Nadine Hilgert (México and Montpellier) and
Onésimo Hernández-Lerma (México)

LIMITING AVERAGE COST CONTROL PROBLEMS
IN A CLASS OF DISCRETE-TIME STOCHASTIC SYSTEMS

Abstract. We consider a class of $\mathbb{R}^d$-valued stochastic control systems, with possibly unbounded costs. The systems evolve according to a discrete-time equation $x_{t+1} = G_n(x_t, a_t) + \xi_t$ ($t = 0, 1, \ldots$), for each fixed $n = 0, 1, \ldots$, where the ξ_t are i.i.d. random vectors, and the G_n are given functions converging pointwise to some function G_∞ as $n \to \infty$. Under suitable hypotheses, our main results state the existence of stationary control policies that are *expected average cost* (EAC) optimal and *sample path average cost* (SPAC) optimal for the *limiting control system* $x_{t+1} = G_\infty(x_t, a_t) + \xi_t$ ($t = 0, 1, \ldots$).

1. Introduction. We are concerned with a discrete-time $\mathbb{R}^d$-valued stochastic control process evolving according to the system equation

$$(1) \qquad x_{t+1} = G_n(x_t, a_t) + \xi_t \qquad (t \in \mathbb{N})$$

(with $\mathbb{N} := \{0, 1, \ldots\}$), for each *fixed* $n \in \mathbb{N}$, where x_t and a_t denote the state and control variables, respectively, and $\{\xi_t\}$ is a sequence of independent and identically distributed (i.i.d.) random vectors. Suppose that $\{G_n\}$ is a sequence of given functions converging pointwise to some function G_∞, that is,

$$(2) \qquad G_n(x, a) \to G_\infty(x, a) \qquad \text{for all } (x, a),$$

and consider the *limiting control system*

$$(3) \qquad x_{t+1} = G_\infty(x_t, a_t) + \xi_t \qquad (t \in \mathbb{N}).$$

In this paper we give conditions for the existence of *expected average cost* (EAC) optimal and *sample path average cost* (SPAC) optimal policies for the limiting control system (3).

The motivation to study models of the type (1) comes from our interest in time-varying control systems of the form

(4) $$x_{n+1} = G_n(x_n, a_n) + \xi_n \quad (n \in \mathbb{N}),$$

where the drift term $G_n(x, a)$ tends to "stabilize" as in (2). Such kind of systems appears, for example, in the modelling of some biotechnological processes ([2, 10]).

Our study is partly a sequel to [9], where we considered related α-*discounted cost* (α-DC) problems, for $\alpha \in (0, 1)$. In fact, our approach in the present average cost (AC) case is similar to that in [9] in the sense that we first consider the AC problems for the *time-invariant* system for each *fixed* $n \in \mathbb{N}$, and then we let $n \to \infty$ to obtain the corresponding AC results for the limiting system (3). A key fact to get the latter results is that, from (2), we obtain the convergence

(5) $$Q_n(B \,|\, x, a) \to Q_\infty(B \,|\, x, a),$$

where, for each $n \in \mathbb{N} \cup \{\infty\}$, Q_n denotes the *transition law* of (1), namely,

(6) $\quad Q_n(B \,|\, x, a) := \mathrm{Prob}(x_{t+1} \in B \,|\, x_t = x, \; a_t = a) \quad (t \in \mathbb{N}).$

(A precise statement of (5) is given in Lemma 10(iii).)

Still another connection with [9] is that the existence of EAC optimal policies is obtained via the well known "vanishing discount" approach (see, for instance, [1, 3, 4, 5], [6, Chapter 4], [7, Chapter 10]). Thus an important ingredient in the proof of Theorem 4 below is the α-DC optimality equation in [9, Theorem 3] for $n = \infty$. Finally, the existence of SPAC optimal policies is obtained as a consequence of Theorem 4 and recent results in [8].

The remainder of the paper is organized as follows. In §2 we introduce some terminology and the AC criteria we are interested in. In §3 we present the hypotheses and state our main results, Theorem 4 and Corollary 8, which are proved in §4.

2. Markov control models. For each *fixed* $n = 0, 1, \ldots, \infty$, we first introduce the Markov control model $\mathcal{M}_n$ associated with the system (1). That is,

(7) $\quad \mathcal{M}_n := (\mathbf{X}, \mathbf{A}, \{A(x) \,|\, x \in \mathbf{X}\}, Q_n, c) \quad$ for $n = 0, 1, \ldots, \infty,$

with state space $\mathbf{X} = \mathbb{R}^d$ (or a Borel subset of $\mathbb{R}^d$) and control (or action) set $\mathbf{A}$, which is assumed to be a Borel space, namely, a Borel subset of some complete separable metric space. The spaces $\mathbf{X}$ and $\mathbf{A}$ are endowed with their Borel σ-algebras $\mathcal{B}(\mathbf{X})$, $\mathcal{B}(\mathbf{A})$. For each $x \in \mathbf{X}$, the set $A(x) \in \mathcal{B}(\mathbf{A})$

in (7) stands for the (nonempty) set of admissible controls (or actions) when the state is x. Moreover, we suppose that the set

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

of admissible state-action pairs is a Borel subset of $\mathbf{X} \times \mathbf{A}$. Finally, Q_n denotes the transition law in (6), defined for all $B \in \mathcal{B}(\mathbf{X})$ and $(x, a) \in \mathbf{K}$, whereas $c : \mathbf{K} \to \mathbb{R}$ is a measurable function that represents the cost-per-stage.

Let $\varPi$ be the set of all (randomized, history-dependent) admissible control policies. If necessary, see [1, 3, 6, 7, 8] for further information on those policies. Let $\mathbf{F}$ be the set of all measurable functions $f : \mathbf{X} \to \mathbf{A}$ with values $f(x)$ in $A(x)$ for all $x \in \mathbf{X}$. As usual, for Markov control processes, we will identify $\mathbf{F}$ with the subfamily $\varPi_{\mathrm{DS}}$ of $\varPi$ that consists of the (deterministic) *stationary* policies. (The Borel measurability of $\mathbf{K}$ and Assumption 2(a) below ensure that $\mathbf{F}$—hence $\varPi$—is nonempty.)

For notational ease, let us define

$$c(x, f) := c(x, f(x)) \quad \text{and} \quad Q_n(\cdot \mid x, f) := Q_n(\cdot \mid x, f(x)),$$

for all $f \in \mathbf{F}$, $x \in \mathbf{X}$ and $n = 0, 1, \ldots, \infty$.

AC criteria. For each fixed $n = 0, 1, \ldots, \infty$, let $J_n^0(\pi, x)$ be the long-run *sample path average cost* (SPAC) using the policy $\pi \in \varPi$, for the initial state $x_0 = x$, when the control model is $\mathcal{M}_n$ in (7); see also (1). That is,

$$
(8) \qquad J_n^0(\pi, x) := \limsup_{t \to \infty} \frac{1}{t} \sum_{i=0}^{t-1} c(x_i, a_i).
$$

A policy $\pi^* \in \varPi$ is said to be *SPAC optimal for the control model $\mathcal{M}_n$* if there exists a constant $\widehat{\varrho}_n$ such that

$$
(9) \qquad
\begin{aligned}
J_n^0(\pi^*, x) &= \widehat{\varrho}_n \quad P_x^{(n)\pi^*}\text{-a.s. } \forall x \in \mathbf{X}, \quad \text{and} \\
J_n^0(\pi, x) &\geq \widehat{\varrho}_n \quad P_x^{(n)\pi}\text{-a.s. } \forall \pi \in \varPi, \ x \in \mathbf{X},
\end{aligned}
$$

where $P_x^{(n)\pi}$ denotes the probability measure corresponding to $\mathcal{M}_n$ when using the policy π, for the initial state $x_0 = x$. The constant $\widehat{\varrho}_n$ is called the *optimal sample path average cost* for the model $\mathcal{M}_n$.

The "expected" analogue of (8) is the long-run *expected average cost* (EAC):

$$
(10) \qquad J_n(\pi, x) := \limsup_{t \to \infty} \frac{1}{t} E_x^{(n)\pi} \Big(\sum_{i=0}^{t-1} c(x_i, a_i) \Big).
$$

The corresponding value function is

$$
(11) \qquad J_n^*(x) := \inf_{\pi \in \varPi} J_n(\pi, x) \quad \forall x \in \mathbf{X}.
$$

A policy $\pi^* \in \Pi$ is said to be *EAC optimal* for the control model $\mathcal{M}_n$ if it attains the infimum in (11).

In the following section we show the existence of stationary policies that are SPAC and EAC optimal for the model $\mathcal{M}_n$, for each fixed $n = 0, 1, \dots, \infty$, and the convergence of the optimal costs J_n^* and $\widehat{\varrho}_n$ as $n \to \infty$.

3. Main results. We shall require three different sets of hypotheses. The first one, Assumption 1, refers to the system (4):

ASSUMPTION 1. (a) For each $n \in \mathbb{N}$, the function $G_n : \mathbf{K} \to \mathbf{X}$ is continuous, and, furthermore, there exists a continuous function $G_\infty : \mathbf{K} \to \mathbf{X}$ for which (2) holds for all $(x, a) \in \mathbf{K}$.

(b) The $\mathbf{X}$-valued random disturbances ξ_t are i.i.d., and their common distribution is absolutely continuous (with respect to the Lebesgue measure) with a continuous, bounded and positive density γ.

The next two assumptions are more specific to the control model (7), and have already been used in previous works—see for instance [3, 8, 15] or [7, Chapter 10].

ASSUMPTION 2. For each state $x \in \mathbf{X}$:

(a) $A(x)$ is a (nonempty) compact subset of $\mathbf{A}$, and
(b) $c(x, \cdot)$ is lower semicontinuous (l.s.c.) on $A(x)$.
(c) Moreover, $c(x, a)$ is nonnegative, and there exist a constant $\overline{c}_1 \geq 0$ and a measurable function $w \geq 1$ on $\mathbf{X}$ such that, for each $x \in \mathbf{X}$ and $n \in \mathbb{N}$,

$$\sup_{a \in A(x)} c(x, a) \leq \overline{c}_1 w(x).$$

NOTATION. Let $w(\cdot)$ be as in Assumption 2, and let $\mathbf{B}_w(\mathbf{X})$ be the Banach space of measurable functions $u : \mathbf{X} \to \mathbb{R}$ with a finite w-norm $\|u\|_w$, which is defined as

(12) $$\|u\|_w := \sup_{x \in \mathbf{X}} |u(x)|/w(x).$$

Further, $\mathbf{B}_0(\mathbf{X}) \subset \mathbf{B}_w(\mathbf{X})$ denotes the subspace of bounded measurable functions on $\mathbf{X}$. We write an integral as $\mu(u) := \int u \, d\mu$, whenever it is well defined.

ASSUMPTION 3. There exist a probability measure ν and two constants $\gamma > 0$ and $\beta \in (0, 1)$ for which the following holds: for each $f \in \mathbf{F}$ and $n \in \mathbb{N}$, there is a measurable function $0 \leq l_{f,n}(\cdot) \leq 1$ such that

(a) $Q_n(B \,|\, x, f) \geq l_{f,n}(x)\nu(B) \; \forall x \in \mathbf{X}, \; B \in \mathcal{B}(\mathbf{X})$;
(b) $\nu(l_{f,n}) := \int_{\mathbf{X}} l_{f,n} \, d\nu \geq \gamma$;
(c) $\nu(w) := \int_{\mathbf{X}} w \, d\nu = \|\nu\|_w < \infty$; and

(d) for all $x \in \mathbf{X}$,

$$(13) \qquad \int_{\mathbf{X}} w(y)\, Q_n(dy \,|\, x, f) \leq \beta w(x) + l_{f,n}(x)\nu(w).$$

For a discussion of Assumption 3, as well as related ergodicity conditions and examples, see, for instance, [7, §10.2.C] or [8, Remark 6.1]. Similar ergodicity conditions have been used by other authors, in particular [11, 14] or [16, 17].

We now state our first result, which is proved in §4.

THEOREM 4. *Suppose that Assumptions* 1 *to* 3 *hold. Then:*

(a) *For each* $n = 0, 1, \ldots, \infty$, *there exist a constant* ϱ_n, *a nonnegative function* h_n *in* $\mathbf{B}_w(\mathbf{X})$, *and a stationary policy* $f_n \in \mathbf{F}$ *such that, for each state* $x \in \mathbf{X}$, *the Average Cost Optimality Inequality* (*ACOI*) *holds, i.e.,*

$$(14) \qquad \varrho_n + h_n(x) \geq \min_{a \in A(x)} \left[c(x, a) + \int_{\mathbf{X}} h_n(y)\, Q_n(dy \,|\, x, a) \right],$$

and, moreover, $f_n(x) \in A(x)$ *attains the minimum in* (14), *so that*

$$(15) \qquad \varrho_n + h_n(x) \geq c(x, f_n) + \int_{\mathbf{X}} h_n(y)\, Q_n(dy \,|\, x, f_n).$$

In addition, f_n *is EAC optimal for the model* $\mathcal{M}_n$, *with* ϱ_n *being the optimal value:*

$$(16) \qquad \varrho_n = J_n^*(x) \qquad \forall x \in \mathbf{X}.$$

(b) *The sequence* $\{\varrho_n\}$ *converges to* ϱ_∞ *as* $n \to \infty$.

REMARK 5. The ACOI (14) is not the same as in the standard EAC optimality results. Indeed, as Assumptions 1 to 3 are stated for each *finite* $n \in \mathbb{N}$, it is well known that the ACOI (14) holds for *finite* $n \in \mathbb{N}$. The interesting point of Theorem 4 is that (14) also holds for $n = \infty$.

REMARK 6. Theorem 4 is proved in §4 via the α-DC optimality (see [9]), using the "vanishing discount" approach (see [1, 3, 4, 6, 7]), so let us recall the definition of the α-DC cost, $\alpha \in (0, 1)$, for the control model $\mathcal{M}_n$, using the policy $\pi \in \Pi$ with the initial state $x_0 = x$:

$$(17) \qquad V_n^\alpha(\pi, x) := E_x^{(n)\pi}\left(\sum_{t=0}^{\infty} \alpha^t c(x_t, a_t) \right).$$

The corresponding optimal α-DC function is

$$(18) \qquad V_n^\alpha(x) := \inf_{\pi \in \Pi} V_n^\alpha(\pi, x).$$

It will be shown that, for all $x \in \mathbf{X}$ and $n = 0, 1, \ldots, \infty$,

$$(19) \qquad \varrho_n = J_n^*(x) = \limsup_{\alpha \nearrow 1}(1 - \alpha)V_n^\alpha(x).$$

Result (b) of Theorem 4 can then be viewed as a further result of [9], where the convergence of $V_n^\alpha(x)$ to $V_\infty^\alpha(x)$ was proved for all $x \in \mathbf{X}$ and each *fixed* $\alpha \in (0, 1)$.

As a consequence of Theorem 4, we may show the existence of SPAC optimal policies under the following assumption.

ASSUMPTION 7. There exists a constant $\overline{c}_2 \geq 0$ such that
$$c^2(x, a) \leq \overline{c}_2 w(x) \quad \forall (x, a) \in \mathbf{K}.$$

With this additional assumption we have:

COROLLARY 8. *Suppose that Assumptions* 1, 2, 3 *and* 7 *are satisfied. Then, for each* $n = 0, 1, \ldots, \infty$, *a stationary policy is EAC optimal for the model* $\mathcal{M}_n$ *if and only if it is SPAC optimal for* $\mathcal{M}_n$; *hence, by Theorem* 4, *there exists a policy* $f_n \in \mathbf{F}$ *that is SPAC optimal for* $\mathcal{M}_n$ *and, furthermore,* ϱ_n *is the optimal sample path average cost, that is,* $\varrho_n = \widehat{\varrho}_n$ *where* $\widehat{\varrho}_n$ *is the constant in* (9). *In particular,*
$$\widehat{\varrho}_\infty = \lim_{n \to \infty} \varrho_n = \lim_{n \to \infty} \limsup_{\alpha \nearrow 1} (1 - \alpha) V_n^\alpha(x) \quad \forall x \in \mathbf{X}.$$

4. Proofs. Before proving Theorem 4 itself we shall state some preliminary facts.

LEMMA 9. *For each fixed* n *in* $\mathbb{N}$, *the state process* $\{x_t\}$ $(t = 0, 1, \ldots)$, *following the model* $\mathcal{M}_n$ *under a stationary policy* $f \in \mathbf{F}$, *is a Markov chain that, under Assumption* 1, *is aperiodic and* λ-*irreducible, where* λ *denotes the Lebesgue measure. If Assumption* 3 *is also satisfied, then the chain is positive Harris-recurrent. We denote by* $\mu_{f,n}$ *the corresponding (unique) invariant probability measure (i.p.m.), which satisfies* $\mu_{f,n}(w) < \infty$ *and the following property: there exist two constants* $R \geq 0$ *and* $r \in (0, 1)$ *such that*

$$(20) \qquad \left| \int_{\mathbf{X}} u(y) \, Q_n^t(dy \,|\, x, f) - \mu_{f,n}(u) \right| \leq R r^t \|u\|_w w(x)$$

for all $f \in \mathbf{F}$, $x \in \mathbf{X}$, $u \in \mathbf{B}_w(\mathbf{X})$, $n \in \mathbb{N}$ *and* $t = 0, 1, \ldots$, *which means, in other words, that* $\{x_t\}$ *is* w-*geometrically ergodic.*

Proof. Choose arbitrary n in $\mathbb{N}$ and f in $\mathbf{F}$. Then, by Assumption 1(b), we may write Q_n as

$$(21) \qquad Q_n(B \,|\, x, f) = \int_B \gamma[s - G_n(x, f(x))] \, ds.$$

As the density γ is positive, $Q_n(B \,|\, x, f) > 0$ for all $x \in \mathbf{X}$ if $\lambda(B) > 0$. Moreover, it is easy to check that $\{x_t\}$ has the weak Feller property, and $\{x_t\}$ is, therefore, λ-irreducible.

The aperiodicity is a straightforward consequence of the λ-irreducibility and the fact that γ is positive.

The positive Harris-recurrence is proved, for example, in [15, Lemma 4.1], and the w-geometric ergodicity of $\{x_t\}$ is due to [3, Lemmas 3.3 and 3.4], which follows ideas of [12, 13]. Furthermore, the fact that R and r are independent of f comes directly from the latter references, while the independence of n results from Assumption 3, because γ, β and ν are supposed to be independent of n, and $\{l_{f,n}\}$ is uniformly bounded in n. ∎

LEMMA 10. (a) *If Assumption 1 holds, then for each* $n = 0, 1, \ldots, \infty$:

(i) *The transition law* Q_n *is strongly continuous, that is,* $Q_n(B \,|\, \cdot)$ *is continuous on* $\mathbf{K}$ *for every Borel set* $B \in \mathcal{B}(\mathbf{X})$. (*Equivalently,* $\int u(y) \, Q_n(dy \,|\, \cdot)$ *is continuous on* $\mathbf{K}$ *for each function* u *in* $\mathbf{B}_0(\mathbf{X})$.)

(ii) *The function* $u'(\cdot) := \int u(y) \, Q_n(dy \,|\, \cdot)$ *is l.s.c. on* $\mathbf{K}$ *for every nonnegative function* u *in* $\mathbf{B}_w(\mathbf{X})$; *in particular,* $u'(x, \cdot)$ *is l.s.c. on* $A(x)$ *for each* $x \in \mathbf{X}$.

(iii) $Q_n(\cdot \,|\, x, a)$ *converges setwise to* $Q_\infty(\cdot \,|\, x, a)$ *as* $n \to \infty$; *that is,* (5) *holds for each* $B \in \mathcal{B}(\mathbf{X})$.

(b) *As a consequence of* (ii) *and Assumption* 2(b), *for each* $n = 0, 1, \ldots, \infty$, $x \in \mathbf{X}$ *and* $u \in \mathbf{B}_w(\mathbf{X})$ *nonnegative, the function* $c(x, \cdot) + \int u(y) \, Q_n(dy|x, \cdot)$ *is l.s.c. on* $A(x)$.

Proof. Part (i) follows from (21) and the continuity of γ and G_n. The Monotone Convergence Theorem and the Dominated Convergence Theorem are respectively used to prove (ii) and (iii). See [9]. ∎

LEMMA 11. *Under Assumptions* 1 *to* 3:

(a) *Assumption* 3 *holds for* $n = \infty$; *that is, for each* f *in* $\mathbf{F}$, *there exists a measurable function* $0 \leq l_{f,\infty}(\cdot) \leq 1$ *that satisfies the inequalities* (a), (b) *and* (d) *of Assumption* 3.

(b) *The results of Lemma* 9 *hold for* $n = \infty$.

Proof. (a) We will show that $l_{f,\infty} := \limsup_{n \to \infty} l_{f,n}$ satisfies inequalities (a), (b) and (d) in Assumption 3.

Let $B \in \mathcal{B}(\mathbf{X})$, $x \in \mathbf{X}$, and $f \in \mathbf{F}$. Then, from Lemma 10(iii), taking the upper limit in Assumption 3(a) we deduce that

$$Q_\infty(B \,|\, x, f) \geq l_{f,\infty}(x) \nu(B).$$

Similarly, taking the upper limit in Assumption 3(b), by Fatou's Lemma we get

$$\gamma \leq \limsup_{n \to \infty} \int_{\mathbf{X}} l_{f,n}(x) \nu(x) \, dx \leq \int_{\mathbf{X}} l_{f,\infty}(x) \nu(x) \, dx.$$

Now, to prove (13) for $n = \infty$, let $\{u_k\}$ be a nondecreasing sequence in $\mathbf{B}_0(\mathbf{X})$ such that $u_k(x) \uparrow w(x)$ for all $x \in \mathbf{X}$. Choose arbitrary $x \in \mathbf{X}$ and $f \in \mathbf{F}$. Then, by (13), for each k, n in $\mathbb{N}$,

$$\int u_k(y) \, Q_n(dy \,|\, x, f) \leq \int w(y) \, Q_n(dy \,|\, x, f) \leq \beta w(x) + l_{f,n}(x)\nu(w).$$

Taking the lower limit as $n \to \infty$, we infer by Lemma 10(iii) and an extension of Fatou's Lemma (see [7, Lemma 8.3.7]) that

$$\begin{aligned}
\int u_k(y) \, Q_\infty(dy \,|\, x, f) &\leq \liminf_{n \to \infty} \int u_k(y) \, Q_n(dy \,|\, x, f) \\
&\leq \beta w(x) + \liminf_{n \to \infty} l_{f,n}(x)\nu(w) \\
&\leq \beta w(x) + l_{f,\infty}(x)\nu(w).
\end{aligned}$$

Thus, letting $k \to \infty$, we obtain (by monotone convergence)

$$\int w(y) \, Q_\infty(dy \,|\, x, f) \leq \beta w(x) + l_{f,\infty}(x)\nu(w).$$

(b) This follows from (a) and Assumption 1. ∎

Choose arbitrary $B \in \mathcal{B}(\mathbf{X})$, $n = 0, 1, \ldots, \infty$, and $f \in \mathbf{F}$. Let $\mu_{n,f}$ be the unique i.p.m. of $Q_n(B \,|\, x, f)$ in Lemmas 9 and 11(b), i.e.,

$$(22) \qquad \mu_{n,f}(\cdot) = \int_{\mathbf{X}} Q_n(\cdot \,|\, x, f) \, \mu_{n,f}(dx).$$

The following Lemma 12 proves the convergence of $\mu_{n,f}$, as $n \to \infty$, for all $f \in \mathbf{F}$. To state it, for each $\alpha \in (0, 1)$, let us first define the α-discounted cost $V_n^\alpha(f, x, I_B)$ when $c(x, a)$ is replaced by the indicator function $I_B(\cdot)$ of $B \in \mathcal{B}(\mathbf{X})$, i.e.,

$$(23) \qquad V_n^\alpha(f, x, I_B) := \sum_{t=0}^{\infty} \alpha^t E_x^{(n)f} I_B(x_t) = \sum_{t=0}^{\infty} \alpha^t Q_n^t(B \,|\, x, f)$$

for all $x \in \mathbf{X}$.

LEMMA 12. (a) *Under Assumption 1, for all $f \in \mathbf{F}$, $x \in \mathbf{X}$, $B \in \mathcal{B}(\mathbf{X})$ and $\alpha \in (0, 1)$,*

$$(24) \qquad V_n^\alpha(f, x, I_B) \to V_\infty^\alpha(f, x, I_B) \qquad \text{as } n \to \infty.$$

(b) *Moreover, if in addition Assumptions 2 and 3 hold, then*

$$(25) \qquad \mu_{n,f}(B) \to \mu_{\infty,f}(B) \qquad \text{as } n \to \infty.$$

Proof. (a) By Lemma 10(iii), and by the formula

$$Q_n^t(B \,|\, x, f) = \int_{\mathbf{X}} Q_n^{t-1}(B \,|\, y, f) \, Q_n(dy \,|\, x, f), \qquad t \geq 1,$$

a straightforward induction argument and the Bounded Convergence Theorem give

$$\lim_{n \to \infty} Q_n^t(B \,|\, x, f) = Q_\infty^t(B \,|\, x, f), \quad t = 0, 1, \ldots$$

Therefore, by (23) and the Bounded Convergence Theorem again, we get (24).

(b) Let $B \in \mathcal{B}(\mathbf{X})$, $n = 0, 1, \ldots, \infty$, $f \in \mathbf{F}$ and $x \in \mathbf{X}$.

From the definition (23) of $V_n^\alpha(f, x, I_B)$ and a straightforward calculation (or from formula (5.3.10) in [6] with $\mu_{n,f}(B)$ in lieu of $J(\pi, x)$), we get

$$V_n^\alpha(f, x, I_B) = \frac{\mu_{n,f}(B)}{1 - \alpha} + (1 - \alpha) \sum_{t=1}^{\infty} \alpha^{t-1} \Big(\sum_{k=0}^{t-1} Q_n^k(B \,|\, x, f) - t\mu_{n,f}(B) \Big),$$

which can be written as

$$\mu_{n,f}(B) = (1-\alpha)V_n^\alpha(f, x, I_B) - (1-\alpha)^2 \sum_{t=1}^{\infty} \alpha^{t-1} \sum_{k=0}^{t-1} (Q_n^k(B \,|\, x, f) - \mu_{n,f}(B)).$$

To estimate $|\mu_{n,f}(B) - \mu_{\infty,f}(B)|$, we write $\mu_{n,f}(B) - \mu_{\infty,f}(B) = I_{\alpha,n} + II_{\alpha,n}$, where

$$I_{\alpha,n} = (1 - \alpha)[(V_n^\alpha(f, x, I_B) - V_\infty^\alpha(f, x, I_B)],$$

and

$$II_{\alpha,n} = -(1 - \alpha)^2 \sum_{t=1}^{\infty} \alpha^{t-1} \sum_{k=0}^{t-1} \{(Q_n^k(B \,|\, x, f) - \mu_{n,f}(B))$$
$$- (Q_\infty^k(B \,|\, x, f) - \mu_{\infty,f}(B))\}.$$

From part (a), we see that for each $\alpha \in (0, 1)$, $I_{\alpha,n} \to 0$ as $n \to \infty$.

Moreover, taking $u(\cdot) = I_B(\cdot)$ in (20) gives

$$|Q_n^k(B \,|\, x, f) - \mu_{f,n}(B)| \le Rr^k w(x) \quad \forall k = 0, 1, \ldots$$

We then deduce that

$$|II_{\alpha,n}| \le (1 - \alpha)^2 \sum_{t=1}^{\infty} \alpha^{t-1} \sum_{k=0}^{t-1} 2Rr^k w(x) \le (1 - \alpha)\frac{2R}{1 - r}w(x).$$

As α is arbitrary, letting $\alpha \to 1$, we obtain the convergence of $|II_{\alpha,n}|$ to 0, for all n, which finally gives (25). ∎

LEMMA 13. *Under Assumptions 1 to 3, for all $f \in \mathbf{F}$, $x \in \mathbf{X}$ and $n = 0, 1, \ldots, \infty$,*

$$(26) \qquad J_n(f) := \int_{\mathbf{X}} c(y, f) \, \mu_{f,n}(dy) = \lim_{N \to \infty} \frac{1}{N} E_x^{(n)f} \Big(\sum_{i=0}^{N-1} c(x_i, f) \Big).$$

Proof. By Assumption 2(c), $c(\cdot, f)$ is in $\mathbf{B}_w(\mathbf{X})$ for all $f \in \mathbf{F}$. Then (26) follows from (20) with $u(\cdot) := c(\cdot, f)$. ∎

LEMMA 14. *Under Assumptions* 1 *to* 3, *for all* $u \in \mathbf{B}_w(\mathbf{X})$, $x \in \mathbf{X}$, $\pi \in \Pi$ *and* $n = 0, 1, \ldots, \infty$,

$$(27) \qquad \lim_{N \to \infty} \frac{1}{N} \sup_{\pi} E_x^{(n)\pi} |u(x_N)| = 0.$$

Proof. Choose arbitrary $u \in \mathbf{B}_w(\mathbf{X})$, $x \in \mathbf{X}$, $\pi \in \Pi$, $t \in \mathbb{N}$, and $n = 0, 1, \ldots, \infty$. Observe that (13) in Assumption 3 yields

$$(28) \qquad \int w(y)\, Q_n(dy \,|\, x, a) \leq \beta w(x) + \overline{\nu} \quad \forall x \in \mathbf{X}, \ a \in A(x),$$

where $\overline{\nu} := \|\nu\|_w$, and so

$$E_x^{(n)\pi} w(x_t) \leq \beta E_x^{(n)\pi} w(x_{t-1}) + \overline{\nu}.$$

Iteration of this inequality gives

$$(29) \qquad E_x^{(n)\pi} w(x_t) \leq \beta^t w(x) + \overline{\nu} \sum_{i=0}^{t-1} \beta^i \leq \left(1 + \frac{\overline{\nu}}{1 - \beta}\right) w(x)$$

as $w(\cdot) \geq 1$. On the other hand, from (12) we have

$$E_x^{(n)\pi} |u(x_t)| \leq \|u\|_w E_x^{(n)\pi} w(x_t),$$

which together with (29) implies (27). ∎

We are now ready for the proof of Theorem 4.

Proof of Theorem 4. (a) With Lemmas 9, 11 and 14 and the results in [9] at hand, the proof of part (a), following the "vanishing discount" approach, is more or less standard—see [1, 3, 4, 5] and [7, Theorem 10.3.1]. Hence we shall only give the main ideas.

Let $\alpha \in (0, 1)$, $x \in \mathbf{X}$, and $n = 0, 1, \ldots, \infty$, and define

$$(30) \qquad M_n(\alpha) := \inf_{x \in \mathbf{X}} V_n^\alpha(x),$$

$$(31) \qquad \varrho_n(\alpha) := (1 - \alpha) M_n(\alpha),$$

$$(32) \qquad u_n^\alpha(x) := V_n^\alpha(x) - M_n(\alpha),$$

where $V_n^\alpha(x)$ is the optimal α-DC function (18) for the model $\mathcal{M}_n$. Observe that $u_n^\alpha(\cdot)$ is a *nonnegative* function.

From (29) and Assumption 2(c), it is easy to verify that

$$0 \leq V_n^\alpha(x) \leq \frac{\overline{c}_1}{1 - \alpha} \left(1 + \frac{\overline{\nu}}{1 - \beta}\right) w(x),$$

which implies

$$0 \leq \varrho_n(\alpha) \leq \overline{c}_1 \left(1 + \frac{\overline{\nu}}{1 - \beta}\right) \inf_{\mathbf{X}} w(x).$$

Therefore, there is a number ϱ_n such that $\limsup_{\alpha \nearrow 1} \varrho_n(\alpha) = \varrho_n$. For each n, let $\{\alpha_{n,k}\} \nearrow 1$ be a sequence of "discount factors" such that $\varrho_n =$

$\lim_{k\to\infty} \varrho_n(\alpha_{n,k})$, and define

(33)
$$h_n(x) := \liminf_{k\to\infty} u_n^{\alpha_{n,k}}(x) \quad \forall x \in \mathbf{X}.$$

Note that $h_n(\cdot)$ is *nonnegative*, by (32).

From the α-DC Optimality Equation obtained in [9], namely,

$$V_n^\alpha(x) = \min_{a\in A(x)} \Big[c(x,a) + \alpha \int_{\mathbf{X}} V_n^\alpha(y)\, Q_n(dy\,|\,x,a) \Big],$$

and (30)–(32), we deduce

(34)
$$\varrho_n(\alpha) + u_n^\alpha(x) = \min_{a\in A(x)} \Big[c(x,a) + \alpha \int_{\mathbf{X}} u_n^\alpha(y)\, Q_n(dy\,|\,x,a) \Big].$$

Now in (34) replace α with $\alpha_{n,k}$ and take the lower limit as $k \to \infty$. Then, by Fatou's Lemma, we get the ACOI (14) for all $n = 0, 1, \dots, \infty$.

The existence of $f_n \in \mathbf{F}$ that satisfies (15) is assured by Lemma 10(b) and Assumption 2(a), if we use a well known measurable selector theorem (see [6, Proposition D.5], for example). Moreover, by Lemma 14 and the fact that the cost $c(\cdot,\cdot)$ is nonnegative, [7, Theorem 10.3.1] gives

(35)
$$\varrho_n = \inf_{f\in\mathbf{F}} J_n(f) = J_n(f_n) = J_n^*(x) \quad \forall x \in \mathbf{X},$$

with $J_n(f)$ as in (26).

Finally, as mentioned in Remark 6, the fact that

$$\varrho_n = \limsup_{\alpha\nearrow 1}(1-\alpha)V_n^\alpha(x) \quad \text{for } all \ x \in \mathbf{X}$$

can be obtained by noting that

$$|(1-\alpha)V_n^\alpha(x) - \varrho_n| \le (1-\alpha)u_n^\alpha(x) + |\varrho_n(\alpha) - \varrho_n| \quad \forall x \in \mathbf{X}.$$

(b) Define $\underline{\varrho} := \liminf_{n\to\infty} \varrho_n$ and $\bar{\varrho} := \limsup_{n\to\infty} \varrho_n$. The basic idea of the proof is to show that

(36)
$$\underline{\varrho} = \bar{\varrho} = \varrho_\infty.$$

For each $n \in \mathbb{N}$, the first equality in (35) yields

(37)
$$\varrho_n \le J_n(f_\infty),$$

where f_∞ is an EAC optimal policy for the model $\mathcal{M}_\infty$. Moreover, from Lemma 13 we get $J_n(f_\infty) := \int_{\mathbf{X}} c(y, f_\infty)\, \mu_{n,f_\infty}(dy)$. Thus, by (25) and Fatou's Lemma, we deduce from (37) that

(38)
$$\bar{\varrho} \le J_\infty(f_\infty) = \varrho_\infty.$$

We now wish to prove

(39)
$$\underline{\varrho} \ge \varrho_\infty.$$

To get this, let $\{n_k\}$ be a subsequence of $\{n\}$ such that

$$\underline{\varrho} = \lim_{k\to\infty} \varrho_{n_k},$$

and in the ACOI (14) replace n with n_k. Thus

$$(40) \qquad \varrho_{n_k} + h_{n_k}(x) \geq \min_{a \in A(x)} \left[c(x,a) + \int_{\mathbf{X}} h_{n_k}(y)\, Q_{n_k}(dy\,|\,x,a) \right] \quad \forall x \in \mathbf{X}.$$

Now take the lower limit in (40) as $k \to \infty$ to get

$$(41) \qquad \underline{\varrho} + \underline{h}(x) \geq \liminf_{k \to \infty} \min_{a \in A(x)} \left[c(x,a) + \int_{\mathbf{X}} h_{n_k}(y)\, Q_{n_k}(dy\,|\,x,a) \right] \quad \forall x \in \mathbf{X},$$

where $\underline{h}(x) := \liminf_{k \to \infty} h_{n_k}(x)$. Then, by Fatou's Lemma, and by applying a general result on the interchange of limits and minima (see [6, Lemma 4.2.4]), we obtain

$$(42) \qquad \underline{\varrho} + \underline{h}(x) \geq \min_{a \in A(x)} \left[c(x,a) + \int_{\mathbf{X}} \underline{h}(y)\, Q_\infty(dy\,|\,x,a) \right] \quad \forall x \in \mathbf{X}.$$

Moreover, as in part (a), Lemma 10(b) and Assumption 2(a) yield the existence of a stationary policy $f \in \mathbf{F}$ that attains the minimum in (42); that is, (42) becomes

$$(43) \qquad \underline{\varrho} + \underline{h}(x) \geq c(x,f) + \int_{\mathbf{X}} \underline{h}(y)\, Q_\infty(dy\,|\,x,f) \quad \forall x \in \mathbf{X}.$$

On the other hand, iteration of the latter inequality yields, for all $N = 1, 2, \ldots,$

$$N\underline{\varrho} + \underline{h}(x) \geq \sum_{t=0}^{N-1} E_x^{(\infty)f} c(x_t, f) + E_x^{(\infty)f} \underline{h}(x_N) \geq \sum_{t=0}^{N-1} E_x^{(\infty)f} c(x_t, f),$$

as $\underline{h}(\cdot)$ is nonnegative. Hence dividing by N and letting $N \to \infty$, from (26) we obtain $\underline{\varrho} \geq J_\infty(f)$. Finally, by (35), $J_\infty(f) \geq \varrho_\infty$ and (39) follows, which together with (38) gives (36). ∎

Proof of Corollary 8. Let $n = 0, 1, \ldots, \infty$. From Lemmas 9 and 11(b), the state process $\{x_t\}$, following the model $\mathcal{M}_n$ under a stationary policy $f \in \mathbf{F}$, is an aperiodic and λ-irreducible Markov chain. The equivalence between EAC optimality and SPAC optimality for $\mathcal{M}_n$ is then a consequence of [8, Theorem 3.7(b)] applied to each model $\mathcal{M}_n$. ∎

References

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

[2] G. Bastin and D. Dochain, *On-line Estimation and Adaptive Control of Bioreactors*, Elsevier, Amsterdam, 1990.

[3]   E. Gordienko and O. Hernández-Lerma, *Average cost Markov control policies with weighted norms*: *existence of canonical policies*, Appl. Math. (Warsaw) 23 (1995), 199–218.

[4]   O. Hernández-Lerma, *Average optimality in dynamic programming on Borel spaces —Unbounded costs and controls*, Systems Control Lett. 17 (1991), 237–242.

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

[6]   —, —, *Discrete-Time Markov Control Processes*: *Basic Optimality Criteria*, Springer, New York, 1996.

[7]   —, —, *Further Topics on Discrete-Time Markov Control Processes*, Springer, New York, 1999.

[8]   O. Hernández-Lerma, O. Vega-Amaya and G. Carrasco, *Sample-path optimality and variance-minimization of average cost Markov control processes*, SIAM J. Control Optim. 38 (1999), 79–93.

[9]   N. Hilgert and O. Hernández-Lerma, *Limiting optimal discounted-cost control of a class of time-varying stochastic systems*, Systems Control Lett. 40 (2000), 37–42.

[10]   N. Hilgert, R. Senoussi and J.-P. Vila, *Nonparametric estimation of time-varying autoregressive nonlinear processes*, C. R. Acad. Sci. Paris Sér. I Math. 323 (1996), 1085–1090.

[11]   A. Hordijk and A. A. Yushkevich, *Blackwell optimality in the class of all policies in Markov decision chains with Borel state space and unbounded rewards*, Math. Methods Oper. Res. 50 (1999), 421–448.

[12]   N. V. Kartashov, *Inequalities in theorems of ergodicity and stability for Markov chains with common phase space*, *II*, Theory Probab. Appl. 30 (1986), 507–515.

[13]   —, *Strong Stable Markov Chains*, VSP, Utrecht, 1996.

[14]   H. U. Küenle, *Markov games with average cost criterion under a geometric drift condition*, paper presented at the 10th INFORMS Applied Probability Conference, University of Ulm, July 26–28, 1999.

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

[16]   A. S. Nowak, *Optimal strategies in a class of zero-sum ergodic stochastic games*, Math. Methods Oper. Res. 50 (1999), 399–419.

[17]   —, *Sensitive equilibria for ergodic stochastic games with countable state-spaces*, ibid., 65–76.

Departamento de Matemáticas   Departamento de Matemáticas
CINVESTAV-IPN   CINVESTAV-IPN
Apartado Postal 14-740   Apartado Postal 14-740
México D.F. 07000, México   México D.F. 07000, México
                                                                   E-mail: ohernand@math.cinvestav.mx

Permanent address:
Laboratoire de Biométrie
INRA-ENSA.M
2 place Viala
34060 Montpellier Cedex 1, France
E-mail: hilgert@ensam.inra.fr