnegative, the well-known formula

$$\sum_{t=0}^{\infty} \alpha^t b_t = (1-\alpha) \sum_{t=0}^{\infty} \alpha^t (b_0 + \cdots + b_t), \quad 0 < \alpha < 1, \quad b_t \ge 0,$$

yields

$$(1-\alpha_n) V_{\alpha_n}(\delta_{\alpha_n}, x) = (1-\alpha_n)^2 \left[\sum_{t=0}^{\infty} (t+1)\alpha_n^t (t+1)^{-1} J_t(\delta_{\alpha_n}, x) \right]$$
$$\ge (1-\alpha_n)^2 \left[\sum_{t=T}^{\infty} (t+1)\alpha_n^t \left[\inf_{t\ge T} (t+1)^{-1} J_t(\delta_{\alpha_n}, x) \right] \right]$$

(by (3.1))

$$= (1-\alpha_n)^2 \sum_{t=T}^{\infty} (t+1)\alpha_n^t \left[J(\delta_{\alpha_n}, x) - H_{\alpha_n}(T, x) \right]$$
$$\ge (1-\alpha_n)^2 \sum_{t=T}^{\infty} (t+1)\alpha_n^t \left[J(\delta_{\alpha_n}, x) - \sup_m H_{\alpha_m}(T, x) \right]$$
$$= (1-\alpha_n)^2 \sum_{t=T}^{\infty} (t+1)\alpha_n^t [J'(x) - \sup_m H_{\alpha_m}(T, x)]$$
$$+ (1-\alpha_n)^2 \sum_{t=T}^{\infty} (t+1)\alpha_n^t [J(\delta_{\alpha_n}, x) - J'(x)], \tag{3.2}$$

where

$$J'(x) := \limsup_{n\to\infty} J(\delta_{\alpha_n}, x).$$

Since, for each T, $(1-\alpha)^2 \sum_{t=T}^{\infty}(t+1)\alpha^t \to 1$ as $\alpha \uparrow 1$, the $\limsup_n$ of the first term of (3.2) equals

$$J'(x) - \sup_m H_{\alpha_m}(T,x) \geq J(\delta,x) - \sup_m H_{\alpha_m}(T,x)$$

by (C4a). Now let $\{\alpha_{n(i)}\}$ be a subsequence of $\{\alpha_n\}$ such that

$$J'(x) = \lim_{i\to\infty} J(\delta_{\alpha_{n(i)}}, x).$$

Then the $\limsup_n$ of the second term of (3.2) is greater than or equal to

$$\lim_{i\to\infty}(1-\alpha_{n(i)})^2 \sum_{t=T}^{\infty}(t+1)\alpha_{n(i)}^t [J(\delta_{\alpha_{n(i)}}, x) - J'(x)] = 0.$$

Combining these facts we obtain:

$$\limsup_{n\to\infty}(1-\alpha_n)V_{\alpha_n}(\delta_{\alpha_n}, x) \geq J(\delta,x) - \sup_m H_{\alpha_m}(T,x).$$

Thus, letting $T \to \infty$, condition (C4b) yields

$$\limsup_{n\to\infty}(1-\alpha_n)V_{\alpha_n}(\delta_{\alpha_n}, x) \geq J(\delta,x) \qquad \forall x \in X. \tag{3.3}$$

On the other hand, as δ_{α_n} is α_n-discount optimal for each n, we obtain, for any policy $\pi \in \Delta$ and $x \in X$,

$$\begin{aligned}\limsup_{n\to\infty}(1-\alpha_n)V_{\alpha_n}(\delta_{\alpha_n}, x) &\leq \limsup_{n\to\infty}(1-\alpha_n)V_{\alpha_n}(\pi, x) \\ &\leq J(\pi, x),\end{aligned} \tag{3.4}$$

where the latter inequality is due to a well-known Abelian Theorem (see, e.g. [21] Theorem 2.2). Thus, combining (3.3) and (3.4), we get

$$J(\pi,x) \geq J(\delta,x) \quad \forall x \in X, \pi \in \Delta;$$

hence, δ is AC-optimal. ∎

(C1)-(C4) are all variants of the so-called "vanishing discount factor" approach, which does not include the following condition. (If (C5b) holds, it is sometimes said that c is a "moment function"; see e.g. [11, 13, 14].)

CONDITION 5 (C5). (a) There exists a policy $\hat{\delta}$ and an initial state $\hat{x}$ such that $J(\hat{\delta}, \hat{x}) < \infty$;

(b) There exist increasing sequences of compact sets $X_n \uparrow X$ and $A_n \uparrow A$ such that $K_n := X_n \times A_n$ is a subset of K and

$$\lim_{n\to\infty} \inf\{c(x,a) \mid (x,a) \notin K_n\} = \infty.$$

Condition (C5) has been studied in several contexts, including MCPs with denumerable state space, controlled diffusions and semi–Markov processes; see [11] and references therein. (C5) and assumption 2.3 ensure the existence of "stable" control policies that are average cost optimal [11].

4. Comparison of optimality conditions

We now state results that describe the majority of the relations between the conditions (C1) to (C5). In Section 5 we present another result, which includes strengthened versions of (C1), (C2) and (C3).

THEOREM (4.1). *Under the Assumption 2.3, (C1) and (C3) are equivalent; hence*

$$(C2) \Longrightarrow [(C1) \Longleftrightarrow (C3)].$$

Proof. In [18, Theorem 4.1] it has been proved that, under Assumption 2.3, each of (C2) and (C3) implies (C1). Thus, to complete the proof it only remains to show that (C1) implies (C3).

Under Assumption 2.3 and condition (C1), Montes-de-Oca and Hernández-Lerma [18] have shown the existence of an AC-optimal policy f^*, with constant optimal cost. Thus, taking $\hat{\delta} = f^*$, we get (C3a) for all initial state. On the other hand, (C1) yields, for every $x \in X$ and $\alpha \in [\alpha_0, 1)$,

$$\begin{aligned} g_\alpha(x) = V_\alpha(x) - m_\alpha &= h_\alpha(x) + g_\alpha(\bar{x}) \\ &\leq b(x) + N < \infty \qquad \text{[by (C1c)]}. \end{aligned}$$

That is, taking $\beta := \alpha_0$, we get (C3b). ■

We now give several examples illustrating that other implications do not necessarily hold between (C1)-(C5).

EXAMPLE 4.2: (C4) $\not\Rightarrow$ (C_i) for i = 1, 2, 3. Take $X = \{0,1\}$ and $A = \{1\}$. The action sets and the one-stage cost are given by $A(x) = A$ for $x = 0, 1$, and $c(0,1) = 1$, $c(1,1) = 0$. The transition law is given by $Q(\{1\} \mid 1,1) = Q(\{0\} \mid 0,1) = 1$. Notice that there is only one policy, namely, $f(x) = 1$, $x \in X$. Therefore, the α-discounted optimal function is given by

$$V_\alpha(x) = V_\alpha(f,x) = \begin{cases} \frac{1}{1-\alpha} & \text{if } x = 0 \\ 0 & \text{if } x = 1 \end{cases} \tag{4.1}$$

Now, fix a sequence $\{\alpha_n\}$ with $\alpha_n \uparrow 1$. We define $\delta_{\alpha_n} = \delta = f$, $n = 1, 2, \ldots$. Obviously, δ_{α_n} is α_n-optimal, for every $n \geq 1$ and, on the other hand, with ε_{xy} being the Kronecker symbol ($\varepsilon_{xy} = 1$ if $x = y$, $= 0$ otherwise), $J(\delta_{\alpha_n}, x) = J(\delta, x) = \varepsilon_{0x}$ $\forall n \geq 1$, which implies (C4a). Moreover,

$$\inf_{t \geq T} (t+1)^{-1} J_t(\delta_{\alpha_n}, x) = \varepsilon_{0x}$$

for any given positive integer T. Therefore, for all n,

$$H_{\alpha_n}(T, x) = 0, \qquad x = 0, 1,$$

which yields (C4b). Hence, (C4) holds. Finally, to see that (C1), (C2) and (C3) do not hold it suffices to show (by Theorem 4.1) that (C2) does not hold. The latter, however, is trivial since, with $\bar{x} = 1$, there is no function b on X that satisfies the first inequality in (C2c). ■

EXAMPLE (4.3): (C5) $\not\Rightarrow$ (C$_i$) for $i = 1, 2, 3$. Let us consider again the Example 4.2. Clearly, (C5a) holds in this case because $J(f, x) = \varepsilon_{0x}$, for $x = 0, 1$. (C5b) also holds, since taking $X_n = X$ and $A_n = A$, $n \geq 1$, we have compact sets $X_n \uparrow X$ and $A_n \uparrow A$ and, moreover, $\lim_{n\to\infty} \inf \{c(x,a) \mid (x,a) \notin K_n\} = \infty$, since $\{c(x,a) \mid (x,a) \notin K_n\}$ is an empty set and $\inf \phi := +\infty$. Therefore, (C5) holds, whereas, as already seen in Example 4.2, (C1), (C2) and (C3) do not hold. ■

EXAMPLE 4.4: (C5) $\not\Rightarrow$ (C4). Take $X = \{0, 1, \ldots.\}$, $A = A(x) = \{1\}$, and $c(x, 1) = x$ for all $x \in X$. The transition law is given by

$$Q(\{0\} \mid 0, 1) = 1$$
$$Q(\{x+1\} \mid x, 1) = 1, \qquad x = 1, 2, \ldots$$

(C5a) holds because there is only one policy, say f, and $J(f, 0) = 0$. To see that (C5b) holds it suffices to take $X_n = \{0, 1, \ldots, n\}$ and $A_n = A$, $n = 1, 2, \ldots$.

We next show that C4 does not hold. Indeed, since

$$(t+1)^{-1} E_1^f \left[\sum_{k=0}^{t} c(x_k, a_k)\right] = (t+1)^{-1} \sum_{k=0}^{t} (k+1) = \frac{t+2}{2},$$

we have $J(f, 1) = \infty$ and, on the other hand,

$$\inf_{t \geq T} (t+1)^{-1} E_1^f \left[\sum_{k=0}^{t} c(x_k, a_k)\right] = \inf_{t \geq T} \left(\frac{t+2}{2}\right) = \frac{T+2}{2}.$$

for any given positive integer T. Therefore, $H_{\alpha_n}(T, 1) = \infty$ for every sequence $\alpha_n \uparrow 1$, which implies that (C4) does not hold. ■

Conditions (C1)-(C4) hold true for denumerable state MCPs under the so-called Lyapunov Function Condition, which does not require any special structure on the cost function —see [6] and references therein. Therefore, it can be seen that Ci ($i = 1, \ldots, 4$) does not imply C5. This is also shown in the following example.

EXAMPLE (4.5): $C_i \not\Rightarrow$ C5 for i = 1, 2, 3, 4. Take $X = A = \{1, 2, \ldots\}$, $A(x) = \{1, 2, \ldots, x\}$ for all $x \in X$, and $c(x, a) = 0$ for all $(x, a) \in K$. The transition law is arbitrary. Clearly, (C1), (C2), (C3), (C4) and (C5a) hold, but (C5b) does not. To see the latter, take (e.g.) $X_n = A_n = \{1, 2, \ldots, n\}$ for n = 1, 2, ... and note that

$$\inf\{c(x, a) \mid (x, a) \notin K_n\} = 0, \qquad \forall n$$ ■

5. Further results

In this section we first introduce a condition M and then we combine M with (C1), (C2) and (C3) and show that each of these combinations implies (C4).

CONDITION M (M). There exist a sequence $\alpha_n \uparrow 1$, and α_n-discount optimal policies $\delta_{\alpha_n} =: \delta_n$ such that (with H_α as in (3.1))

$$\sup_n H_{\alpha_n}(T, x) \leq G(T) \qquad \forall T \quad \text{and} \quad x,$$

where $G(T) \to 0$ as $T \to \infty$.

Condition M can be shown to hold for some LQ problems [19], as well as in the following strengthened versions of (C1), (C2) and (C3).

CONDITION 1* (C1*). C1 and M hold.

CONDITION 2* (C2*). C2 and M hold.

CONDITION 3* (C3*). C3 and M hold.

THEOREM (5.1). *Under the Assumption 2.3, C2* (hence C1*, C3* —see Theorem (4.1)) implies C4.*

Proof. Under the Assumption 2.3 and C2, Montes-de-Oca and Hernández-Lerma [18] have shown the existence of a sequence of discount factors $\alpha_n \uparrow 1$ and stationary policies δ_{α_n} and δ such that

(a) δ_{α_n} is α_n-optimal, for every n;
(b) δ is average cost optimal.

Evidently, (b) yields

$$\limsup_{n \to \infty} J(\delta_{\alpha_n}, x) \geq J(\delta, x) \qquad \forall x \in X,$$

which proves C4(a). Finally, C4(b) follows from M. ■

REMARK 5.2. It is easy to prove that (C4b) is equivalent to the existence of the limit

$$\lim_{t} t^{-1} J_{t-1}(\delta_{\alpha_n}, x) \qquad \forall x \in X \text{ and } n \geq 1,$$

i.e.,

$$\liminf_{t\to\infty} t^{-1} J_{t-1}(\delta_{\alpha_n}, x) = \limsup_{t\to\infty} t^{-1} J_{t-1}(\delta_{\alpha_n}, x) \tag{5.1}$$

$\forall x \in X$ and $n \geq 1$. By [25], the above limit exist if and only if

$$\lim_{\beta\uparrow 1} (1-\beta) \sum_{t=0}^{\infty} \beta^t E_x^{f_{\alpha_n}}[c(x_t, a_t)],$$

exists $\forall n, x$. An obvious sufficient condition for (5.1) is that the following limit exists

$$\lim_{t\to\infty} E_x^{\delta_{\alpha_n}}[c(x_t, a_t)] \qquad \forall x, n.$$

This condition is true, for instance, when c is a bounded function and a suitable ergodicity condition holds (see e.g. [9] page 56).

REMARK 5.3 It is clear that Example 4.5 satisfies condition M since $c(x, a) = 0 \ \forall (x, a) \in K$. Hence, by Theorem 5.1, Ci^* $\not\Rightarrow$ C5 for $i = 1, 2, 3, 4$.

6. Conclusions and open problems

In the previous sections we have presented a comparison between several conditions that ensure the existence of AC-optimal policies for MCPs on Borel spaces, with unbounded costs. The conditions C1-C4 are based on the "vanishing discount factor" approach, whereas C5 imposes a special structure on the one-stage cost. There remain, however, several open problems:

1. In Theorem 4.1 we have seen that C2 implies (C1) $\Longleftrightarrow$ (C3). Is the converse true? If not, one should be able to provide a counterexample.
2. Similarly, C4 does *not* imply C1, C2, C3; see Example 4.2. Does C1 or C2 imply C4?. If not, give a counterexample.
3. Does the convex analytic approach [4] or the "ergodicity/recurrence" approach (cf. [17]) give AC-optimal policies for MCPs on *Borel* spaces with *unbounded* cost? If yes, how do the corresponding assumptions relate to C1-C5?

In relation to problem 3, the reference [14] provides some results on the vanishing–discount–factor approach and *linear* programming.

R. MONTES-DE-OCA
DEPARTAMENTO DE MATEMÁTICAS,
UNIVERSIDAD AUTÓNOMA METROPOLITANA, UNIDAD IZTAPALAPA,
A. POSTAL 55534, MÉXICO, D.F. 09340, MÉXICO.

J.A. MINJÁREZ SOSA
DEPARTAMENTO DE MATEMÁTICAS,
UNIVERSIDAD DE SONORA,
ROSALES Y BLVD. LUIS ENCINAS,
83000 HERMOSILLO, SONORA, MEXICO.

O. HERNÁNDEZ-LERMA
DEPARTAMENTO DE MATEMÁTICAS,
CINVESTAV-IPN, A. POSTAL 14-740,
MÉXICO, D.F. 07000, MEXICO.

REFERENCES

[1] A. ARAPOSTATHIS, V.S. BORKAR, E. FERNÁNDEZ-GAUCHERAND, M.K. GHOSH AND S.I. MARCUS, *Discrete-time controlled Markov processes with an average cost criterion: a survey,* SIAM J. Control Optim. **31,** (1993), 282–344.

[2] D.P. BERTSEKAS, Dynamic Programming: Deterministic and Stochastic Models, Prentice Hall, Englewood Cliffs, NJ, (1987).

[3] ________, AND S.E. SHREVE, Stochastic Optimal Control: The Discrete Time Case, Academic Press, New York, (1978).

[4] V.S. BORKAR, *A convex analytic approach to Markov decision processes, Probab. Th. Rel. Fields* **78,** (1988), 583–602.

[5] R. CAVAZOS-CADENA, *Recent results on conditions for the existence of average optimal stationary policies,* Ann. Oper. Res. **28,** (1991), 3–28.

[6] ________, *A note on the vanishing interest rate approach in average Markov decision chains with continuous and bounded costs,* preprint (1993).

[7] D. GATAREK AND L. STETTNER, *On the compactness method in general ergodic impulsive control of Markov processes,* Stochastics **31,** (1990), 15–27.

[8] R. HARTLEY, *Dynamic programming and an undiscounted, infinite horizon, convex stochastic control problem,* In: Recent Developments in Markov Decision Processes (R. Hartley, L.C. Thomas and D.J. White, Eds.), Academic Press, London, (1980), 277–300.

[9] O. HERNÁNDEZ-LERMA, Adaptive Markov Control Processes, Springer-Verlag, New York, (1989).

[10] ________, *Average optimality in dynamic programming on Borel spaces —unbounded costs and controls,* Syst. Control Lett. **17** (1991), 237–242.

[11] ________, *Existence of average optimal policies in Markov control processes with strictly unbounded costs,* Kybernetika (Prague) **29,** (1993), 1–17.

[12] ________, 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.

[13] ________, AND ________, *Linear programming and average optimality of Markov control processes on Borel spaces—unbounded costs,* SIAM J. Control Optim. **32,** (1994), 480–500.

[14] ________, AND ________, *Average optimality in Markov control processes via discounted cost problems and linear programming,* Rapport LAAS, LAAS-CNRS, Toulouse, (1993).

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

[16] A. HORDIJK AND F.A. VAN DER DUYN SCHOUTEN, *Average optimal policies in Markov decision drift processes with applications to a queuing and replacement model,* Adv. Appl. Prob. **15** (1983), 274–303.

[17] M. KURANO, *The existence of a minimun pair of state and policy for Markov decision processes under hypothesis of Doeblin,* SIAM J. Control Optim. **27**, (1989), 296–307.

[18] R. MONTES-DE-OCA AND O. HERNÁNDEZ-LERMA, *Conditions for average optimality in Markov control processes with unbounded costs and controls,* J. Math. Syst., Estimation, and Control **4**, (1994), 1–19.

[19] ———, J.A. MINJÁREZ SOSA AND O. HERNÁNDEZ-LERMA, *Conditions for average optimality in Markov control processes on Borel spaces.* Reporte Interno 145, Depto. de Matemáticas, CINVESTAV-IPN, (1993).

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

[21] L.I. SENNOTT, *A new condition for the existence of optimal stationary policies in average cost Markov decisions processes,* Oper. Res. Lett. **5**, (1986), 17-23.

[22] ———, *Average cost optimal stationary policies in infinite state Markov decisions processes with unbounded costs,* Oper. Res. **37**, (1989), 626–633.

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

[24] R.H. STOCKBRIDGE, *Time-average control of martingale problems: a linear programming formulation,* Ann. Probab. **18**, (1990), 206–217.

[25] R. SZNAJDER AND J.A. FILAR, *Some comments on a theorem of Hardy and Littlewood,* J. Optim. Theory Appl. **75**, (1992).