Average-Cost Markov Decision Processes with Weakly Continuous Transition Probabilities

Eugene A. Feinberg [1], Pavlo O. Kasyanov[2], and Nina V. Zadoianchuk[3]

Abstract

This paper presents sufficient conditions for the existence of stationary optimal policies for average-cost Markov Decision Processes with Borel state and action sets and with weakly continuous transition probabilities. The one-step cost functions may be unbounded, and action sets may be noncompact. The main contributions of this paper are: (i) general sufficient conditions for the existence of stationary discount-optimal and average-cost optimal policies and descriptions of properties of value functions and sets of optimal actions, (ii) a sufficient condition for the average-cost optimality of a stationary policy in the form of optimality inequalities, and (iii) approximations of average-cost optimal actions by discount-optimal actions.

1 Introduction

This paper provides sufficient conditions for the existence of stationary optimal policies for average-cost Markov Decision Processes (MDPs) with Borel state and action sets and with weakly continuous transition probabilities. The cost functions may be unbounded and action sets may be noncompact. The main contributions of this paper are: (i) general sufficient conditions for the existence of stationary discount-optimal and average-cost optimal policies and descriptions of properties of value functions and sets of optimal actions (Theorems 3.1, 5.2, and 5.6), (ii) a new sufficient condition of average-cost optimality based on optimality inequalities (Theorem 4.1), and (iii) approximations of average-cost optimal actions by discount-optimal actions (Theorem 6.1).

For infinite-horizon MDPs there are two major criteria: average costs per unit time and expected total discounted costs. The former is typically more difficult to analyze. The so-called vanishing discount factor approach is often used to approximate average costs per unit time by normalized

[1]Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11794-3600, USA, eugene.feinberg@sunysb.edu

[2]Institute for Applied System Analysis, National Technical University of Ukraine “Kyiv Polytechnic Institute”, Peremogy ave., 37, build, 35, 03056, Kyiv, Ukraine, kasyanov@i.ua.

[3]Institute for Applied System Analysis, National Technical University of Ukraine “Kyiv Polytechnic Institute”, Peremogy ave., 37, build, 35, 03056, Kyiv, Ukraine, ninellll@i.ua.

expected total discounted costs. The literature on average-cost MDPs is vast. Most of the earlier results are surveyed in Arapostathis et al. [1]. Here we mention just a few references.

For finite state and action sets, Derman [10] proved the existence of stationary average-cost optimal policies. This result follows from Blackwell [6] and it also was independently proved by Viskov and Shiryaev [29]. When either the state set or the action set is infinite, even ϵ-optimal policies may not exist for some $\epsilon > 0$; Ross [23], Dynkin and Yushkevich [11, Chapter 7], Feinberg [12, Section 5]. For a finite state set and compact action sets, optimal policies may not exist; Bather [2], Chitashvili [9], Dynkin and Yushkevich [11, Chapter 7].

For MDP with finite state and action sets, there exist stationary policies satisfying optimality equations (see Dynkin and Yushkevich [11, Chapter 7], where these equations are called canonical), and, furthermore, any stationary policy satisfying optimality equations is optimal. The latter is also true for MDPs with Borel state and an action sets, if the value and weight (also called bias) functions are bounded; Dynkin and Yushkevich [11, Chapter 7]. When the optimal value of average costs per unit time does not depend on the initial state (the optimal value function is constant), the pair of optimality equations becomes a single equation. For bounded one-step costs, Taylor [28], Ross [21] for a countable state space and Ross [22], Gubenko and Statland [15] for a Borel state space provided sufficient conditions for the validity of optimality equations with a bounded bias function; see also Dynkin and Yushkevich [11, Chapter 7]. Under all known sufficient conditions for the existence of average-cost optimal policies for infinite-state MDPs, the value function is constant.

In many applications of infinite-state MDPs, one-step costs are unbounded from above. For example, holding costs may be unbounded in queueing and inventory systems. Sennott [25, 26] (and references therein) developed a theory for countable-state problems with unbounded one-step costs. For unbounded costs, optimality inequalities are used instead of optimality equations to construct a stationary average-cost optimal policy. Cavazos-Cadena [7] provided an example, when optimality inequalities hold while optimality equations do not.

Schäl [24] developed a theory for Borel state spaces and compact action sets. Two types of continuity assumptions for transition probabilities are considered in Schäl [24]: the setwise and weak continuity. For a countable state space these assumptions coincide; see Chen and Feinberg [8, Appendix]. Setwise convergence of probability measures is stronger than weak convergence; Hernández-Lerma and Lasserre [17, p. 186]. Formally speaking, the setwise continuity assumption for MDPs is not stronger than the weak continuity assumption, since the former claims that the transition probabilities are continuous in actions, while they are jointly continuous in states and actions in the latter. However, the joint continuity of transition probabilities in states and actions often holds in applications. For example, for inventory control problems with uncountable state spaces, setwise continuity of transition probabilities takes place if demand is a continuous random variable, while weak continuity holds for arbitrarily distributed demand; see Feinberg and Lewis [14, Section 4]. The importance of weak convergence for practical applications is mentioned in Hernández-Lerma and Lasserre [18, p. 141].

In many applications action sets are not compact. Hernández-Lerma [16] extended Schäl's [24] results under the setwise continuity assumptions to possibly noncompact action sets. Schäl's [24] assumptions on compactness of action sets and lower semi-continuity of cost functions in the action argument are replaced in Hernández-Lerma [16] by a more general assumption, namely, that the cost functions are inf-compact in the action argument. For weakly continuous transition probabilities and possibly noncompact action sets, Feinberg and Lewis [14] proved the existence of stationary optimal policies for MDPs with cost functions being inf-compact in both state and action arguments when, in addition to Schäl's [24] boundness assumption on the relative discounted value at each state, the so-called local boundness condition was assumed.

The original goal of this study was to show that the results from Feinberg and Lewis [14] hold without local boundness condition. However, the results of this paper are more general. This paper provides a weaker boundness condition on the relative discounted value (Assumption ($\underline{\mathbf{B}}$) in Section 5) than Assumption ($\mathbf{B}$) introduced in Schäl [24]. It also provides a more general and natural assumption (Assumption ($\mathbf{W}^*$) in Section 3) than inf-compactness of the one-step cost function in both arguments. The main result of this paper, Theorem 5.2, establishes the validity of optimality inequalities and the existence of stationary optimal policies under Assumptions ($\mathbf{W}^*$) and ($\underline{\mathbf{B}}$).

While inf-compactness of the cost function in the action parameter is a natural assumption, inf-compactness in the state argument is a more restrictive condition. For example, when the state space is unbounded (e.g., the set of nonnegative numbers) and action sets are compact, the assumption, that the cost function is inf-compact in both arguments, does not cover the case of bounded costs functions studied by Ross [22], Gubenko and Shtatland [15], and Dynkin and Yushkevich [11, Chapter 7]. Assumption ($\mathbf{W}^*$) covers this case as well as unbounded costs and noncompact action sets.

As follows from the example presented in Luque-Vásquez and Hernández-Lerma (1995), MDPs with lower-semicontinuous cost functions may possess pathological properties, even if the one-step cost function is inf-compact in the action variable. Assumption ($\mathbf{W}^*$)(ii) removes this difficulty. As stated in Lemma 3.2, this assumption is weaker than Schäl's [24] compactness and continuity assumptions for weakly continuous transition probabilities and than inf-compactness of one-step cost functions in both arguments (state and action) assumed in Feinberg and Lewis [14].

2 Model Description

For a metric space S, let $\mathcal{B}(S)$ be a Borel σ-field on S, that is, the σ-field generated by all open sets of metric space S. For a set $E \subset S$, we denote by $\mathcal{B}(E)$ the σ-field whose elements are intersections of E with elements of $\mathcal{B}(S)$. Observe that E is a metric space with the same metric as on S, and $\mathcal{B}(E)$ is its Borel σ-field. For a metric space S, we denote by $\mathbb{P}(S)$ the set of probability measures on $(S, \mathcal{B}(S))$. A sequence of probability measures $\{\mu_n\}$ from $\mathbb{P}(S)$ converges weakly to

$\mu \in \mathbb{P}(S)$ if for any bounded continuous function f on S

$$\int_S f(s)\mu_n(ds) \to \int_S f(s)\mu(ds) \qquad \text{as } n \to \infty.$$

Consider a discrete-time MDP with a *state space* $\mathbb{X}$, an *action space* $\mathbb{A}$, one-step costs c, and transition pobabilities q. Assume that $\mathbb{X}$ and $\mathbb{A}$ are *Borel subsets* of Polish (complete separable metric) spaces with the corresponding metrics ρ and γ. For all $x \in \mathbb{X}$ a nonempty Borel subset $A(x)$ of $\mathbb{A}$ represents the *set of actions* available at x. Define the graph of A by

$$\mathrm{Gr}(A) = \{(x,a) \,:\, x \in \mathbb{X}, a \in A(x)\}.$$

Assume also that

(i) $\mathrm{Gr}(A)$ is a measurable subset of $\mathbb{X} \times \mathbb{A}$, that is, $\mathrm{Gr}(A) \in \mathcal{B}(\mathrm{Gr}(A))$, where $\mathcal{B}(\mathrm{Gr}(A)) = \mathcal{B}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{A})$;

(ii) there exists a measurable mapping $\phi : \mathbb{X} \to \mathbb{A}$ such that $\phi(x) \in A(x)$ for all $x \in \mathbb{X}$;

The *one step cost,* $c(x,a) \le +\infty$, for choosing an action $a \in A(x)$ in a state $x \in \mathbb{X}$, is a *bounded below measurable* function on $\mathrm{Gr}(A)$. Let $q(B|x,a)$ be the *transition kernel* representing the probability that the next state is in $B \in \mathcal{B}(\mathbb{X})$, given that the action a is chosen in the state x. This means that:

- $q(\cdot|x,a)$ is a probability measure on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ for all $(x,a) \in \mathbb{X} \times \mathbb{A}$;
- $q(B|\cdot,\cdot)$ is a Borel function on $(\mathrm{Gr}(A), \mathcal{B}(\mathrm{Gr}(A)))$ for all $B \in \mathcal{B}(\mathbb{X})$.

The decision process proceeds as follows:

- at each time epoch $n = 0, 1, ...$ the current state $x \in \mathbb{X}$ is observed;
- a decision-maker chooses an action $a \in A(x)$;
- the cost $c(x,a)$ is incurred;
- the system moves to the next state according to the probability law $q(\cdot|x,a)$.

As explained in the text following the proof of Lemma 3.3, if for each $x \in \mathbb{X}$ there exists $a \in A(x)$ with $c(x,a) < \infty$, the measurability of $\mathrm{Gr}(A)$ and inf-compactness of the cost function c in the action variable a assumed later imply that assumption (ii) holds.

Let $\mathbb{H}_n = (\mathbb{X} \times \mathbb{A})^n \times \mathbb{X}$ be the *set of histories* by time $n = 0, 1, ...$ and $\mathcal{B}(\mathbb{H}_n) = (\mathcal{B}(\mathbb{X}) \otimes \mathcal{B}(\mathbb{A}))^n \otimes \mathcal{B}(\mathbb{X})$. A *randomized decision rule* at epoch $n = 0, 1, ...$ is a regular transition probability $\pi_n : H_n \to \mathbb{A}$ concentrated on $A(\xi_n)$, that is, (i) $\pi_n(\cdot \mid h_n)$ is a probability on $(\mathbb{A}, \mathcal{B}(\mathbb{A}))$, given the history $h_n = (\xi_0, u_0, \xi_1, u_1, ..., u_{n-1}, \xi_n) \in \mathbb{H}_n$, satisfying $\pi_n(A(\xi_n)|h_n) = 1$, and (ii) for all $B \in \mathcal{B}(A)$, the function $\pi_n(B|\cdot)$ is Borel on $(\mathbb{H}_n, \mathcal{B}(\mathbb{H}_n))$. A *policy* is a sequence $\pi = \{\pi_n\}_{n=0,1,...}$ of decision rules. Moreover, π is called *nonrandomized*, if each probability measure $\pi_n(\cdot|h_n)$ is concentrated at one point. A nonrandomized policy is called *Markov*, if all of the decisions depend on the current state and time only. A Markov policy is called *stationary*, if all the decisions depend on the current state only. Thus, a Markov policy ϕ is defined by a sequence $\phi_0, \phi_1, \ldots$ of Borel mappings $\phi_n : \mathbb{X} \to \mathbb{A}$ such that $\phi_n(x) \in A(x)$ for all $x \in \mathbb{X}$. A stationary policy ϕ is defined by a Borel mapping $\phi : \mathbb{X} \to \mathbb{A}$ such that $\phi(x) \in A(x)$ for all $x \in \mathbb{X}$. Let

$$\mathbb{F} = \{\phi : \mathbb{X} \to \mathbb{A} \,:\, \phi \text{ is Borel and } \phi(x) \in A(x) \text{ for all } x \in \mathbb{X}\}$$

be the *set of stationary policies*.

The Ionescu Tulcea theorem (Bertsekas and Shreve [4, pp. 140-141] or Hernández-Lerma and Lassere [17, p.178]) implies that an initial state x and a policy π define a unique probability P_x^π on the set of all trajectories $\mathbb{H}_\infty = (\mathbb{X} \times \mathbb{A})^\infty$ endowed with the product of σ-field defined by Borel σ-field of $\mathbb{X}$ and $\mathbb{A}$. Let $\mathbb{E}_x^\pi$ be an expectation with respect to P_x^π.

For a finite horizon $N = 0, 1, ...,$ let us define the *expected total discounted costs*

$$v_{N,\alpha}^\pi := \mathbb{E}_x^\pi \sum_{n=0}^{N-1} \alpha^n c(\xi_n, u_n), \qquad x \in \mathbb{X}, \tag{2.1}$$

where $\alpha \geq 0$ is the discount factor and $v_{0,\alpha}^\pi(x) = 0$. When $\alpha = 1$, we shall write $v_N^\pi(x)$ instead of $v_{N,1}^\pi(x)$. When $N = \infty$ and $\alpha \in [0, 1)$, (2.1) defines an *infinite horizon expected total discounted cost* denoted by $v_\alpha^\pi(x)$.

The *average cost per unit time* is defined as

$$w^\pi(x) := \limsup_{N \to +\infty} \frac{1}{N} v_N^\pi(x), \qquad x \in \mathbb{X}. \tag{2.2}$$

For any function $g^\pi(x)$, including $g^\pi(x) = v_{N,\alpha}^\pi(x)$, $g^\pi(x) = v_\alpha^\pi(x)$, and $g^\pi(x) = w^\pi(x)$, define the *optimal cost*

$$g(x) := \inf_{\pi \in \Pi} g^\pi(x), \qquad x \in \mathbb{X},$$

where Π is *the set of all policies*.

A policy π is called *optimal* for the respective criterion, if $g^\pi(x) = g(x)$ for all $x \in \mathbb{X}$. For $g^\pi = v_{n,\alpha}^\pi$, the optimal policy is called n*-horizon discount-optimal*; for $g^\pi = v_\alpha^\pi$, it is called *discount-optimal*; for $g^\pi = w^\pi$, it is called *average-cost optimal*.

It is well known (see, e.g., Bertsekas and Shreve [4, Proposition 8.2]) that the functions $v_{n,\alpha}(x)$ recursively satisfy the following *optimality equations* with $v_{0,\alpha}(x) = 0$ for all $x \in \mathbb{X}$,

$$v_{n+1,\alpha}(x) = \inf_{a \in A(x)} \left\{ c(x, a) + \alpha \int_{\mathbb{X}} v_{n,\alpha}(y) q(dy|x, a) \right\}, \quad x \in \mathbb{X}, \ n = 0, 1, \dots . \tag{2.3}$$

In addition, a Markov policy ϕ, defined at the first N steps by the mappings $\phi_0, ...\phi_{N-1}$, that satisfy for all $n = 1, ..., N$ the equations

$$v_{n,\alpha}(x) = c(x, \phi_{N-n}(x)) + \alpha \int_{\mathbb{X}} v_{n-1,\alpha}(y) q(dy|x, \phi_{N-n}(x)), \quad x \in \mathbb{X}, \tag{2.4}$$

is optimal for the horizon N; see e.g. Bertsekas and Shreve [4, Lemma 8.7].

It is also well known (Bertsekas and Shreve [4, Propositions 9.8 and 9.12]) that v_α, where $\alpha \in (0, 1]$, satisfies the following discounted cost optimality equation (DCOE):

$$v_\alpha(x) = \inf_{a \in A(x)} \left\{ c(x, a) + \alpha \int_{\mathbb{X}} v_\alpha(y) q(dy|x, a) \right\}, \quad x \in \mathbb{X}, \tag{2.5}$$

and a stationary policy ϕ_α is discount-optimal if and only if

$$v_\alpha(x) = c(x, \phi_\alpha(x)) + \alpha \int_{\mathbb{X}} v_\alpha(y) q(dy|x, \phi_\alpha(x)), \quad x \in \mathbb{X}. \tag{2.6}$$

3 General Assumptions and Auxiliary Results

Following Schäl [24], consider the following assumption.

Assumption (G). $w^* := \inf_{x\in\mathbb{X}} w(x) < +\infty$.

This assumption is equivalent to the existence of $x \in \mathbb{X}$ and $\pi \in \Pi$ with $w^\pi(x) < \infty$. If Assumption (**G**) does not hold then the problem is trivial, because $w(x) = \infty$ for all $x \in \mathbb{X}$ and any policy π is average-cost optimal. Define the following quantities for $\alpha \in [0,1)$:

$$m_\alpha = \inf_{x\in\mathbb{X}} v_\alpha(x), \quad u_\alpha(x) = v_\alpha(x) - m_\alpha,$$

$$\underline{w} = \liminf_{\alpha\uparrow 1}(1-\alpha)m_\alpha, \quad \overline{w} = \limsup_{\alpha\uparrow 1}(1-\alpha)m_\alpha.$$

Observe that $u_\alpha(x) \ge 0$ for all $x \in \mathbb{X}$. According to Schäl [24, Lemma 1.2], Assumption (**G**) implies

$$0 \le \underline{w} \le \overline{w} \le w^* < +\infty. \tag{3.1}$$

According to Schäl [24, Proposition 1.3], under Assumption (**G**), if there exists a measurable function $u : \mathbb{X} \to [0, +\infty)$ and a stationary policy ϕ such that

$$\underline{w} + u(x) \ge c(x, \phi(x)) + \int_{\mathbb{X}} u(y) q(dy|x, \phi(x)), \quad x \in \mathbb{X}, \tag{3.2}$$

then ϕ is *average-cost optimal* and $w(x) = w^* = \underline{w} = \overline{w}$ for all $x \in \mathbb{X}$. Here need a different form of such a statement.

Theorem 3.1. *Let Assumption (G) hold. If there exists a measurable function $u : \mathbb{X} \to [0, +\infty)$ and a stationary policy ϕ such that*

$$\overline{w} + u(x) \ge c(x, \phi(x)) + \int_{\mathbb{X}} u(y) q(dy|x, \phi(x)), \quad x \in \mathbb{X}, \tag{3.3}$$

then ϕ is average-cost optimal and

$$w(x) = w^\phi(x) = \limsup_{\alpha\uparrow 1}(1-\alpha)v_\alpha(x) = \overline{w} = w^*, \quad x \in \mathbb{X}. \tag{3.4}$$

Proof. Similarly to Hernández-Lerma [16, p. 239] or Schäl [24, Proposition 1.3], since u is non-negative, by iterating (3.3) we obtain

$$n\overline{w} + u(x) \ge v_n^\phi(x), \quad n \ge 1,\ x \in \mathbb{X}.$$

Therefore, after dividing the last inequality by n and setting $n \to \infty$, we have

$$\overline{w} \ge w^\phi(x) \ge w(x) \ge w^*, \quad x \in \mathbb{X}, \tag{3.5}$$

where the second and the third inequalities follow from the definitions of w and w^* respectively. Since $\overline{w} \ge w^*$, inequalities (3.1) imply that for all $\pi \in \Pi$

$$w^* = \overline{w} \le \limsup_{\alpha\uparrow 1}(1-\alpha)v_\alpha(x) \le \limsup_{\alpha\uparrow 1}(1-\alpha)v_\alpha^\pi(x) \le w^\pi(x), \qquad \pi \in \Pi,\ x \in \mathbb{X}.$$

Finally, we obtain that

$$w^* = \overline{w} \leq \limsup_{\alpha\uparrow 1}(1-\alpha)v_\alpha(x) \leq \inf_{\pi\in\Pi} w^\pi(x) = w(x) \leq w^\phi(x) \leq \overline{w}, \qquad x \in \mathbb{X}, \tag{3.6}$$

where the last inequality follows from (3.5). Thus all the inequalities in (3.6) are equalities. □

Let us set $\mathbb{R} = [-\infty, +\infty)$, $\mathbb{R}_+ = [0, \infty)$, and $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$. For an $\overline{\mathbb{R}}$-valued function f, defined on a Borel subset U of a Polish space $\mathbb{Y}$, consider the level sets

$$\mathcal{D}_f(\lambda) = \{y \in U \,:\, f(y) \leq \lambda\}, \tag{3.7}$$

$-\infty < \lambda < +\infty$. We recall that the function f is *lower semi-continuous on* U if all the level sets $\mathcal{D}_f(\lambda)$ are closed and the function is *inf-compact on* U if all these sets are compact. The level sets $\mathcal{D}_f(\lambda)$ satisfy the following properties that are used in this paper:

(a) if $\lambda_1 > \lambda$ then $\mathcal{D}_f(\lambda) \subseteq \mathcal{D}_f(\lambda_1)$;

(b) if g, f are functions on U satisfying $g(y) \geq f(y)$ for all $y \in U$ then $\mathcal{D}_g(\lambda) \subseteq \mathcal{D}_f(\lambda)$.

A set is called σ-compact if it is a union of a countable number of compact sets. Denote by $K(\mathbb{A})$ the *family of all nonempty compact subsets* of $\mathbb{A}$ and by $K_\sigma(\mathbb{A})$ *family of all* σ*-compact subsets* of $\mathbb{A}$; $K(\mathbb{A}) \subset K_\sigma(\mathbb{A})$. Also denote by $S(\mathbb{A})$ the set of nonempty subsets of $\mathbb{A}$.

A set-valued mapping $F : \mathbb{X} \to S(\mathbb{A})$ is *upper semi-continuous* at $x \in \mathbb{X}$ if, for any neighborhood G of the set $F(x)$, there is a neighborhood of x, say $U(x)$, such that $F(y) \subseteq G$ for all $y \in U(x)$ (see e.g., Berge [3, p. 109] or Zgurovsky et al. [30, Chapter 1, p. 7]). A set-valued mapping is called *upper semi-continuous*, if it is upper semi-continuous at all $x \in \mathbb{X}$.

For weakly continuous transition probabilities, the following basic assumptions were considered in Schäl [24].

Assumption (W).

(i) c is lower semi-continuous and bounded below on $\mathrm{Gr}(A)$;

(ii) $A(x) \in K(\mathbb{A})$ for $x \in \mathbb{X}$ and $A : \mathbb{X} \to K(\mathbb{A})$ is upper semi-continuous;

(iii) the transition probability $q(\cdot|x, a)$ is weakly continuous in $(x, a) \in \mathrm{Gr}(A)$.

Weak continuity of q in (x, a) means that

$$\int_{\mathbb{X}} f(z)q(dz|x_k, a_k) \to \int_{\mathbb{X}} f(z)q(dz|x, a), \qquad k = 1, 2, \ldots,$$

for any sequence $\{(x_k, a_k), k \geq 0\}$ converging to (x, a), where (x_k, a_k), $(x, a) \in \mathrm{Gr}(A)$, and for any bounded continuous function $f : \mathbb{X} \to \mathbb{R}$. We notice that there is an additional assumption in Schäl [24], namely, that $\mathbb{X}$ is a locally compact space with countable base. However, as follows from this paper, the assumption is not necessary here as well as in Feinberg and Lewis [14], since there exists at least one stationary policy. We also remark that the assumptions in **(W)** were presented in a different order here than in Schäl [24], and that it is assumed in Schäl [24] that c is nonnegative. Since for discounted and average cost criteria the cost function can be shifted by adding any constant, the boundness and nonnegativity of c are equivalent assumptions. We consider Assumption (**Wu**) from Feinberg and Lewis [14] without assuming that $\mathbb{X}$ is locally compact.

Assumption (Wu).

(i) c is inf-compact on $\mathrm{Gr}(A)$;

(ii) Assumption (**W**)(iii) holds.

Assumption (W*).

(i) Assumption (**W**)(i) holds;

(ii) if a sequence $\{x_n\}_{n=1,2,\dots}$ with values in $\mathbb{X}$ converges and its limit x belongs to $\mathbb{X}$ then any sequence $\{a_n\}_{n=1,2,\dots}$ with $a_n \in A(x_n)$, $n = 1, 2, \dots$, satisfying the condition that the sequence $\{c(x_n, a_n)\}_{n=1,2,\dots}$ is bounded above, has a limit point $a \in A(x)$;

(iii) Assumption (**W**)(iii) holds.

Lemma 3.2. *The following statements hold:*

*(i) Assumption (**W**) implies Assumption (**W***);*

*(ii) Assumption (**Wu**) implies Assumption (**W***).*

Proof. (i) Let $x_n \to x$ as $n \to \infty$, where $x \in \mathbb{X}$ and $x_n \in \mathbb{X}$, $n = 1, \dots$. We show that under Assumption (**W**)(ii) any sequence $\{a_n\}_{n=1,2,\dots}$ with $a_n \in A(x_n)$ has a limit point $a \in A(x)$. Indeed, since $\mathcal{K} := (\cup_{n\geq 1}\{x_n\}) \cup \{x\}$ is a compact set and set-valued mapping $A : \mathbb{X} \to K(\mathbb{A})$ is upper semi-continuous, then Berge [3, Theorem 3 on p. 110] implies that the image $A(\mathcal{K})$ is also compact. As $\{a_n\}_{n\geq 1} \subset A(\mathcal{K})$ then the sequence $\{a_n\}_{n\geq 1}$ has a limit point $a \in \mathbb{A}$. Consider a sequence $n_k \to \infty$ such that $a_{n_k} \to a$. Since $A(z) \in K(\mathbb{A})$ for all $z \in X$, the upper-semicontinuous set-valued mapping A is closed and, since A is closed, $a \in A(x)$; Berge [3, Theorems 5 and 6 on pp. 111, 112].

(ii) Since c is inf-compact, it is lower-semicontinuous and bounded below. We just need to show that Assumption (**W***)(ii) holds. Let us consider $x_n \to x$ as $n \to +\infty$ and $a_n \in A(x_n)$, $n = 1, , 2, \dots$, such that $x_n, x \in \mathbb{X}$ and for some $\lambda < \infty$ the inequality $c(x_n, a_n) \leq \lambda$ holds for all $n = 1, 2, \dots$. Then, by inf-compactness of c on $\mathrm{Gr}(A)$, the level set $\mathcal{D}_c(\lambda)$ is compact. Thus the sequence $\{x_n, a_n\}_{n\geq 1}$ has a limit point $(x, a) \in \mathcal{D}_c(\lambda) \subseteq \mathrm{Gr}(A)$. Since $(x, a) \in \mathrm{Gr}(A)$, we have $a \in A(x)$. $\square$

For any $\alpha \geq 0$ and lower semi-continuous nonnegative function $u : \mathbb{X} \to \overline{\mathbb{R}}$, we consider an *operation* η_u^α,

$$\eta_u^\alpha(x, a) = c(x, a) + \alpha \int_{\mathbb{X}} u(y) q(dy|x, a), \quad (x, a) \in \mathrm{Gr}(A). \tag{3.8}$$

Let $L(\mathbb{X})$ be the class of all lower semi-continuous and bounded below functions $\varphi : \mathbb{X} \to \overline{\mathbb{R}}$ with $\mathrm{dom}\, \varphi := \{x \in \mathbb{X} \,:\, \varphi(x) < +\infty\} \neq \emptyset$. Observe that $\eta_u^\alpha = \eta_{\alpha u}^1$.

Lemma 3.3. *For any $x \in \mathbb{X}$ the following statements hold:*

*(a) under Assumption **W***(ii), the function $c(x, \cdot)$ is inf-compact on $A(x)$;*

(b) under Assumptions $\mathbf{W}^*$*(ii,iii), for any* $u \in L(\mathbb{X})$ *and* $\alpha \geq 0$*, the function* $\eta_u^\alpha(x,\cdot)$ *is inf-compact on* $A(x)$.

Proof. (a) For an arbitrary $\lambda \in \mathbb{R}$ and fixed $x \in \mathbb{X}$, consider the set $\mathcal{D}_{c(x,\cdot)}(\lambda) = \{a \in A(x) : c(x,a) \leq \lambda\}$. Assumption $\mathbf{W}^*$(ii) means, that this set is compact. Thus, (i) is proved.

(b) Fix $x \in \mathbb{X}$ again. Since $u \in L(\mathbb{X})$ and q is weakly continuous in a, the second summand in (3.8) is a lower semi-continuous function on $A(x)$ (Hernńdez-Lerma and Lasserre [17, p. 185]) and it is bounded below by the same constant as u. According to statement (i), $c(x,\cdot)$ is inf-compact on $A(x)$. The sum of an inf-compact function and a bounded below lower semi-continuous function is an inf-continuous function. ☐

A measurable mapping $\phi : \mathbb{X} \to \mathbb{A}$, such that $\phi(x) \in A(x)$ for all $x \in \mathbb{X}$, is called a selector (or a measurable selector). In our case, selectors and decision rules are the same objects. Since we identify a stationary policy with a decision rule, selectors and stationary policies are the same objects. The existence of selector for the mapping A is the necessary and sufficient condition for the existence of a policy. Let $E \subseteq \mathbb{X} \times \mathbb{A}$ and $\mathrm{proj}_{\mathbb{X}}\, E = \{x \in \mathbb{X} : (x,a) \in E \text{ for some } a \in E\}$ be a projection of E on X. A Borel map $f :\ \mathrm{proj}_{\mathbb{X}}\, E \to \mathbb{A}$ is called a Borel uniformization of E, if $(x, f(x)) \in E$ for all $x \in \mathrm{proj}_{\mathbb{X}}\, E$. Let $E_x = \{a : (x,a) \in E\}$ be a cut of E at $x \in \mathbb{X}$.

Arsenin-Kunugui Theorem (Kechris [19, p. 297]) *If* E *is a Borel subset of* $\mathbb{X} \times \mathbb{A}$ *and* $E_x \in K_\sigma(\mathbb{A})$ *for all* $x \in \mathbb{X}$ *then there exists a Borel uniformization of* E *and* $\mathrm{proj}_{\mathbb{X}}\, E$ *is a Borel set.*

We remark that it is assumed in Kechris [19, p. 297]) that $\mathbb{X}$ is a standard Borel space (that is, isomorphic to a Borel subset of a Polish space) and $\mathbb{A}$ is a Polish space. Here $\mathbb{X}$ and $\mathbb{A}$ are Borel subsets of Polish spaces. These two formulations are obviously equivalent.

We recall that $\mathrm{Gr}(A)$ is assumed to be Borel and $A(x) \neq \emptyset$, $x \in \mathbb{X}$. With $E = \mathrm{Gr}(A)$, Arsenin-Kunugui Theorem implies the existence of a stationary policy under the assumption $A(x) \in K(\mathbb{A})$, $x \in \mathbb{X}$. Thus, Assumption (**W**) implies the existence of a policy for the MDP.

Let Assumption ($\mathbf{W}^*$) hold. Set $F(x) = \{a \in A(x) : c(x,a) < \infty\}$, $x \in \mathbb{X}$. In view of Lemma 3.3, $F(x) = \cup_{n \in \{1,2,\dots\}} \mathcal{D}_{c(x,\cdot)}(n) \in K_\sigma(\mathbb{A})$. In addition, $\mathrm{Gr}(F) = \{(x,a) \in \mathrm{Gr}(A) : c(x,a) < \infty\}$ is a Borel subset of $\mathbb{X} \times \mathbb{A}$. Thus, if the function c takes only finite values, a stationary policy exists in view of Arsenin-Kunugui Theorem.

Of course, if it is possible that $c(x,a) = \infty$, a uniformization may not exist. For example, this takes place when $c(x,a) = \infty$ for all $(x,a) \in \mathrm{Gr}(A)$ and $\mathrm{Gr}(A)$ does not have a measurable selector. However $c(x,a) = \infty$ means from a modeling prospective that this state-action pair should be excluded, because selecting a in x leads to the worst possible result. If there are state-action pairs (x,a) with $c(x,a) = \infty$ and $\mathrm{Gr}(A)$ does not have a uniformization, the MDP can be transformed into an MDP modeling the same problem and with a nonempty set of policies. Let us exclude the situation when $c(x,a) = \infty$ for all $(x,a) \in \mathrm{Gr}(A)$, because it is trivial: all the actions are bad. Define $X = \mathrm{proj}_{\mathbb{X}}\, \mathrm{Gr}(F)$ and $Y = \mathbb{X} \setminus X$. Under Assumption ($\mathbf{W}^*$), Arsenin-Kunigui Theorem implies that X is Borel and there exist a Borel mapping f from X to $\mathbb{A}$ such that

$f(x) \in F(x)$ for all $x \in X$. If $Y = \emptyset$ (that is, there exists an action $a \in A(x)$ with $c(x,a) < \infty$ for each $x \in \mathbb{X}$) then $\phi = f$ is a stationary policy.

Let us consider the situation when $Y \neq \emptyset$. In such an MDP, as soon as the state is in Y, the losses are infinite and there is no reason to model the process after this. Let us transform the model by choosing any $x^* \in Y$ and any $a^* \in \mathbb{A}$ and setting the new state set $\mathbb{X}^* = X \cup \{x^*\}$, keeping the original action set $\mathbb{A}$, setting new action sets $A^*(x) = F(x)$ for $x \in X$ and $A^*(x^*) = \{a^*\}$, defining the new cost function

$$c^*(x,a) = \begin{cases} c(x,a), & \text{if } x \in Y \text{ and } a \in F(x), \\ \infty, & \text{if } x = x^* \text{ and } a = a^*. \end{cases}$$

and considering new transition probabilities defined for $x \in X^*$ and $a \in A^*(x)$ by

$$q^*(B|x,a) = \begin{cases} q(B|x,a), & \text{if } B \subseteq X,\ B \in \mathcal{B}(\mathbb{X}),\ \text{and }\ x \in X, \\ q(Y|x,a), & \text{if } B = \{x^*\},\ \text{and } x \in X, \\ 1, & \text{if } B = \{x^*\} \text{ and } x = x^*. \end{cases}$$

The new MDP is nontrivial in the sense that the set of policies is not empty. Finding an optimal policy for this MDP is equivalent to finding a policy for the original MDP until its first exit time from X, and in both cases the process incurs infinite losses, if it leaves X. So, the original and the new MDP model are the same problem.

Lemma 3.4. *If Assumption (*$\mathbf{W}^*$*) holds and* $u \in L(\mathbb{X})$, *then the function*

$$u^*(x) := \inf_{a \in A(x)} \left[c(x,a) + \int_{\mathbb{X}} u(y) q(dy|x,a) \right], \qquad x \in \mathbb{X}, \tag{3.9}$$

belongs to $L(\mathbb{X})$, *and there exists* $f \in \mathbb{F}$ *such that*

$$u^*(x) = c(x, f(x)) + \int_{\mathbb{X}} u(y) q(dy|x, f(x)), \qquad x \in \mathbb{X}. \tag{3.10}$$

Moreover, infimum in (3.9) can be replaced by minimum, and the nonempty sets

$$A_*(x) = \left\{ a \in A(x) :\ u^*(x) = c(x,a) + \int_{\mathbb{X}} u(y) q(dy|x,a) \right\}, \qquad x \in \mathbb{X}, \tag{3.11}$$

satisfy the following properties:

(a) the graph $\mathrm{Gr}(A_*) = \{(x,a) :\ x \in \mathbb{X}, a \in A_*(x)\}$ *is a Borel subset of* $\mathbb{X} \times \mathbb{A}$*;*

(b) if $u^*(x) = +\infty$, *then* $A_*(x) = A(x)$, *and, if* $u^*(x) < +\infty$, *then* $A_*(x)$ *is compact.*

Proof. Under Assumption ($\mathbf{W}^*$), for any lower semi-continuous on $\mathbb{X}$, bounded below function $u : \mathbb{X} \to \overline{\mathbb{R}}$ and $\alpha \in (0,1]$, the function $\eta^{\alpha}_{u(x,\cdot)}$ is inf-compact on $A(x)$, $x \in \mathbb{X}$. This follows from Lemma 3.3. Thus, infimum in (3.9) can be replaced by minimum and $A^*(x)$ is nonempty for any $x \in \mathbb{X}$.

Now we show that u^* is lower semi-continuous on $\mathbb{X}$. Let us fix an arbitrary $x \in \mathbb{X}$ and any sequence $x_n \to x$ as $n \to +\infty$. We need to prove the inequality

$$u^*(x) \le \liminf_{n\to+\infty} u^*(x_n). \tag{3.12}$$

If $\liminf_{n\to+\infty} u^*(x_n) = +\infty$, then (3.12) obviously holds. Thus we consider the case, when $\liminf_{n\to+\infty} u^*(x_n) < +\infty$. There exists a subsequence $\{x_{n_k}\}_{k\ge1} \subseteq \{x_n\}_{n\ge1}$ such that

$$\liminf_{n\to+\infty} u^*(x_n) = \lim_{k\to+\infty} u^*(x_{n_k}).$$

Setting $\lambda = \lim_{k\to+\infty} u^*(x_{n_k}) + 1$, we get the inequality $u^*(x_{n_k}) \le \lambda$ for all $k \ge K$, where K is some natural number. Since the function η_u^1 is inf-compact on $\mathrm{Gr}(A)$, equation (3.9) can be rewritten as

$$u^*(x) := \min_{a\in A(x)} \eta_u^1(x,a), \ \ x \in \mathbb{X}.$$

Thus, for any $k \ge K$ there exists $a_k \in A(x_{n_k})$ such that $u^*(x_{n_k}) = \eta_u^1(x_{n_k}, a_k)$. Therefore,

$$c(x_{n_k}, a_k) \le \eta_u^1(x_{n_k}, a_k) \le \lambda, \qquad k \ge K.$$

In view of Assumption $(\mathbf{W}^*)$(ii), there exists a convergent subsequence $\{a_{k_m}\}_{m\ge1}$ of the sequence $\{a_k\}_{k\ge1}$ such that $a_{k_m} \to a \in A(x)$ as $m \to +\infty$. Due to lower semi-continuity of η_u^1 on $\mathrm{Gr}(A)$,

$$\liminf_{n\to+\infty} u^*(x_n) = \lim_{k\to+\infty} u^*(x_{n_k}) = \lim_{m\to+\infty} u^*(x_{n_{k_m}}) = \lim_{m\to+\infty} \eta_u^1(x_{n_{k_m}}, a_{k_m}) \ge \eta_u^1(x,a) \ge u^*(x).$$

Inequality (3.12) holds. Thus, u^* is lower semi-continuous on $\mathbb{X}$.

Now we consider the nonempty sets $A_*(x)$, $x \in \mathbb{X}$, defined in (3.11). The graph $\mathrm{Gr}(A_*)$ is a Borel subset of $\mathbb{X} \times \mathbb{A}$, because $\mathrm{Gr}(A_*) = \{(x,a) \,:\, u^*(x) = \eta_u^1(x,a)\}$, and the functions η_u^1 and u^* are lower semi-continuous on $\mathrm{Gr}(A)$ and $\mathbb{X}$ respectively, and therefore they are Borel.

We remark that, if $u^* = +\infty$, then $A_*(x) = A(x)$. If $u^*(x) < \infty$, then Lemma 3.3 implies that the set $A_*(x)$ is compact. Indeed, fix any $x \in \mathbb{X}_f := \{x \in \mathbb{X} \,:\, u^*(x) < \infty\}$ and set $\lambda = u^*(x)$. Then the set $A_*(x) = \{a \in A(x) \,:\, \eta_u^1(x,a) \le \lambda\} = \mathcal{D}_{\eta_u^1(x,\cdot)}(\lambda)$ is compact, because $\eta_u^1(x,\cdot)$ is inf-compact on $A(x)$.

Let us prove the existence of $f \in \mathbb{F}$ satisfying (3.10). Since the function u^* is lower-semicontinuous, it is Borel and the sets $X_\infty := \{x \in \mathbb{X} \,:\, u^*(x) = +\infty\}$ and $\mathbb{X}_f$ are Borel. Therefore, the graph of the mapping $\mathbb{X}_f \to A_*$ is the Borel set $\mathrm{Gr}(A_*) \setminus (\mathbb{X}_\infty \times \mathbb{A})$. Since the nonempty sets $A_*(x)$ are compact for all $x \in \mathbb{X}_f$, the Arsenin-Kunugui Theorem implies the existence of a Borel selector $f_1 : \mathbb{X}_f \to \mathbb{A}$ such that $f_1(x) \in A_*(x)$ for all $x \in \mathbb{X}$. Consider any Borel mapping f_2 from $\mathbb{X}$ to $\mathbb{A}$ satisfying $f_2(x) \in A(x)$ for all $x \in \mathbb{X}$ and set

$$f(x) = \begin{cases} f_1(x), & \text{if } x \in \mathbb{X}_f, \\ f_2(x), & \text{if } x \in \mathbb{X}_\infty. \end{cases}$$

Then $f \in \mathbb{F}$ and $f(x) \in A_*(x)$ for all $x \in \mathbb{X}$. □

The following Lemma 3.5 is formulated in Schäl [24, Lemma 2.3(ii)] without proof. Reference Serfozo [27] mentioned in Schäl [24, Lemma 2.3(ii)] contains relevant facts, but it does not contain this statement. Therefore we provide the proof. Recall that for a metric space S, the family of all probability measures on $(S,\mathcal{B}(S))$ is denoted by $\mathbb{P}(S)$.

Lemma 3.5. *Let S be an arbitrary metric space, $\{\mu_n\}_{n\ge 1}\subset\mathbb{P}(S)$ converges weakly to $\mu\in\mathbb{P}(S)$, and $\{h_n\}_{n\ge1}$ be a sequence of measurable nonnegative $\overline{\mathbb{R}}$-valued functions on S. Then*

$$\int_S \underline{h}(s)\mu(ds)\le \liminf_{n\to+\infty}\int_S h_n(s)\mu_n(ds),$$

where $\underline{h}(s)=\liminf_{n\to+\infty,\,s'\to s} h_n(s')$, $s\in S$.

Proof. See Appendix **A**. □

We remark that $\liminf_{n\to+\infty,\,s'\to s} h_n(s')$ is the least upper bound of the set of all $\lambda\in\mathbb{R}$ such that there exist $N=1,2,\dots$ and a neighborhood $U(s)$ of s such that $\lambda\le\inf\{h_n(s'):\,n\ge N, s'\in U(s)\}$.

4 Expected Total Discounted Costs

In this section, we establish under Assumption (**W***) the standard properties of discounted MDPs: the existence of stationary optimal policies, description of the sets of stationary optimal policy, and convergence of value iterations. Theorem 4.1 strengthens Feinberg and Lewis [14, Proposition 3.1], where these facts are proved under Assumption (**Wu**). In terms of applications to inventory and queuing control, Assumption (**W***) does not require that holding costs increase to infinity as the inventory level (or workload, or the number of customers in queue) increases to infinity.

Theorem 4.1. *Let Assumption (**W***) hold. Then*

(i) the functions $v_{n,\alpha}$, $n=1,2,\dots$, and v_α are lower semi-continuous on $\mathbb{X}$, and $v_{n,\alpha}(x)\uparrow v_\alpha(x)$ as $n\to+\infty$ for all $x\in\mathbb{X}$;

(ii)

$$v_{n+1,\alpha}(x)=\min_{a\in A(x)}\left\{c(x,a)+\alpha\int_{\mathbb{X}} v_{n,\alpha}(y)q(dy|x,a)\right\},\quad x\in\mathbb{X},\ n=0,1,..., \tag{4.1}$$

where $v_{0,\alpha}(x)=0$ for all $x\in\mathbb{X}$, and the nonempty sets $A_{n,\alpha}(x):=\{a\in A(x):\ v_{n+1,\alpha}(x)=\eta^\alpha_{v_{n,\alpha}}(x,a)\}$, $x\in\mathbb{X}$, $n=0,1,\dots$, satisfy the following properties: (a) the graph $\mathrm{Gr}(A_{n,\alpha})=\{(x,a):\,x\in\mathbb{X},a\in A_\alpha(x)\}$, $n=0,1,\dots$, is a Borel subset of $\mathbb{X}\times\mathbb{A}$, and (b) if $v_{n+1,\alpha}(x)=+\infty$, then $A_{n,\alpha}(x)=A(x)$ and, if $v_{n+1,\alpha}(x)<+\infty$, then $A_{n,\alpha}(x)$ is compact;

(iii) for any $N=1,2,\dots$, there exists a Markov optimal N-horizon policy $(\phi_0,\dots,\phi_{N-1})$ and if, for an N-horizon Markov policy $(\phi_0,\dots,\phi_{N-1})$ the inclusions $\phi_{N-1-n}(x)\in A_{\alpha,n}(x)$, $x\in\mathbb{X}$, $n=0,\dots,N-1$, hold then this policy is N-horizon optimal;

(iv) for $\alpha \in [0,1)$

$$v_\alpha(x) = \min_{a\in A(x)} \left\{ c(x,a) + \alpha \int_{\mathbb{X}} v_\alpha(y) q(dy|x,a) \right\}, \qquad x \in \mathbb{X}, \tag{4.2}$$

and the nonempty sets $A_\alpha(x) := \{a \in A(x) : v_\alpha(x) = \eta^\alpha_{v_\alpha}(x,a)\}$, $x \in \mathbb{X}$, *satisfy the following properties: (a) the graph* $\mathrm{Gr}(A_\alpha) = \{(x,a) : x \in \mathbb{X}, a \in A_\alpha(x)\}$ *is a Borel subset of* $\mathbb{X}\times\mathbb{A}$, *and (b) if* $v_\alpha(x) = +\infty$, *then* $A_\alpha(x) = A(x)$ *and, if* $v_\alpha(x) < +\infty$, *then* $A_\alpha(x)$ *is compact.*

(v) for an infinite-horizon there exists a stationary discount-optimal policy ϕ_α, *and a stationary policy is optimal if and only if* $\phi_\alpha(x) \in A_\alpha(x)$ *for all* $x \in \mathbb{X}$.

(vi) (Feinberg and Lewis [14, Proposition 3.1(iv)]) *under Assumption (*$\mathbf{Wu}$*), the functions* $v_{n,\alpha}$, $n = 1, 2, \ldots$, *and* v_α *are inf-compact on* $\mathbb{X}$.

Proof. (i)–(v). First, we prove these statements for a nonnegative cost function c. In this case, $v_{n,\alpha}(x) \geq 0$, $n = 0, 1, \ldots$, and $v_\alpha(x) \geq 0$ for all $x \in \mathbb{X}$.

By (2.3) and Lemma 3.4, $v_{1,\alpha} \in L(\mathbb{X})$, since $v_{0,\alpha} = 0 \in L(\mathbb{X})$. By the same arguments, if $v_{n,\alpha} \in L(\mathbb{X})$ then $v_{n+1,\alpha} \in L(\mathbb{X})$. Thus $v_{n,\alpha} \in L(\mathbb{X})$ for all $n = 0, 1, \ldots$. By Lemma 3.3, for any $n = 1, 2, \ldots$, $x \in \mathbb{X}$, and $\lambda \in \mathbb{R}$, the set $\mathcal{D}_{\eta^\alpha_{v_{n,\alpha}}(x,\cdot)}(\lambda)$ is a compact subset of $\mathbb{A}$. By Bertsekas and Shreve [4, Proposition 9.17], $v_{n,\alpha} \uparrow v_\alpha$ as $n \to +\infty$. Since the limit of a monotone increasing sequence of lower semi-continuous functions is again a lower semi-continuous function, $v_\alpha \in L(\mathbb{X})$. Lemma 3.4, applied to equations (2.3) and (2.5), implies statements (ii) and (iv) respectively. Statement (iii) follows from (2.4) and statement (v) follows from (2.6).

Now let $c(x,a) \geq K$ for all $(x,a) \in \mathrm{Gr}(A)$ and for some $K > -\infty$. For $K \geq 0$, statements (i)–(v) are proved. For $K < 0$, consider the value functions $\tilde{c} = c - K \geq 0$. If the cost function c substituted with $\tilde{c}$, we substitute the notation v with $\tilde{v}$. Then $v^\pi_{n,\alpha} = \tilde{v}^\pi_{n,\alpha} + \frac{1-\alpha^n}{1-\alpha}K$, $n = 0, 1, \ldots$, for all policies π. Thus, $v_{n,\alpha} = \tilde{v}_{n,\alpha} + \frac{1-\alpha^n}{1-\alpha}K$, $n = 0, 1, \ldots$, and $v_\alpha = \tilde{v}_\alpha + \frac{K}{1-\alpha}$. Since statements (i)–(v) hold for the shifted costs $\tilde{c}$ and the value functions $\tilde{v}_{n,\alpha}$ and $\tilde{v}_\alpha$, they also hold for the initial cost function c and the value functions $v_{n,\alpha}$ and v_α. □

We remark that the conclusions of Theorem 4.1 and its proof remain correct when $\alpha = 1$ and the function c is nonnegative.

5 Average Costs Per Unit Time

In this section we show that Assumption ($\mathbf{W}^*$) and boudness assumption Assumption ($\underline{\mathbf{B}}$) on the function u_α, which is weaker boundness Assumption ($\mathbf{B}$) introduced by Schäl [24], lead to the validity of stationary average-cost optimal inequalities and the existence of stationary policies. Stronger results hold under Assumption ($\mathbf{B}$).

Assumption ($\underline{\mathbf{B}}$). (i) Assumption ($\mathbf{G}$) holds, and (ii) $\liminf\limits_{\alpha\uparrow 1} u_\alpha(x) < \infty$ for all $x \in \mathbb{X}$.

Assumption ($\underline{\mathbf{B}}$)(ii) is weaker than the assumption $\sup_{\alpha\in[0,1)} u_\alpha(x) < \infty$ for all $x \in \mathbb{X}$ considered in Schäl [24]. This assumption and Assumption ($\mathbf{G}$) were combined in Feinberg and Lewis [14] into the following assumption.

Assumption (B). (i) Assumption ($\mathbf{G}$) holds, and (ii) $\sup_{\alpha\in[0,1)} u_\alpha(x) < \infty$ for all $x \in \mathbb{X}$.

It seems natural to consider the assumption $\limsup_{\alpha\uparrow 1} u_\alpha(x) < \infty$ for all $x \in \mathbb{X}$, which is stronger than Assumption ($\underline{\mathbf{B}}$)(ii) and weaker than Assumption ($\mathbf{B}$)(ii). However, as the following lemma shows, under Assumption ($\mathbf{G}$) this assumption is equivalent to Assumption ($\mathbf{B}$)(ii).

Lemma 5.1. *Let the cost function c be bounded below and Assumption ($\mathbf{G}$) hold. Then for each $x \in \mathbb{X}$ the following two inequalities are equivalent:*

(i) $\sup_{\alpha\in[0,1)} u_\alpha(x) < \infty$,

(ii) $\limsup_{\alpha\uparrow 1} u_\alpha(x) < \infty$.

Proof. Obviously, (i)→(ii). Let us prove (ii)→(i). Let (ii) hold. Assume that (i) does not hold. Since $\sup_{\alpha\in[0,1)} u_\alpha(x) = \max\{\sup_{\alpha\in[0,\alpha^*)} u_\alpha(x), \sup_{\alpha\in[\alpha^*,1)} u_\alpha(x)\}$ for any $\alpha^* \in [0,1)$, there exists $\alpha^* \in [0,1)$ such that $\sup_{\alpha\in[0,\alpha^*)} u_\alpha(x) = \infty$.

Since the function u_α remains unchanged, if a finite constant is added to the cost function c, we assume without loss of generality that $c(x,a) \ge 0$ for all $(x,a) \in \mathrm{Gr}(A)$. Since $c \ge 0$, the functions $v_\alpha(x)$ and m_α are nonnegative nondecreasing functions in $\alpha \in [0,1)$. Since $v_\alpha(x) = u_\alpha(x) + m_\alpha \ge u_\alpha(x)$, we have $\sup_{\alpha\in[0,\alpha^*)} v_\alpha(x) = \infty$ and therefore $v_\alpha(x) = \infty$ for all $\alpha \in [\alpha^*, 1)$, because of the monotonicity of v_α in α. Thus, $\limsup_{\alpha\uparrow 1}(1-\alpha)v_\alpha(x) = \infty$. However, $\limsup_{\alpha\uparrow 1}(1-\alpha)v_\alpha(x) = \limsup_{\alpha\uparrow 1}(1-\alpha)(u_\alpha(x)+m_\alpha) \le \limsup_{\alpha\uparrow 1}(1-\alpha)u_\alpha(x) + \overline{w} < \infty$, where the last inequality follows from (ii) and (3.1). The obtained contradiction completes the proof. □

Until the end of this section we assume that Assumption ($\underline{\mathbf{B}}$) holds. Let us set

$$u(x) := \liminf_{\alpha\uparrow 1,\, y\to x} u_\alpha(y), \quad x \in \mathbb{X}, \tag{5.1}$$

where $\liminf_{\alpha\uparrow 1,\, y\to x} u_\alpha(y)$ is the least upper bound of the set of all $\lambda \in \mathbb{R}_+$ such that there exist $\beta \in [0,1)$ and a neighborhood $U(x)$ of x such that $\lambda \le \inf\{u_\alpha(y) :\ \alpha \in [\beta,1), y \in U(x) \cap \mathbb{X}\}$.

Also define the following nonnegative functions on $\mathbb{X}$:

$$U_\beta(x) = \inf_{\alpha\in[\beta,1)} u_\alpha(x), \quad \underline{u}_\beta(x) = \liminf_{y\to x} U_\beta(y), \qquad \beta \in [0,1),\ x \in \mathbb{X}. \tag{5.2}$$

Observe that all the three defined functions take finite values at $x \in \mathbb{X}$. Indeed,

$$\underline{u}_\beta(x) \le U_\beta(x) \le \sup_{\beta\in[0,1)} \inf_{\alpha\in[\beta,1)} u_\alpha(x) = \liminf_{\alpha\uparrow 1} u_\alpha(x) < \infty, \quad \beta \in [0,1),\ x \in \mathbb{X}, \tag{5.3}$$

where the first two inequalities follow from the definitions of $\underline{u}_\beta$ and U_β respectively, and the last inequality follows from Assumption ($\underline{\mathbf{B}}$). For $x \in \mathbb{X}$

$$\begin{aligned} u(x) &= \sup_{\beta\in[0,1),\ R>0}\left[\inf_{\alpha\in[\beta,1),\ y\in B_R(x)} u_\alpha(y)\right] = \sup_{\beta\in[0,1)}\ \sup_{R>0}\ \inf_{y\in B_R(x)}\ \inf_{\alpha\in[\beta,1)} u_\alpha(y) \\ &= \sup_{\beta\in[0,1)}\ \sup_{R>0}\ \inf_{y\in B_R(x)} U_\beta(y) = \sup_{\beta\in[0,1)}\ \liminf_{y\to x} U_\beta(y) = \sup_{\beta\in[0,1)} \underline{u}_\beta(x) < \infty, \end{aligned} \tag{5.4}$$

where $B_R(x) = \{y \in \mathbb{X} : \rho(y,x) < R\}$, the first equality is (5.1), the second equality follows from the properties of infinums, the third and the fifth equalities follow from (5.2), the fourth equality follows from the definition of $\limsup$, and the inequality follows from (5.3). In view of (5.2), the functions $U_\beta(x)$ and $\underline{u}_\beta(x)$ are nondecreasing in β. Therefore, in view of (5.4),

$$u(x) = \lim_{\beta\uparrow 1} \underline{u}_\beta(x), \qquad\qquad x \in \mathbb{X}. \tag{5.5}$$

We also set for u from (5.5)

$$A^*(x) := \left\{a \in A(x) : \overline{w} + u(x) \geq c(x,a) + \int_{\mathbb{X}} u(y)q(dy|x,a)\right\},\ x \in \mathbb{X}, \tag{5.6}$$

and let $A_*(x)$, $x \in \mathbb{X}$, be the sets defined in (3.11) for this function u; $A_*(x) \subseteq A^*(x)$.

Theorem 5.2. *Suppose Assumptions ($\mathbf{W}^*$) and ($\underline{\mathbf{B}}$) hold. There exist a stationary policy ϕ satisfying (3.3) with u defined in (5.1). Thus, equalities (3.4) hold for this policy ϕ. Furthermore, the following statements hold:*

(**a**) *the function $u : \mathbb{X} \to \mathbb{R}_+$, defined in (5.1), is lower semi-continuous;*

(**b**) *the nonempty sets $A^*(x)$, $x \in \mathbb{X}$, satisfy the following properties:*
($\mathbf{b_1}$) *the graph $\mathrm{Gr}(A^*) = \{(x,a) : x \in \mathbb{X}, a \in A^*(x)\}$ is a Borel subset of $\mathbb{X} \times \mathbb{A}$;*
($\mathbf{b_2}$) *for each $x \in \mathbb{X}$ the set $A^*(x)$ is compact;*

(**c**) *a stationary policy ϕ is optimal for average costs and satisfies (3.3) with u defined in (5.1), if $\phi(x) \in A^*(x)$ for all $x \in \mathbb{X}$;*

(**d**) *there exists a stationary policy ϕ with $\phi(x) \in A_*(x) \subseteq A^*(x)$ for all $x \in \mathbb{X}$;*

(**e**) *if, in addition, Assumption ($\mathbf{Wu}$) holds, then the function u, defined in (5.1), is inf-compact.*

Before the proof of Theorem 5.2, we establish some auxiliary facts.

Lemma 5.3. *Under Assumption ($\underline{\mathbf{B}}$), the functions $u, \underline{u}_\alpha : \mathbb{X} \to \mathbb{R}_+$, $\alpha \in [0,1)$, are lower semi-continuous on $\mathbb{X}$. If additionally Assumption ($\mathbf{W}^*$) holds, the functions $u_\alpha : \mathbb{X} \to \mathbb{R}_+$, $\alpha \in [0,1)$, are lower semi-continuous on $\mathbb{X}$. Under Assumptions ($\mathbf{Wu}$) and ($\underline{\mathbf{B}}$), the functions $u, u_\alpha, \underline{u}_\alpha : \mathbb{X} \to \mathbb{R}_+$, $\alpha \in [0,1)$, are inf-compact on $\mathbb{X}$.*

Proof. Since $\underline{u}_\alpha(x) \geq 0$, $\alpha \in [0,1)$ and $x \in \mathbb{X}$, the functions $\underline{u}_\alpha$, $\alpha \in [0,1)$, are lower semi-continuous; Feinberg and Lewis [14, Lemma 3.1]. Since supremum over any set of lower semi-continuous functions is a lower semi-continuous function, the function u is lower semi-continuous.

According to (3.1), $\overline{w} := \limsup_{\alpha\uparrow 1}(1-\alpha)m_\alpha = \inf_{\alpha\in(0,1)} \sup_{\alpha\in[\alpha,1)}(1-\alpha)m_\alpha < \infty$. Thus, there exists $\alpha_0 \in [0,1)$ such that

$$\lambda' := \sup_{\alpha\in[\alpha_0,1)} (1-\alpha)m_\alpha < \infty. \tag{5.7}$$

Let us assume that the function c is bounded below. As explained in the proof of Lemma 5.1, without loss of generality we can assume that $c \geq 0$. Then m_α is a nonnegative, nondecreasing function. Thus, $(1-\alpha)m_\alpha \leq (1-\alpha)m_{\alpha_0} \leq \lambda'/(1-\alpha_0)$, $\alpha \in [0,\alpha_0)$, and (5.7) implies that

$$\lambda^* = \sup_{\alpha\in[0,1)} (1-\alpha)m_\alpha < \infty. \tag{5.8}$$

According to Theorem 4.1(i, iv,v), under Assumption (**W***), the function $u_\alpha(x) = v_\alpha(x) - m_\alpha$ is lower semi-continuous, and a stationary policy ϕ_α is α-discount optimal if and only if for all $x \in \mathbb{X}$

$$v_\alpha(x) = \min_{a\in A(x)} \left\{ c(x,a) + \alpha \int_{\mathbb{X}} v_\alpha(y)q(dy|x,a) \right\} = c(x,\phi_\alpha(x)) + \alpha \int_{\mathbb{X}} v_\alpha(y)q(dy|x,\phi_\alpha(x)). \tag{5.9}$$

The first equality in (5.9) is equivalent to

$$(1-\alpha)m_\alpha + u_\alpha(x) = \min_{a\in A(x)} \left[c(x,a) + \alpha \int_{\mathbb{X}} u_\alpha(y)q(dy|x,a) \right], \quad x \in \mathbb{X}. \tag{5.10}$$

Let Assumption (**Wu**) hold. The function $u_\alpha(x) = v_\alpha(x) - m_\alpha$ is inf-compact by Theorem 4.1(vi). Consider an arbitrary $\lambda \in \mathbb{R}_+$. Since $u(x) \geq \underline{u}_{\alpha_1}(x) \geq \underline{u}_{\alpha_2}(x)$, $x \in \mathbb{X}$, for all $\alpha_1, \alpha_2 \in [0,1)$, $\alpha_1 \geq \alpha_2$, then $\mathcal{D}_u(\lambda) \subseteq \mathcal{D}_{\underline{u}_\alpha}(\lambda) \subseteq \mathcal{D}_{\underline{u}_0}(\lambda)$, $\alpha \in [0,1)$. Since the functions u and $\underline{u}_\alpha$ are lower semi-continuous, the sets $\mathcal{D}_u(\lambda)$ and $\mathcal{D}_{\underline{u}_\alpha}(\lambda)$ are closed, $\alpha \in [0,1)$. Therefore, if the set $\mathcal{D}_{\underline{u}_0}(\lambda)$ is compact then those sets are also compact and the functions u and $\underline{u}_\alpha$, $\alpha \in [0,1)$, are inf-compact.

Observe that (5.8) and (5.10) imply that $u_\alpha(x) \geq v_1(x) - \lambda^*$, $x \in X$, for all $\alpha \in [0,1)$. This implies $U_0(x) \geq v_1(x) - \lambda^*$, $x \in X$. Since $\underline{u}_0$ is the largest lower-semicontinuous function that is less than or equal to U_0 at all $x \in \mathbb{X}$, we have $\underline{u}_0(x) \geq v_1(x) - \lambda^*$, $x \in X$. Since the function $\underline{u}_0$ is lower semi-continuous, the set $\mathcal{D}_{\underline{u}_0}(\lambda)$ is closed. In addition, $\mathcal{D}_{\underline{u}_0}(\lambda) \subseteq \mathcal{D}_{v_1}(\lambda+\lambda^*)$, where the set $\mathcal{D}_{v_1}(\lambda+\lambda^*)$ is compact. Thus, the set $\mathcal{D}_{\underline{u}_0}(\lambda)$ is compact, and the functions u and $\underline{u}_\alpha$, $\alpha \in [0,1)$, are inf-compact. □

Corollary 5.4. *Under Assumption (*$\underline{\mathbf{B}}$*), for every sequence* $\alpha_n \uparrow 1$ *as* $n \to +\infty$ *and for every* $x \in \mathbb{X}$*,*

$$u(x) = \liminf_{n\to+\infty,\ y\to x} \underline{u}_{\alpha_n}(y).$$

Proof. Let $\alpha_n \uparrow 1$ as $n \to +\infty$, and $x \in \mathbb{X}$. Similar to (5.4)

$$\liminf_{n\to+\infty,\, y\to x} \underline{u}_{\alpha_n}(y) = \sup_{n=1,2,\dots} \sup_{R>0} \inf_{y\in B_R(x)} \inf_{m\ge n} \underline{u}_{\alpha_m}(y) = \sup_{n=1,2,\dots} \sup_{R>0} \inf_{y\in B_R(x)} \underline{u}_{\alpha_n}(y)$$
$$= \sup_{n=1,2\dots} \liminf_{y\to x} \underline{u}_{\alpha_n}(y) = \lim_{n\to\infty} \underline{u}_{\alpha_n}(x) = u(x),$$

where the second equality holds because the function $\underline{u}_\alpha(y)$ is nondecreasing in α, the fourth equality holds because it is lower semi-continuous, and the last equality follows from (5.5). □

Lemma 5.5. *Under Assumptions (*$\mathbf{W}^*$*) and (*$\underline{\mathbf{B}}$*), the following inequalities hold*

$$\overline{w} + u(x) \ge \min_{a\in A(x)} \left[c(x,a) + \int_{\mathbb{X}} u(y) q(dy|x,a) \right], \qquad x \in \mathbb{X}. \tag{5.11}$$

Proof. Let us fix an arbitrary $\varepsilon^* > 0$. Since $\overline{w} = \limsup_{\alpha\uparrow 1} (1-\alpha) m_\alpha$, there exists $\alpha_0 \in [0,1)$ such that

$$\overline{w} + \varepsilon^* > (1-\alpha) m_\alpha, \qquad \alpha \in [\alpha_0, 1). \tag{5.12}$$

Our next goal is to prove the inequality

$$\overline{w} + \varepsilon^* + u(x) \ge \min_{a\in A(x)} \left[c(x,a) + \alpha \int_{\mathbb{X}} \underline{u}_\alpha(y) q(dy|x,a) \right], \qquad x \in \mathbb{X},\ \alpha \in [\alpha_0, 1). \tag{5.13}$$

Indeed, by (5.10) and (5.12) for every $\alpha, \beta \in [\alpha_0, 1)$, such that $\alpha \le \beta$, and for every $x \in \mathbb{X}$

$$\overline{w} + \varepsilon^* + u_\beta(x) > (1-\beta) m_\beta + u_\beta(x) = \min_{a\in A(x)} \left[c(x,a) + \beta \int_{\mathbb{X}} u_\beta(y) q(dy|x,a) \right] \ge$$
$$\ge \min_{a\in A(x)} \left[c(x,a) + \alpha \int_{\mathbb{X}} U_\alpha(y) q(dy|x,a) \right].$$

As right-hand side does not depend on $\beta \in [\alpha, 1)$, we have for all $x \in \mathbb{X}$ and for all $\alpha \in [\alpha_0, 1)$

$$\overline{w} + \varepsilon^* + U_\alpha(x) = \inf_{\beta\in[\alpha,1)} [\overline{w} + \varepsilon^* + u_\beta(x)] \ge \min_{a\in A(x)} \left[c(x,a) + \alpha \int_{\mathbb{X}} U_\alpha(y) q(dy|x,a) \right] \ge$$
$$\ge \min_{a\in A(x)} \left[c(x,a) + \alpha \int_{\mathbb{X}} \underline{u}_\alpha(y) q(dy|x,a) \right] = \min_{a\in A(x)} \eta^\alpha_{\underline{u}_\alpha}(x,a).$$

By Lemma 3.4, the function $x \to \min_{a\in A(x)} \eta^\alpha_{\underline{u}_\alpha}(x,a)$ is lower semi-continuous on $\mathbb{X}$. Thus,

$$\liminf_{y\to x} \min_{a\in A(y)} \eta^\alpha_{\underline{u}_\alpha}(y,a) \ge \min_{a\in A(x)} \eta^\alpha_{\underline{u}_\alpha}(x,a), \qquad x \in \mathbb{X},\ \alpha \in [0,1).$$

and, as, by definition (5.2), $\underline{u}_\alpha(x) = \liminf_{y\to x} U_\alpha(y)$, we finally obtain

$$\overline{w} + \varepsilon^* + \underline{u}_\alpha(x) \ge \min_{a\in A(x)} \eta^\alpha_{\underline{u}_\alpha}(x,a), \qquad x \in \mathbb{X}, \alpha \in [\alpha_0, 1). \tag{5.14}$$

As, by (5.2), $u(x) = \sup_{\alpha\in[\alpha_0,1)} \underline{u}_\alpha(x)$ for all $x \in \mathbb{X}$, (5.14) yields (5.13).

To complete the proof of the lemma, we fix an arbitrary $x \in \mathbb{X}$. By Lemma 3.4, for any $\alpha \in [0, 1)$ there exists $a_\alpha \in A(x)$ such that $\min_{a\in A(x)} \eta^\alpha_{\underline{u}_\alpha}(x, a) = \eta^\alpha_{\underline{u}_\alpha}(x, a_\alpha)$. Since $\underline{u}_\alpha \geq 0$, for $\alpha \in [\alpha_0, 1)$ the inequality (5.13) can be continued as

$$\overline{w} + \varepsilon^* + u(x) \geq \eta^\alpha_{\underline{u}_\alpha}(x, a_\alpha) \geq c(x, a_\alpha). \tag{5.15}$$

Thus, for all $\alpha \in [\alpha_0, 1)$

$$a_\alpha \in \mathcal{D}_{\eta^\alpha_{\underline{u}_\alpha}(x,\cdot)}(\overline{w} + \varepsilon^* + u(x)) \subseteq \mathcal{D}_{c(x,\cdot)}(\overline{w} + \varepsilon^* + u(x)) \subseteq A(x).$$

By Lemma 3.3, the set $\mathcal{D}_{c(x,\cdot)}(\overline{w} + \varepsilon^* + u(x))$ is compact. Thus, for every sequence $\beta_n \uparrow 1$ of numbers from $[\alpha_0, 1)$ there is a subsequence $\{\alpha_n\}_{n\geq1}$ such that the sequence $\{a_{\alpha_n}\}_{n\geq1}$ converges and $a_* := \lim_{n\to\infty} a_{\alpha_n} \in A(x)$.

Consider a sequence $\alpha_n \uparrow 1$ such that $a_{\alpha_n} \to a_*$ for some $a_* \in A(x)$. Due to Lemmas 3.5 and Corollary 5.4,

$$\liminf_{n\to+\infty} \alpha_n \int_{\mathbb{X}} \underline{u}_{\alpha_n}(y)q(dy|x, a_n) \geq \int_{\mathbb{X}} u(y)q(dy|x, a_*). \tag{5.16}$$

Since the function c is lower semi-continuous, (5.15) and (5.16) imply

$$\overline{w} + \varepsilon^* + u(x) \geq \limsup_{n\to\infty} \eta^{\alpha_n}_{\underline{u}_{\alpha_n}}(x, a_{\alpha_n}) \geq c(x, a_*) + \int_{\mathbb{X}} u(y)q(dy|x, a_*) \geq \min_{a\in A(x)} \eta^1_u(x, a).$$

Since $\overline{w} + \varepsilon^* + u(x) \geq \min_{a\in A(x)} \eta^1_u(x, a)$ for any $\varepsilon^* > 0$, this is also true when $\varepsilon^* = 0$. □

Proof of Theorem 5.2. Lemma 5.3 contains statements (**a**) and (**e**). Since $\mathrm{Gr}(A^*) = \{(x, a) \in \mathrm{Gr}(A) :\ g(x, a) \geq 0\}$, where $g(x, a) = \overline{w} + u(x) - c(x, a) - \int_{\mathbb{X}} u(y)q(dy|x, a)$ is a Borel function, the set $\mathrm{Gr}(A^*)$ is Borel. The sets $A^*(x)$, $x \in \mathbb{X}$, are compact in view of Lemma 3.3(b). Thus, the statement (**b**) is proved. The Arsenin-Kunugui theorem implies the existence of a stationary policy ϕ such that $\phi(x) \in A^*(x)$ for all $x \in \mathbb{X}$. Statement (**e**) follows from Lemma 3.4 and the Arsenin-Kunugui theorem. The rest follows from Theorem 3.1. □

Theorem 5.6. *Suppose Assumptions ($\mathbf{W}^*$) and ($\mathbf{B}$) hold. Then all the conclusions of Theorem 5.2 hold and, in addition, for a stationary policy ϕ satisfying (3.3) with u defined in (5.1),*

$$w^\phi(x) = \underline{w} = \lim_{\alpha\uparrow1}(1 - \alpha)v_\alpha(x) = \lim_{N\to\infty} \frac{1}{N} v^\phi_N(x), \qquad x \in \mathbb{X}. \tag{5.17}$$

Proof. Consider a sequence $\{\alpha(n)\}_{n\geq1}$ such that $\alpha(n) \uparrow 1$ as $n \to +\infty$, and

$$\lim_{n\to+\infty}(1 - \alpha(n))m_{\alpha(n)} = \underline{w}.$$

Define the following nonnegative functions on $\mathbb{X}$:

$$\tilde{U}_n(x) = \inf_{m\geq n} u_{\alpha(m)}(x),\ \underline{\tilde{u}}_n(x) = \liminf_{y\to x} \tilde{U}_n(y), \quad n \geq 1,\ x \in \mathbb{X},$$

and

$$\tilde{u}(x) = \sup_{n\ge1} \underline{\tilde{u}}_n(x),\ x\in\mathbb{X}. \tag{5.18}$$

Observe that

$$\underline{\tilde{u}}_n(x) \le \tilde{U}_n(x) \le \limsup_{m\to+\infty} u_{\alpha(m)}(x) < \infty, \quad x\in\mathbb{X},\ n=1,2,\dots, \tag{5.19}$$

where the first two inequalities follow from the definitions of $\underline{\tilde{u}}_n$ and $\tilde{U}_n$ respectively, and the last inequality follows from Assumption (**B**). As follows from (5.18) and (5.19), $\tilde{u}(x) \le \limsup_{m\to+\infty} u_{\alpha(m)}(x) < +\infty$. According to Feinberg and Lewis [14, Lemma 3.1], the functions $\underline{\tilde{u}}_n$, $n\ge 1,$ are lower semi-continuous on $\mathbb{X}$. Therefore, their supremum $\tilde{u}$ is also lower semi-continuous. In addition,

$$\tilde{u}(x) = \sup_{n\ge1}\sup_{R>0}\inf_{y\in B_R(x)}\inf_{m\ge n} u_{\alpha_m}(y) = \liminf_{n\to+\infty,\, y\to x} u_{\alpha(n)}(y), \quad x\in\mathbb{X},$$

where the first equality follows from the definitions of $\tilde{U}_n$, $\underline{\tilde{u}}_n$, and $\tilde{u}$, and the second equality is the definition of the $\liminf$. Since $\tilde{U}_n(x)\uparrow$, we have $\underline{\tilde{u}}_n(x)\uparrow\tilde{u}(x)$ as $n\to\infty$ for all $x\in\mathbb{X}$.

We show next that for each $x\in\mathbb{X}$

$$\underline{w} + \tilde{u}(x) \ge \inf_{a\in A(x)}\left[c(x,a) + \int_{\mathbb{X}} \tilde{u}(y) q(dy|x,a)\right]. \tag{5.20}$$

Indeed let us fix any $\varepsilon^* > 0$. By the definition of $\underline{w}$, there exists a subsequence $\{\alpha(n_k)\}_{k\ge1} \subseteq \{\alpha(n)\}_{n\ge1}$ such that for $k=1,2,\dots$

$$\underline{w} + \varepsilon^* \ge (1-\alpha(n_k)) m_{\alpha(n_k)}.$$

Let $x\in\mathbb{X}$ be an arbitrary state. By Theorem 4.1 for each $k\ge1$ there exists $a_{n_k}\in A_{\alpha(n_k)}(x)$ such that

$$(1-\alpha(n_k)) m_{\alpha(n_k)} + u_{\alpha(n_k)}(x) = c(x,a_{n_k}) + \alpha(n_k)\int_{\mathbb{X}} u_{\alpha(n_k)}(y) q(dy|x,a_{n_k}).$$

Thus, similarly to the proof of Lemma 5.5, we get (5.20).

From Lemma 3.4 and the Arsenin-Kunugui theorem there exists a stationary policy $\tilde{\phi}\in\mathbb{F}$ such that for any $x\in\mathbb{X}$

$$\underline{w} + \tilde{u}(x) \ge c(x,\tilde{\phi}(x)) + \int_{\mathbb{X}} \tilde{u}(y) q(dy|x,\tilde{\phi}(x)). \tag{5.21}$$

Thus, by Schäl [24, Proposition 1.3] described in (3.2), for all $x\in\mathbb{X}$

$$\overline{w} = \underline{w} = w(x) = w^{\tilde{\phi}}(x) = \lim_{\alpha\uparrow1}(1-\alpha)v_\alpha(x) = w^*. \tag{5.22}$$

Let us choose any stationary policy ϕ such that inequalities (3.2) and (3.3) hold with the function u defined in (5.1). Since $\overline{w}=\underline{w}$, according to Theorem 5.2, such a stationary policy exists. Theorem 3.1 implies that the stationary policy ϕ satisfies (3.4), and Schäl [24, Proposition 1.3] (see (3.2)) implies that (5.22) holds with $\tilde{\phi}=\phi$.

In addition, (5.22) with $\tilde{\phi} = \phi$ implies that for all $x \in \mathbb{X}$

$$w^{\phi}(x) = \lim_{\alpha \uparrow 1}(1-\alpha)m_{\alpha} = \lim_{\alpha \uparrow 1}(1-\alpha)(v_{\alpha}(x) - u_{\alpha}(x)) = \lim_{\alpha \uparrow 1}(1-\alpha)v_{\alpha}(x),$$

where the last equality follows from Assumption ($\mathbf{B}$). Thus, for all $x \in \mathbb{X}$

$$\begin{aligned} w^{\phi}(x) &= \limsup_{n \to \infty} \frac{1}{n} v_n^{\phi}(x) \geq \limsup_{\alpha \uparrow 1}(1-\alpha)v_{\alpha}^{\phi}(x) \geq \liminf_{\alpha \uparrow 1}(1-\alpha)v_{\alpha}^{\phi}(x) \\ &\geq \lim_{\alpha \uparrow 1}(1-\alpha)v_{\alpha}(x) = w^{\phi}(x), \end{aligned}$$

where the first inequality follows from the Tauberian theorem (see Sennott [25, Section A.4] or [26, Proposition 5.7]), and the last inequality follows from $v_{\alpha}^{\phi}(x) \geq v_{\alpha}(x)$ and the existence of the limit. So, we have, the existence of $\lim_{\alpha \uparrow 1}(1-\alpha)v_{\alpha}^{\phi}(x)$. Thus, the Karamata Tauberian theorem (Sennott [25, Section A.4] or [26, Proposition 5.7]) implies $w^{\phi}(x) = \lim_{n \to \infty} \frac{1}{n} v_n^{\phi}(x)$. □

Corollary 5.7. *Under Assumptions ($\mathbf{W}^*$) and ($\underline{\mathbf{B}}$), the conclusions of Theorems 5.2 and 5.6 remain correct, if the function u is substituted with the function $\tilde{u}$ defined in (5.18).*

Proof. As shown in the proof of Theorem 5.6, there exists a stationary policy $\tilde{\phi}$ satisfying (5.21). The function $\tilde{u}$ is nonnegative, lower semi-continuous, and takes finite values. Thus, both [24, Proposition 1.3] (see (3.2)) and Theorem 3.1 can be applied to this function. The proof of statements (a)–(d) of Theorem 5.2 uses just these properties of u. Statement (e) follows from Lemma 5.3, whose proof remains unchanged if u is replaced with $\tilde{u}$. □

6 Approximation of Average Cost Optimal Strategies by α-discount Optimal Strategies

For a family of sets $\{\mathrm{Gr}(A_{\alpha})\}_{\alpha \in (0,1)}$, $x \in \mathbb{X}$, considered in Theorem 4.1, we pay our attention to its upper topological limit

$$\overline{\operatorname{Lim}}_{\alpha \uparrow 1} \mathrm{Gr}(A_{\alpha}) = \left\{ (x,a) \in \mathbb{X} \times \mathbb{A} \ : \ \begin{array}{l} \exists \alpha_n \uparrow 1,\ n \to +\infty,\ \exists (x_n, a_n) \in \mathrm{Gr}(A_{\alpha_n}),\ n \geq 1, \\ \text{such that } (x,a) = \lim_{n \to +\infty} (x_n, a_n) \end{array} \right\},$$

defined, for example, in Zgurovsky et al. [30, Chapter 1, p. 3]. Let us set

$$A^{app}(x) := \left\{ a \in A^*(x) \ : \ (x,a) \in \overline{\operatorname{Lim}}_{\alpha \uparrow 1} \mathrm{Gr}(A_{\alpha}) \right\}, \qquad x \in \mathbb{X}.$$

Theorem 6.1. *Under Assumptions ($\mathbf{W}^*$) and ($\underline{\mathbf{B}}$), the graph $\mathrm{Gr}(A^{app})$ is a Borel subset of $\mathrm{Gr}(A^*)$, and for each $x \in \mathbb{X}$ the set $A^{app}(x)$ is nonempty and compact. Furthermore, there exists a stationary policy ϕ^{app} such that $\phi^{app}(x) \in A^{app}(x)$ for all $x \in X$, and any such policy is average-cost optimal.*

Proof. Let us fix an arbitrary $x \in \mathbb{X}$. From (5.1) (the definition of u), there exists $\{y_n, \alpha_n\}_{n\geq 1} \subseteq \mathbb{X} \times (0,1)$ such that $y_n \to x$, $\alpha_n \uparrow 1$, $u_{\alpha_n}(y_n) \to u(x)$, $n \to +\infty$.

Let us choose an arbitrary $\varepsilon^* > 0$ and $b_n \in A_{\alpha_n}(y_n)$, $n \geq 1$. Since $\overline{w} = \limsup_{\alpha\uparrow 1}(1-\alpha)m_\alpha$, there exists $N \geq 1$ such that $u(x) + \frac{\varepsilon^*}{2} \geq u_{\alpha_n}(y_n)$ and $\overline{w} + \frac{\varepsilon^*}{2} \geq (1-\alpha_n)m_{\alpha_n}$ for all $n \geq N$.

By definition of the sets $A_\alpha(\cdot)$, for each $n \geq N$

$$(1-\alpha_n)m_{\alpha_n} + u_{\alpha_n}(y_n) = c(y_n, b_n) + \alpha_n \int_{\mathbb{X}} u_{\alpha_n}(y) q(dy|y_n, b_n) = \eta^{\alpha_n}_{u_{\alpha_n}}(y_n, b_n).$$

Thus, for all $n \geq N$

$$\overline{w} + \varepsilon^* + u(x) > \eta^{\alpha_n}_{u_{\alpha_n}}(y_n, b_n) \geq \eta^{\alpha_n}_{U_{\alpha_n}}(y_n, b_n) \geq \eta^{\alpha_n}_{\underline{u}_{\alpha_n}}(y_n, b_n) \geq c(y_n, b_n).$$

Therefore, because of Assumption ($\mathbf{W}^*$)(ii), the sequence $\{b_n\}_{n\geq 1}$ has a subsequence $\{b_{n_k}\}_{k\geq 1}$ such that $b_{n_k} \to a$, as $k \to +\infty$, for some $a \in A(x)$. Thus, $(x,a) \in \overline{\operatorname{Lim}}_{\alpha\uparrow 1} \operatorname{Gr}(A_\alpha)$.

Let us prove that $(x,a) \in \operatorname{Gr}(A^*)$. Indeed, as $\alpha_{n_k}\underline{u}_{\alpha_{n_k}}(\cdot) \uparrow u(\cdot)$, $k \to +\infty$, then due to Lemma 3.5 and Corollary 5.4,

$$\liminf_{k\to+\infty} \alpha_{n_k} \int_{\mathbb{X}} \underline{u}_{\alpha_{n_k}}(x) q(dy|y_{n_k}, b_{n_k}) \geq \int_{\mathbb{X}} u(x) q(dy|x,a).$$

Thus, by Lemma 3.4, $\overline{w} + \varepsilon^* + u(x) \geq \eta^1_u(x,a)$, and this is true for any $\varepsilon^* > 0$. This implies $\overline{w} + u(x) \geq \eta^1_u(x,a)$. This inequality means that $(x,a) \in \operatorname{Gr}(A^*)$ and $A^{app}(x) \neq \emptyset$, since $(x,a) \in \overline{\operatorname{Lim}}_{\alpha\uparrow 1} \operatorname{Gr}(A_\alpha)$. The set $A^{app}(x)$ is compact because of the closureness of $\overline{\operatorname{Lim}}_{\alpha\uparrow 1} \operatorname{Gr}(A_\alpha)$ (see Zgurovsky et al. [30, Chapter 1, p. 3]) and Theorem 5.2(b). The second statement of the theorem follows from the Arsenin-Kunugui theorem. □

Corollary 6.2. *Under Assumptions ($\mathbf{W}^*$) and ($\underline{\mathbf{B}}$), for any stationary average-cost optimal policy ϕ^{app}, such that $\phi^{app}(x) \in A^{app}(x)$ for all $x \in \mathbb{X}$, for every $x \in \mathbb{X}$ there exist $\alpha_n(x) \uparrow 1$ and $y_n(x) \to x$ as $n \to +\infty$ such that $a_n(x) \in A_{\alpha_n(x)}(y_n(x))$, $n \geq 1$, and $\phi^{app}(x) = \lim_{n\to+\infty} a_n(x)$.*

Proof. Following Theorem 6.1, consider a stationary average-cost optimal policy ϕ^{app} such that $\phi^{app}(x) \in A^{app}(x)$ for all $x \in X$. Furthermore, since $A^{app}(x) \subseteq A^*(x)$ for all $x \in \mathbb{X}$, any such a policy is optimal. Let us fix an arbitrary $x \in \mathbb{X}$. By definition of $A^{app}(x)$, we have that $(x, \phi^{app}(x)) \in \overline{\operatorname{Lim}}_{\alpha\uparrow 1} \operatorname{Gr}(A_\alpha)$. Then, there exist $\alpha_n(x) \uparrow 1$, $n \to +\infty$, and $(y_n(x), a_n(x)) \in \operatorname{Gr}(A_{\alpha_n})$, $n \geq 1$, such that $(x, \phi^{app}(x)) = \lim_{n\to+\infty}(y_n(x), a_n(x))$, i.e. $\phi^{app}(x) = \lim_{n\to+\infty} a_n(x)$, where $a_n(x) \in A_{\alpha_n(x)}(y_n(x))$, $n \geq 1$, $\alpha_n(x) \uparrow 1$ and $y_n(x) \to x$ as $n \to +\infty$. □

We remark that, if we replace in (5.6) the function u with $\tilde{u}$ defined in (5.18), Theorem 6.1 and Corollary 6.2 remain correct.

Let us set

$$X_\alpha := \{x \in \mathbb{X} : v_\alpha(x) = m_\alpha\}, \quad \alpha \in [0,1).$$

Under Assumptions $(\mathbf{G})$, $m_\alpha < \infty$. If Assumptions $(\mathbf{G})$ and $(\mathbf{Wu})$ hold then Theorem 4.1 implies that X_α is a compact set for each $\alpha \in [0,1)$. This fact is useful to establish the validity of Assumptions $(\mathbf{G})$; see Feinberg and Lewis [14, Lemma 5.1] and references therein.

Theorem 6.3. *Let Assumptions $(\mathbf{G})$ and $(\mathbf{Wu})$ hold. Then there exists a compact set $\mathcal{K} \subseteq \mathbb{X}$ such that $X_\alpha \subseteq \mathcal{K}$ for each $\alpha \in [0,1)$.*

Proof. From Assumption $(\mathbf{G})$ and Theorem 4.1 we have that for each $\alpha \in [0,1)$

$$\emptyset \neq X_\alpha = \{x \in \mathbb{X} : u_\alpha(x) = 0\} = \mathcal{D}_{u_\alpha}(0) \subseteq \mathcal{D}_{U_\alpha}(0) \subseteq \mathcal{D}_{\underline{u}_\alpha}(0) \subseteq \mathcal{D}_{\underline{u}_0}(0).$$

In virtue of Lemma 5.3, we have that $\underline{u}_0 : \mathbb{X} \to [0, +\infty)$ is inf-compact function on $\mathbb{X}$. Setting $\mathcal{K} = \mathcal{D}_{\underline{u}_0}(0)$, we obtain the statement of the theorem. □

7 Illustrative Example

The following example is from Hernández-Lerma [16]. Let

$$x_{n+1} = \gamma x_n + \beta a_n + \xi_n, \qquad n = 0, 1, ...,$$

and

$$c(x,a) = qx^2 + ra^2,$$

where (a) q and r are positive constants, γ and β are two constants satisfying $\gamma\beta > 0$, and (b) ξ_n are independent and identically distributed (iid) random variables with zero mean, finite variance, and continuous density.

This problem is solved in Hernández-Lerma [16], where a stationary average-cost optimal policy is computed. This problem corresponds to an MDP with $\mathbb{X} = \mathbb{A} = \mathbb{R}$ and with setwise continuous transition probabilities. However, if ξ_n do not have a density, the transition probability may not be setwise continuous, but they are weakly continuous; see Feinberg and Lewis [13, p. 48] for detail. If ξ_n are arbitrary iid random variables with zero mean and finite variance, this problem satisfies Assumption $(\mathbf{Wu})$ and, similarly to the case when there are densities, it satisfies Assumption $(\mathbf{B})$. Thus, Theorem 5.6 can be applied. The optimal policy provided in Hernández-Lerma [16] is also optimal when ξ_n may not have a density.

A Proof of Lemma 3.5

Proof. First, we prove the lemma for uniformly bounded above functions h_n. Let $h_n(s) \le K < \infty$ for all $n = 1, 2, ...$ and all $s \in S$. For $n = 1, 2, \ldots$ and $s \in S$, define

$$H_n(s) = \inf_{m \ge n} h_m(s) \quad \text{and} \quad \underline{h}_n(s) = \liminf_{s' \to s} H_n(s').$$

The functions $\underline{h}_n : S \to [0, +\infty)$, $n = 1, 2, \ldots$, are lower semi-continuous; see, for example, Feinberg and Lewis [14, Lemma 3.1]). In addition, for $s \in S$

$$\underline{h}_n(s) \downarrow \underline{h}(s) \qquad \text{as} \quad n \to \infty. \tag{A.1}$$

Weak convergence of $\{\mu_n\}_{n\geq 1}$ to μ is equivalent to

$$\liminf_{n\to+\infty} \mu_n(A) \geq \mu(A) \qquad \text{for all } A \in \mathcal{O}, \tag{A.2}$$

where $\mathcal{O}$ is the family of all open subsets of the space S; Billingsley [5, Theorem 2.1].

Fix an arbitrary $t > 0$. By (A.1), if $h(s) > t$ then $\underline{h}_n(s) > t$, $n = 1, 2, \ldots$, and

$$\{s \in S : h(s) > t\} = \bigcup_{n\geq 1} S_n, \tag{A.3}$$

where

$$S_n = \{s \in S : \underline{h}_n(s) > t\}, \quad n = 1, 2, \ldots,$$

are open sets, since the functions $\underline{h}_n : S \to \mathbb{R}_+$ are lower semi-continuous. In addition,

$$S_n \subseteq S_{n+1}, \qquad n = 1, 2, \ldots . \tag{A.4}$$

Thus,

$$\mu(\{s \in S : h(s) > t\}) = \lim_{n\to+\infty} \mu(S_n) \leq \lim_{n\to+\infty} \liminf_{m\to+\infty} \mu_m(S_n)$$
$$\leq \limsup_{n\to+\infty} \liminf_{m\to+\infty} \mu_m(S_m) = \liminf_{n\to+\infty} \mu_n(S_n) = \liminf_{n\to+\infty} \mu_n(\{s \in S : \underline{h}_n(s) > t\}),$$

where the first equality follows from (A.4) and (A.3), the first inequality follows from to (A.2), and the second inequality follows from (A.4).

Thus Serfozo [27, Lemma 2.1] yields

$$\int_S \underline{h}(s)\mu(ds) \leq \liminf_{n\to+\infty} \int_S \underline{h}_n(s)\mu_n(ds) \leq \liminf_{n\to+\infty} \int_S h_n(s)\mu_n(ds),$$

where the second inequality is fulfilled due to

$$\underline{h}_n(s) \leq H_n(s) \leq h_n(s), \qquad s \in S, \ n = 1, 2, \ldots .$$

Case 2. Consider a sequence $\{h_n\}_{n\geq 1}$ of measurable nonnegative $\overline{\mathbb{R}}$-valued functions on S. For $\lambda > 0$ set $h_n^\lambda(s) := \min\{h_n(s), \lambda\}$, $s \in S$, $n = 1, 2, \ldots$. Since the functions h_n^λ are uniformly bounded above,

$$\int_S \underline{h}^\lambda(s)\mu(ds) \leq \liminf_{n\to+\infty} \int_S h_n^\lambda(s)\mu_n(ds) \leq \liminf_{n\to+\infty} \int_S h_n(s)\mu_n(ds),$$

where $\underline{h}^\lambda(s) = \liminf_{n\to+\infty,\, s'\to s} h_n^\lambda(s')$, $\lambda > 0$, $s \in S$.

Then, using Fatou's lemma,

$$\int_S \underline{h}(s)\mu(ds) \leq \liminf_{\lambda\to+\infty} \int_S \underline{h}^\lambda(s)\mu(ds).$$

□

Acknowledgements. Research of the first author was partially supported by NSF grant CMMI-0900206. The authors thank Professor M.Z. Zgurovsky for initiating their research cooperation.

References

[1] Arapostathis, A., V. S. Borkar, E. Fernandez-Gaucherand, M. K. Ghosh and S. I. Marcus. 1993. Discrete time controlled Markov processes with average cost criterion: a survey, *SIAM J. Control Optim.* **31**(2) 282–344.

[2] Bather, J. 1973. Optimal decision procedures for finite Markov chains. Part I: Examples. *Adv. in Appl. Probab. 5*(2) 328–339.

[3] Berge, E. 1963. *Topological Spaces.* Macmillan, New York.

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

[5] Billingsley, P. 1968. *Convergence of Probability Measures*. Jonh Wiley, New York.

[6] Blackwell, D. 1962. Discrete dynamic programming. *Ann. Math. Statist.* **33**(2) 719–726.

[7] Cavazos-Cadena, R. 1991. A counterexample on the optimality equation in Markov decision chains with the average cost criterion. *Systems & Control Lett.* **16**(5) 387-392.

[8] Chen, R. C., E. A. Feinberg. 2010. Compactness of the space of non-randomized policies in countable-state sequential decision processes. *Math. Methods Oper. Res.* **71**(2) 307–323.

[9] Chitashvili, R. Y. 1975. A controlled finite Markov chain with an arbitrary set of decisions. *Theor. Probability Appl.* **20**(4) 839–847.

[10] Derman, C. 1962. On sequential decisions and Markov chains. *Management Sci. 9*(1) 16–24.

[11] Dynkin, E. B., A. A. Yushkevich. 1979. *Controlled Markov Processes*. Springer-Verlag, New York.

[12] Feinberg, E. A. 1980. An ϵ-optimal control of a finite Markov chain. *Theor. Probability Appl.* **25**(1) 70–81.

[13] Feinberg, E. A., M. E. Lewis. 2004. Optimality of four-threshold policies in inventory systems with customer returns and borrowing/storage options. *Probab. Engrg. Inform. Sci.* **19**(1) 45–71.

[14] Feinberg, E. A., M. E. Lewis. 2007. Optimality inequalities for average cost Markov decision processes and the stochastic cash balance problem. *Math. Oper. Res.* **32**(4) 769–783.

[15] Gubenko, L. G., E. S. Shtatland 1975. On controlled, discrete-time Markov decision processes. *Theory Probab. Math. Statist.* **7** 47–61.

[16] Hernández-Lerma, O. 1991. Averege optimality in dynamic programming on Borel spaces - Unbounded costs and controls. *Systems & Control Lett.* **17**(3) 237–242.

[17] Hernández-Lerma, O., J. B. Lassere. 1996. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Springer, New York.

[18] Hernández-Lerma, O., J. B. Lassere. 2000. Fatou's lemma and Lebesgue's convergence theorem for measures. *J. Appl. Math. Stoch. Anal.* **13**(2) 137–146.

[19] Kechris, A.S. 1995. *Classical Descriptive Set Theory.* Springer-Verlag, New York.

[20] Luque-Vásquez, F., O. Hernández-Lerma. 1995. A counterexample on the semicontinuity of minima. *Proc. Amer. Math. Soc.* **123**(10) 3175–3176.

[21] Ross, S. M. 1968. non-discounted denumerable Markovian decision model. *Ann. Math. Statist.* **39**(2) 412–424.

[22] Ross, S. M. 1968a. Arbitrary state Markovian decision processes. *Ann. Math. Statist.* **39**(6) 2118–2122.

[23] Ross, S. M. 1971. On the nonexistence of ϵ-optimal randomized stationary policies in average cost Markov decision models. *Ann. Math. Statist.* **42**(5) 1767–1768.

[24] Schäl, M. 1993. Average optimality in dynamic programming with general state space. *Math. Oper. Res*. **18**(1) 163–172.

[25] Sennott, L. I. 1999. *Stochastic Dynamic Programming and the Control of Queueing Systems*. John Wiley and Sons, New York.

[26] Sennott, L. I. 2002. Average reward optimization theory for denumerable state spaces. E. A. Feinberg, A. Shwartz, eds. *Handbook of Markov Decision Processes. Methods and Applications.* Kluwer, Boston, 153-172.

[27] Serfozo, R. 1982. Convergence of Lebesgue integrals with varying measures. *The Indian Journal of Statistics (Series A)*. **44** 380–402.

[28] Taylor, III, H. M.. 1965. Markovian sequential replacement processes. *Ann. Math. Statist.* **36**(6) 1677-1694.

[29] Viskov, O. V., A. N. Shiryaev. 1964. On controls which reduce to optimal stationary regimes, *Trudy Mat. Inst. Steklov*. **71** 3545 (in Russion; English translation: Report Number FTD-HT-67-69, National Technical Information Service, U.S. Department of Commerce).

[30] Zgurovsky, M. Z., V. S. Mel'nik, P. O. Kasyanov. 2011. *Evolution Inclusions and Variation Inequalities for Earth Data Processing I.* Springer, Berlin.