Commun. Optim. Theory 2026 (2026) 17 https://doi.org/10.23952/cot.2026.17

MATHRES

Communications in Optimization Theory

Available online at http://cot.mathres.org

AVERAGE-COST MDPS WITH INFINITE STATE AND ACTION SETS: NEW SUFFICIENT CONDITIONS FOR OPTIMALITY INEQUALITIES AND EQUATIONS

EUGENE A. FEINBERG[1,*], PAVLO O. KASYANOV[2], LILIIA S. PALIICHUK[2]

[1]Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11794-3600, USA

[2]Institute for Applied System Analysis, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Beresteiskyi Ave., 37, build 35, Kyiv, 03056 Ukraine

Dedicated to the memory of Uriel G. Rothblum

Abstract. This paper studies discrete-time average-cost infinite-horizon Markov decision processes (MDPs) with Borel state and action sets. It introduces new sufficient conditions for the validity of optimality inequalities and optimality equations for MDPs with weakly and setwise continuous transition probabilities. These inequalities and equations imply the existence of deterministic optimal policies.

Keywords. Average cost per unit time; Markov decision process; Optimality inequality; Optimality equation.

2020 Mathematics Subject Classification. 90C39, 90C40.

1. Introduction

This paper establishes sufficient conditions for the existence of deterministic optimal policies minimizing expected costs per unit time for infinite-horizon Markov Decision Processes with infinite state and action sets. Such policies exist for problems with finite state and actions sets [5, 9, 32], and deterministic policies were called stationary or randomized stationary in earlier publications. However, if either the state space or the action space is infinite, optimal policies may not exist. In particular, for countable-state MDPs with finite action sets, there are examples demonstrating nonexistence of optimal policies [10, 12, 26]. For finite-state MDPs there are examples when optimal policies do not exist when action sets are compact, costs do not depend on actions, and one-step transition probabilities depend continuously on actions [2, 8, 10].

For finite-state MDPs with compact action sets, deterministic optimal policies exist in the following two cases: (i) all sets of transition probabilities have finite sets of extreme points [11], (ii) the MDP is communicating [3], that is, any state can be reached from any state.

*Corresponding author.

E-mail address: eugene.feinberg@stonybrook.edu (E.A. Feinberg).

Received September 30, 2024; Accepted February 2, 2025.

For countable state MDPs, Sennott [29, 30] proved the validity of optimality inequalities under conditions generalizing the communicating condition, and these inequalities imply the existence of deterministic optimal policies. Cavazos-Cadena [7] provides an example of a countable-state MDP, for which optimality inequalities hold, and the optimality equation does not. Schäl [28] extended these results to MDPs with possibly uncountable state sets and compact action sets by considering Assumption B formulated below. In view of Hernández-Lerma and Lasserre [25, Theorem 5.4.6], Assumptio B is equivalent to the assumptions for the validity of the optimality inequality in Sennott [29, 30]. Feinberg et al. [17] and Feinberg and Kasyanov [14] extended Schäl's [28] results to MDPs with possibly noncompact action sets and introduced a weaker Assumption $\underline{\mathrm{B}}$, which also implies an optimality inequality in a weaker form, which also implies the existence of deterministic optimal policies. [14, Example 4.1] demonstrates that Assumption $\underline{\mathrm{B}}$ is indeed weaker than Assumption B. The results for noncompact action sets are important for inventory control [13, 21].

Schäl [28] studied MPDs with weakly and setwise continuous transition probabilities. Though weak continuity is more general than setwise continuity, MDPs with weakly continuous transition probabilities are not more general since continuity of costs and transition probabilities is assumed with respect to state-action pairs, while, for MDPs with setwise transition probabilities, continuity of costs and transition probabilities is assumed only with respect to actions. Feinberg et al. [17] studied MDPs with weakly continuous transition probabilities, and Feinberg and Kasyanov [14] studied MDPs with setwise continuous transition probabilities. Both models have important applications. For example, MDPs with weakly continuous transition probabilities are used for partially observable MDPs [19, 20]. An MDP with finite action sets and with arbitrary transition probabilities and arbitrary costs is an example of an MDP with setwise continuous transition probabilities. Hernández-Lerma [24] studied MDPs with setwise continuous transition probabilities with possibly noncompact action sets under Assumption B, but the optimality equation for discounted MDPs, which was used in the proofs, was formulated there without a proof, and the only proof known to the authors follows from the optimal selection theorem proved later in [14].

In this paper we introduce sufficient conditions, which are weaker than Assumption B, and which lead to the same or weaker conclusions on the validity of optimality inequalities and optimality equations as Assumption B. There is a significant literature on MDPs with average costs per unit time, which includes three surveys [1, 6, 30]. Recently Guo et al. [23] established new conditions for the existence of optimal policies for non-stationary MDPs.

2. Preliminaries

Let $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$, $\mathbb{N}^* := \{0, 1, \ldots\} = \mathbb{N} \cup \{0\}$. Consider a discrete-time MDP with a state space $\mathbb{X}$, an action space $\mathbb{A}$, one-step costs c, and transition probabilities q. Assume that $\mathbb{X}$ and $\mathbb{A}$ are Borel subsets of Polish (complete separable metric) spaces. Let $c(x,a) : \mathbb{X} \times \mathbb{A} \mapsto \overline{\mathbb{R}}$ be the one-step cost and $q(B|x,a)$ be the transition kernel representing the probability that the next state is in $B \in \mathscr{B}(\mathbb{X})$, given that the action a is chosen at the state x. The cost function c is assumed to be measurable and bounded below.

The decision process proceeds as follows: at each time epoch $t = 0, 1, \ldots$, the current state of the system, x, is observed. A decision-maker chooses an action a, the cost $c(x,a)$ is accrued, and the system moves to the next state according to $q(\cdot|x,a)$. Let $H_t = (\mathbb{X} \times \mathbb{A})^t \times \mathbb{X}$ be the set of histories for $t = 0,1,\dots$. A (randomized) decision rule at period $t = 0,1,\dots$ is a regular transition probability $\pi_t : H_t \mapsto \mathbb{A}$, that is, (i) $\pi_t(\,\cdot\,|h_t)$ is a probability distribution on $\mathbb{A}$, where $h_t = (x_0,a_0,x_1,\dots,a_{t-1},x_t)$, and (ii) for any measurable subset $B \subset \mathbb{A}$, the function $\pi_t(B|\cdot)$ is measurable on H_t. A policy π is a sequence $(\pi_0,\pi_1,\dots)$ of decision rules. Let Π be the set of all policies. A policy π is called non-randomized if each probability measure $\pi_t(\,\cdot\,|h_t)$ is concentrated at one point. A non-randomized policy is called deterministic if all decisions depend only on the current state.

The Ionescu Tulcea theorem 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-fields defined by Borel σ-fields of $\mathbb{X}$ and $\mathbb{A}$; see Bertsekas and Shreve [4, pp. 140–141] or Hernández-Lerma and Lasserre [25, p. 178]. Let $\mathbb{E}_x^\pi$ be an expectation w.r.t. P_x^π.

For a finite-horizon $N \in \mathbb{N}^* := \{1,2,\dots\}$, let us define the expected total discounted costs,

$$v_{N,\alpha}^\pi(x) := \mathbb{E}_x^\pi \sum_{t=0}^{N-1} \alpha^t c(x_t,a_t), \;\; x\in\mathbb{X}, \tag{2.1}$$

where $\alpha \in [0,1]$ is the discount factor. When $N = \infty$ and $\alpha \in [0,1)$, equation (2.1) defines an infinite-horizon expected total discounted cost denoted by $v_\alpha^\pi(x)$. We always assume that $\alpha \in [0,1)$ when $N = \infty$. We observe that the expectation in (2.1) is well-defined in the following two cases: (i) $N < \infty$, (ii) $N = \infty$ and $\alpha \in [0,1)$. This is true because the sum in (2.1) is a bounded below measurable function since the function c is bounded below and measurable.

Let $v_\alpha(x) := \inf_{\pi\in\Pi} v_\alpha^\pi(x)$, $x \in \mathbb{X}$. A policy π is called optimal for the discount factor α if $v_\alpha^\pi(x) = v_\alpha(x)$ for all $x \in \mathbb{X}$.

The *average cost per unit time* is defined as

$$w^\pi(x) := \limsup_{N\to\infty} \frac{1}{N} v_{N,1}^\pi(x), \;\; x\in\mathbb{X}.$$

Define the optimal value function $w(x) := \inf_{\pi\in\Pi} w^\pi(x)$, $x\in\mathbb{X}$. A policy π is called average-cost optimal if $w^\pi(x) = w(x)$ for all $x \in \mathbb{X}$.

We remark that in general action sets may depend on current states, and usually the state-dependent sets $A(x)$ are considered for all $x \in \mathbb{X}$. In our problem formulations $A(x) = \mathbb{A}$ for all $x\in\mathbb{X}$. This problem formulation is simpler than a formulation with the sets $A(x)$, and these two problem formulations are equivalent because we allow that $c(x,a) = +\infty$ for some $(x,a)\in\mathbb{X}\times\mathbb{A}$ and can set $A(x) = \{a\in\mathbb{A} : c(x,a) < +\infty\}$. For a formulation with the sets $A(x)$, one may define $c(x,a) = +\infty$ when $a \in \mathbb{A}\setminus A(x)$ and use the action set $\mathbb{A}$ instead of $A(x)$.

To establish the existence of average-cost optimal policies for problems with compact action sets, Schäl [28] considered two continuity Assumptions W and S for problems with weakly and setwise continuous transition probabilities, respectively. For setwise continuous transition probabilities, Hernández-Lerma [24] generalized Assumption S to Assumption S* to cover MDPs with possibly noncompact action sets. For the similar purpose, when transition probabilities are weakly continuous, Feinberg et al. [17] generalized Assumption **W** to Assumption W*.

We recall that a function $f : \mathbb{U} \mapsto \overline{\mathbb{R}}$ defined on a metric space $\mathbb{U}$ is called inf-compact (on $\mathbb{U}$), if for every $\lambda\in\mathbb{R}$ the level set $\{u\in\mathbb{U} : f(u) \le \lambda\}$ is compact. A measurable subset of a metric space is also a metric space with respect to the same metric. For $U \subset \mathbb{U}$, if the domain of f is narrowed to U, then this function is called the restriction of f to U.

Definition 2.1 (Feinberg et al. [18, Definition 1.1], Feinberg [13, Definition 2.1])**.** A function $f : \mathbb{X}\times\mathbb{A} \mapsto \overline{\mathbb{R}}$ is called $\mathbb{K}$-inf-compact, if for every nonempty compact subset $\mathscr{K}$ of $\mathbb{X}$ the restriction of f to $\mathscr{K}\times\mathbb{A}$ is an inf-compact function.

Assumption W* (Feinberg et al. [17, 20], Feinberg and Lewis [21], or Feinberg [13])**.**
(i) the function c is $\mathbb{K}$-inf-compact;
(ii) the transition probability $q(\cdot\,|x,a)$ is weakly continuous in $(x,a)\in\mathbb{X}\times\mathbb{A}$.

Assumption S* (Hernández-Lerma [24, Assumption 2.1] or Feinberg and Kasyanov [14])
(i) the function $c(x,a)$ is inf-compact in $a\in\mathbb{A}$ for each $x\in\mathbb{X}$;
(ii) the transition probability $q(\cdot\,|x,a)$ is setwise continuous in $a\in\mathbb{A}$ for each $x\in\mathbb{X}$.

Let

$$\begin{aligned} m_\alpha &:= \inf_{x\in\mathbb{X}} v_\alpha(x), \quad u_\alpha(x) := v_\alpha(x) - m_\alpha,\\ \underline{w} &:= \liminf_{\alpha\uparrow 1}(1-\alpha)m_\alpha, \quad \bar{w} := \limsup_{\alpha\uparrow 1}(1-\alpha)m_\alpha. \end{aligned} \tag{2.2}$$

The function u_α is called the discounted relative value function. If either Assumption W* or Assumption S* holds, let us consider the following assumption.

Assumption B. (Schäl [28]). (i) $w^* := \inf_{x\in\mathbb{X}} w(x) < +\infty$; and (ii) $\sup\limits_{\alpha\in[0,1)} u_\alpha(x) < +\infty$, $x\in\mathbb{X}$.

We recall that $\sup\limits_{\alpha\in[0,1)} u_\alpha(x) < +\infty$ if and only if $\limsup\limits_{\alpha\uparrow 1} u_\alpha(x) < +\infty$; [17, Lemma 5]. As follows from Schäl [28, Lemma 1.2(a)], Assumption **B**(i) implies that $m_\alpha < +\infty$ for all $\alpha\in[0,1)$. Thus, all the quantities in (2.2) are defined.

It is known [17, Theorem 1] that, if a deterministic policy ϕ satisfies the weakened average-cost optimality inequality (WACOI):

$$c(x,\phi(x)) + \int_{\mathbb{X}} u(y)q(dy\,|\,x,\phi(x)) \le \overline{w} + u(x), \qquad x\in\mathbb{X}, \tag{2.3}$$

for some nonnegative measurable function $u:\mathbb{X}\to\mathbb{R}$, then the deterministic policy ϕ is average-cost optimal, and

$$w(x) = w^\phi(x) = \lim_{\alpha\uparrow 1}(1-\alpha)v_\alpha(x) = \overline{w} = w^*, \quad x\in\mathbb{X}. \tag{2.4}$$

Let us consider the following assumption.

Assumption $\underline{\mathbf{B}}$. (Feinberg et al. [17]). (i) Assumption B(i) holds, and (ii) $\liminf\limits_{\alpha\uparrow 1} u_\alpha(x) < +\infty$ for all $x\in\mathbb{X}$.

Assumption $\underline{\mathrm{B}}$(ii) is weaker than Assumption B(ii); see [14, Example 4.1]. If a deterministic policy ϕ satisfies the average-cost optimality inequality (ACOI):

$$c(x,\phi(x)) + \int_{\mathbb{X}} u(y)q(dy\,|\,x,\phi(x)) \le \underline{w} + u(x), \qquad x\in\mathbb{X}, \tag{2.5}$$

for some nonnegative measurable function $u:\mathbb{X}\to\mathbb{R}$, which is a stronger version of (2.3) because $\underline{w}\le\overline{w}$ always holds, then, according to [28], the deterministic policy ϕ is average-cost optimal, and in addition to (2.4), it follows that $\underline{w}=\overline{w}$. A nonnegative measurable function

$u(x)$ satisfying inequality (2.5) with some deterministic policy ϕ is called an average-cost relative value function. “Boundedness” Assumption $\underline{\mathrm{B}}$ on the function u_α, which is weaker than boundedness Assumption B, and either Assumption W* or Assumption S* lead to the validity of WACOI (2.3) and the existence of optimal deterministic policies [17, Theorem 3] and [14, Theorem 3.3]. Stronger results, namely, the validity of ACOI (2.5) hold if Assumption B holds instead of Assumption $\underline{\mathrm{B}}$; see [17, Theorem 4] and [24].

We recall that $\alpha \in [0,1)$ for infinite-horizon problems, and everywhere in this paper, if we consider a discount factor α_n, we assume that $\alpha_n \in [0,1)$.

3. Main Results

We study MDPs either with weakly continuous transition probabilities satisfying Assumption W* or with setwise continuous transition probabilities satisfying Assumption S*. In either case, each of the Assumptions B or $\underline{\mathrm{B}}$ imply the validity of optimality inequalities and the existence of deterministic optimal policies [14, 17]. However, the results are stronger under Assumption B. In addition, under additional conditions, Assumption B implies the validity of the optimality equation [16]. We prove in Corollaries 3.4 and 3.12 that the results on the validity of optimality inequalities and optimality equations, that were established under Assumption B, hold under more general assumptions introduced in this section.

Theorems 3.2 and 3.3 state the validity of WACOI (2.3) under Assumption W* or Assumption S* and under essentially weakened version of Assumption **B**. For this purpose for an arbitrary fixed sequence $\alpha_n \uparrow 1$ we set:

$$\underline{w}_{\{\alpha_n\}} := \liminf_{n\to\infty}(1-\alpha_n)m_{\alpha_n}, \quad \overline{w}_{\{\alpha_n\}} := \limsup_{n\to\infty}(1-\alpha_n)m_{\alpha_n}. \tag{3.1}$$

According to Schäl [28, Lemma 1.2], Assumption B(i) implies

$$0 \le \underline{w} \le \underline{w}_{\{\alpha_n\}} \le \overline{w}_{\{\alpha_n\}} \le \overline{w} \le w^* < +\infty. \tag{3.2}$$

The following theorem formulates average-cost optimality inequality (3.3) in a different form than ACOI (2.5) and WACOI (2.3).

Theorem 3.1. *Let Assumption B(i) hold and $\{\alpha_n \uparrow 1\}_{n\in\mathbb{N}^*}$ be an arbitrary fixed sequence. If there exists a measurable function $u:\mathbb{X}\to[0,+\infty)$ and a deterministic policy ϕ such that*

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

then ϕ is average-cost optimal,

$$w(x) = w^\phi(x) = \limsup_{n\to\infty}(1-\alpha_n)v_{\alpha_n}(x) = \limsup_{\alpha\uparrow 1}(1-\alpha)v_\alpha(x) = \overline{w} = \overline{w}_{\{\alpha_n\}} = w^*, \tag{3.4}$$

for each $x\in\mathbb{X}$, and WACOI (2.3) hold for the same policy ϕ and function u as in (3.3).

We remark that for $\alpha_n\uparrow 1$, if $(1-\alpha_n)m_{\alpha_n}\to\overline{w}$, then (3.3) coincides with WACOI (2.3), which is already stated in Theorem 3.1, and, if $(1-\alpha_n)m_{\alpha_n}\to\underline{w}$, then (3.3) coincides with ACOI (2.5), which is an additional property; see Corollary 3.4.

Proof of Theorem 3.1. Similarly to Feinberg et al. [17, Theorem 1], since u is nonnegative, by iterating (3.3) we obtain

$$v^\phi_{n,1}(x) \le n\overline{w}_{\{\alpha_n\}} + u(x), \quad n\ge 1,\ x\in\mathbb{X}.$$

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

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

where the first and the second inequalities follow from the definitions of w and w^* respectively. Since $w^* \le \overline{w}_{\{\alpha_n\}}$, inequalities (3.2) imply that for all $\pi\in\Pi$ and for all $x\in\mathbb{X}$

$$w^* = \overline{w} = \overline{w}_{\{\alpha_n\}} \le \limsup_{n\to\infty}(1-\alpha_n)v_{\alpha_n}(x) \le \limsup_{\alpha\uparrow 1}(1-\alpha)v_\alpha(x) \le \limsup_{\alpha\uparrow 1}(1-\alpha)v_\alpha^\pi(x) \le w^\pi(x),$$

where the last inequality follows from the Tauberian theorem. Finally, we obtain that

$$\begin{aligned} w^* &= \overline{w} = \overline{w}_{\{\alpha_n\}} \le \limsup_{n\to\infty}(1-\alpha_n)v_{\alpha_n}(x) \le \limsup_{\alpha\uparrow 1}(1-\alpha)v_\alpha(x) \\ &\le \inf_{\pi\in\Pi} w^\pi(x) = w(x) \le w^\phi(x) \le \overline{w}_{\{\alpha_n\}}, \end{aligned} \tag{3.6}$$

for each $x\in\mathbb{X}$, where the last inequality follows from (3.5). Thus, all the inequalities in (3.6) are equalities. □

For a sequence $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$ of discount factors, consider the following assumption.

Assumption $\underline{\mathrm{B}}_{\{\alpha_n\}}$. (i) Assumption B(i) holds, and (ii) for a sequence $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$ of discount factors, the inequality $\liminf\limits_{n\to\infty} u_{\alpha_n}(x) < +\infty$ holds for all $x\in\mathbb{X}$.

Assumption B is equivalent to the statement that Assumption $\underline{\mathrm{B}}_{\{\alpha_n\}}$ holds for an arbitrary sequence $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$ because Assumption B obviously implies this statement and, conversely, by contradiction, if Assumption B does not hold, then $\limsup_{n\to\infty} u_{\alpha_n}(x)\to+\infty$ for some sequence $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$, and Assumption $\underline{\mathrm{B}}_{\{\beta_n\}}$ does not hold for the subsequence $\{\beta_n\}_{n\in\mathbb{N}^*}$ of the sequence $\{\alpha_n\}_{n\in\mathbb{N}^*}$ such that $\lim_{n\to\infty} u_{\beta_n}(x)\to+\infty$. The existence of a sequence $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$ satisfying Assumption $\underline{\mathrm{B}}_{\{\alpha_n\}}$ implies Assumption $\underline{\mathrm{B}}$. Moreover, we note that $\liminf_{n\to\infty,y\to x} u_{\alpha_n}(y)$ is the least upper bound of the set of all $\lambda\in\mathbb{R}_+$ such that there exist $m\in\mathbb{N}$ and a neighborhood $V(x)$ of x satisfying

$$\lambda \le \inf\{u_{\alpha_n}(y) : n\ge m,\, y\in V(x)\}.$$

This holds because

$$\liminf_{n\to\infty,y\to x} u_{\alpha_n}(y) = \sup_{V(x),m}\ \inf_{y\in V(x),n\ge m} u_{\alpha_n}(y),$$

where the supremum is taken over all neighborhoods $V(x)$ of x and $m = 1,2,\ldots$.

Theorem 3.2. *Let Assumptions W* holds and let Assumption $\underline{\mathrm{B}}_{\{\alpha_n\}}$ hold for a sequence $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$. Let*

$$u(x) := \liminf_{n\to\infty,y\to x} u_{\alpha_n}(y), \qquad x\in\mathbb{X}. \tag{3.7}$$

Then there exists a deterministic policy ϕ satisfying WACOI (2.3) with the function u defined in (3.7). Therefore, ϕ is a deterministic average-cost optimal policy. In addition, the function u is lower semi-continuous, and equalities (3.4) hold.

According to definition (3.7) the function u depends on the sequence $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$. We do not write this dependence explicitly. A natural question, which we do not study in this paper, is under which conditions functions u defined in (3.7) coincide for two sequences of discount factors converging to 1.

Note that the following properties take place in Example 4.1 from [14]: (a) Assumption B does not hold; (b) Assumptions W* and $\underline{\mathrm{B}}_{\{\alpha_n\}}$ for some sequence $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$ hold; (c) $\underline{w} =$

$\overline{w}$ and, therefore, there exists a deterministic policy ϕ satisfying ACOI (3.3), (2.5) with the function u defined in (3.7).

Let Assumption $\underline{\mathrm{B}}_{\{\alpha_n\}}$ hold for a sequence $\{\alpha_n \uparrow 1\}_{n\in\mathbb{N}^*}$. We define the following nonnegative functions on $\mathbb{X}$:

$$U_m(x) = \inf_{n\ge m} u_{\alpha_n}(x), \quad \underline{u}_m(x) = \liminf_{y\to x} U_m(y), \qquad m = 1,2,\dots,\ x\in\mathbb{X}. \tag{3.8}$$

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

$$\underline{u}_m(x) \le U_m(x) \le \sup_{m=1,2,\dots} \inf_{n\ge m} u_{\alpha_n}(x) = \liminf_{n\to\infty} u_{\alpha_n}(x) < +\infty, \quad m = 1,2,\dots,\ x\in\mathbb{X}, \tag{3.9}$$

where the first two inequalities follow from the definitions of $\underline{u}_m$ and U_m respectively, and the last inequality follows from Assumption $\underline{\mathrm{B}}_{\{\alpha_n\}}$. For $x\in\mathbb{X}$

$$\begin{aligned} u(x) &= \sup_{m=1,2,\dots,\ R>0} \left[\inf_{n\ge m,\ y\in B_R(x)} u_{\alpha_n}(y) \right] = \sup_{m=1,2,\dots} \sup_{R>0} \inf_{y\in B_R(x)} \inf_{n\ge m} u_{\alpha_n}(y) \\ &= \sup_{m=1,2,\dots} \sup_{R>0} \inf_{y\in B_R(x)} U_m(y) = \sup_{m=1,2,\dots} \liminf_{y\to x} U_m(y) = \sup_{m=1,2,\dots} \underline{u}_m(x) < +\infty, \end{aligned} \tag{3.10}$$

where $B_R(x) = \{y\in\mathbb{X} : \rho(y,x) < R\}$, the first equality is (3.7), the second equality follows from the properties of infima, the third and the fifth equalities follow from (3.8), the fourth equality follows from the definition of $\limsup$, and the last inequality follows from (3.9). In view of (3.8), the functions $U_m(x)$ and $\underline{u}_m(x)$ are nondecreasing in m. Therefore, in view of (3.10),

$$u(x) = \lim_{m\to\infty} \underline{u}_m(x), \qquad x\in\mathbb{X}. \tag{3.11}$$

We also set for u from (3.11)

$$A^*(x) := \left\{ a\in A(x) \,:\, c(x,a) + \int_{\mathbb{X}} u(y) q(dy|x,a) \le \overline{w}_{\{\alpha_n\}} + u(x) \right\},\ x\in\mathbb{X}. \tag{3.12}$$

Proof of Theorem 3.2. By replacing $\alpha\in[0,1)$ with $\alpha\in\{\alpha_n\}_n$ in Lemma 6 from Feinberg et al. [17], we obtain that the functions u, u_{α_m}, and $\underline{u}_m : \mathbb{X}\to\mathbb{R}_+$, $m = 1,2,\dots,$ are lower semi-continuous on $\mathbb{X}$.

Let us prove that u satisfies (3.3). For this purpose, let us fix an arbitrary $\varepsilon^* > 0$. Since $\overline{w}_{\{\alpha_n\}} = \limsup_{n\to\infty} (1-\alpha_n) m_{\alpha_n}$, there exists $n_0\in[0,1)$ such that

$$\overline{w}_{\{\alpha_n\}} + \varepsilon^* > (1-\alpha_n) m_{\alpha_n}, \quad n = n_0, n_0+1, \dots. \tag{3.13}$$

Our next goal is to prove the inequality

$$\overline{w}_{\{\alpha_n\}} + \varepsilon^* + u(x) \ge \min_{a\in A(x)} \left[c(x,a) + \alpha_m \int_{\mathbb{X}} \underline{u}_m(y) q(dy|x,a) \right], \quad x\in\mathbb{X},\ m\ge n_0. \tag{3.14}$$

Indeed, by

$$(1-\alpha) m_\alpha + u_\alpha(x) = \min_{a\in A(x)} \left[c(x,a) + \alpha \int_{\mathbb{X}} u_\alpha(y) q(dy|x,a) \right], \quad x\in\mathbb{X}. \tag{3.15}$$

and (3.13) for every $n, m\ge n_0$, such that $n\ge m$, and for every $x\in\mathbb{X}$

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

$$\geq \inf_{a\in A(x)}\left[c(x,a)+\alpha_m\int_{\mathbb{X}}U_m(y)q(dy|x,a)\right]$$

because $\alpha_n \geq \alpha_m$ since $\alpha_n \uparrow 1$. As the right-hand side does not depend on $n \geq m$, we have for all $x \in \mathbb{X}$ and for all $\alpha \in [\alpha_0, 1)$

$$\overline{w}_{\{\alpha_n\}}+\varepsilon^*+U_m(x)=\inf_{n\geq m}\left[\overline{w}_{\{\alpha_n\}}+\varepsilon^*+u_{\alpha_n}(x)\right]\geq \inf_{a\in A(x)}\left[c(x,a)+\alpha_m\int_{\mathbb{X}}U_m(y)q(dy|x,a)\right]$$

$$\geq \min_{a\in A(x)}\left[c(x,a)+\alpha_m\int_{\mathbb{X}}\underline{u}_m(y)q(dy|x,a)\right]=\min_{a\in A(x)}\eta_{\underline{u}_m}^{\alpha_m}(x,a),$$

where

$$\eta_{\underline{u}_m}^{\alpha_m}(x,a):=c(x,a)+\alpha_m\int_{\mathbb{X}}\underline{u}_m(y)q(dy|x,a).$$

By Feinberg et al. [17, Lemma 3], the function $x\mapsto \min_{a\in A(x)}\eta_{\underline{u}_m}^{\alpha_m}(x,a)$ is lower semi-continuous on $\mathbb{X}$. Thus,

$$\liminf_{y\to x}\min_{a\in A(y)}\eta_{\underline{u}_m}^{\alpha_m}(y,a)\geq \min_{a\in A(x)}\eta_{\underline{u}_m}^{\alpha_m}(x,a),\qquad x\in\mathbb{X},\ m=1,2,\ldots.$$

and, as, by definition (3.8), $\underline{u}_m(x)=\liminf_{y\to x}U_m(y)$, we finally obtain

$$\overline{w}_{\{\alpha_n\}}+\varepsilon^*+\underline{u}_m(x)\geq \min_{a\in A(x)}\eta_{\underline{u}_m}^{\alpha_m}(x,a),\quad x\in\mathbb{X}, m\geq n_0. \tag{3.16}$$

Since $u(x)=\sup_{m\geq n_0}\underline{u}_m(x)$ for all $x\in\mathbb{X}$, (3.16) yields (3.14).

To complete the proof of the theorem, we fix an arbitrary $x\in\mathbb{X}$. By Feinberg et al. [17, Lemma 3], for any $m=1,2,\ldots$ there exists $a_m\in A(x)$ such that $\min_{a\in A(x)}\eta_{\underline{u}_m}^{\alpha_m}(x,a)=\eta_{\underline{u}_m}^{\alpha_m}(x,a_m)$. Since $\underline{u}_m\geq 0$, for $m\geq n_0$ the inequality (3.14) can be continued as

$$\overline{w}_{\{\alpha_n\}}+\varepsilon^*+u(x)\geq \eta_{\underline{u}_m}^{\alpha_m}(x,a_m)\geq c(x,a_m). \tag{3.17}$$

Thus, for all $m\geq n_0$

$$a_m\in\mathscr{D}_{\eta_{\underline{u}_m}^{m}(x,\cdot)}(\overline{w}_{\{\alpha_n\}}+\varepsilon^*+u(x))\subset\mathscr{D}_{c(x,\cdot)}(\overline{w}_{\{\alpha_n\}}+\varepsilon^*+u(x))\subset A(x),$$

where $\mathscr{D}_f(\lambda)=\{y\in U\,:\,f(y)\leq\lambda\}$ is the level set. By Feinberg et al. [17, Lemma 2], the set $\mathscr{D}_{c(x,\cdot)}(\overline{w}_{\{\alpha_n\}}+\varepsilon^*+u(x))$ is compact. Thus, there is a subsequence $\{\alpha_m\}_m\subset\{\alpha_n\}_{n\geq 1}$ such that the sequence $\{a_{\alpha_m}\}_m$ converges and $a_*:=\lim_m a_{\alpha_m}\in A(x)$.

Consider a subsequence $\{\alpha_m\}_m\subset\{\alpha_n\}_{n\geq 1}$ such that $a_{\alpha_m}\to a_*$ for some $a_*\in A(x)$. Due to Fatou's lemma for weakly converging probabilities [15],

$$\liminf_{m\to+\infty}\alpha_m\int_{\mathbb{X}}\underline{u}_m(y)q(dy|x,a_m)\geq\int_{\mathbb{X}}u(y)q(dy|x,a_*). \tag{3.18}$$

Since the function c is lower semi-continuous, (3.17) and (3.18) imply

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

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

The Arsenin-Kunugui theorem implies the existence of a deterministic policy ϕ such that $\phi(x)\in A^*(x)$ for all $x\in\mathbb{X}$, where the sets $A^*(x)$ are defined in (3.12). $\square$

Analyzing the proofs of Hernández-Lerma [24, Section 4, Theorem] and Feinberg and Kasyanov [14, Theorem 3.3], we obtain the following theorem.

Theorem 3.3. *Let Assumptions S* hold, and let Assumption* $\underline{\mathrm{B}}_{\{\alpha_n\}}$ *hold for a sequence* $\{\alpha_n \uparrow 1\}_{n\in\mathbb{N}^*}$*. Let*

$$u(x) := \liminf_{n\to\infty} u_{\alpha_n}(x), \qquad x \in \mathbb{X}. \tag{3.19}$$

Then there exists a deterministic policy ϕ *satisfying WACOI* (2.3) *with the function* u *defined in* (3.19)*. Therefore,* ϕ *is a deterministic average-cost optimal policy.*

Proof. The proof of optimality inequality (3.3) follows the original proof of Hernández-Lerma [24, Section 4, Theorem] with minor modifications; see, also, Feinberg and Kasyanov [14, Theorem 3.3]). Inequality (3.3) implies WACOI (2.3) in view of Theorem 3.1. □

The following corollary provides under Assumptions W* or S* sufficient conditions for the validity of ACOI (2.5) under weaker conditions than Assumption B.

Corollary 3.4. *Let Assumption W* or S* hold, and let there exist a sequence* $\{\alpha_n \uparrow 1\}_{n\in\mathbb{N}^*}$ *such that Assumption* $\underline{\mathrm{B}}_{\{\alpha_n\}}$ *holds, and* $(1-\alpha_n)m_{\alpha_n} \to \underline{w}$ *as* $n\to\infty$*. Then ACOI* (2.5) *holds.*

Proof. Theorems 3.2 and 3.3 imply that WACOI (2.3) holds. In addition, since $\underline{w} = \overline{w}$, we see that ACOI (2.5) holds. □

Recall the following definitions.

Definition 3.5 (Semi-equicontinuity [16])**.** A sequence $\{f_n\}_{n\in\mathbb{N}^*}$ of real-valued functions on a metric space $(\mathbb{S},\rho)$ is called *lower semi-equicontinuous at the point* $s\in\mathbb{S}$ if for each $\varepsilon>0$ there exists $\delta>0$ such that

$$f_n(s') > f_n(s) - \varepsilon \qquad \text{for all } n\in\mathbb{N}^* \text{ if } \rho(s,s') < \delta.$$

The sequence $\{f_n\}_{n\in\mathbb{N}^*}$ is called *lower semi-equicontinuous* (*on* $\mathbb{S}$) if it is lower semi - equicontinuous at all $s\in\mathbb{S}$. A sequence $\{f_n\}_{n\in\mathbb{N}^*}$ of real-valued functions on a metric space $\mathbb{S}$ is called *upper semi-equicontinuous at the point* $s\in\mathbb{S}$ (*on* $\mathbb{S}$) if the sequence $\{-f_n\}_{n\in\mathbb{N}^*}$ is lower semi-equicontinuous at the point $s\in\mathbb{S}$ (on $\mathbb{S}$).

Definition 3.6 (Equicontinuity)**.** A sequence $\{f_n\}_{n\in\mathbb{N}^*}$ of real-valued functions on a metric space $\mathbb{S}$ is called *equicontinuous at the point* $s\in\mathbb{S}$ (*on* $\mathbb{S}$) if this sequence is both lower and upper semi-equicontinuous at the point $s\in\mathbb{S}$ (on $\mathbb{S}$).

The following corollary from Theorem 3.2 provides a sufficient condition for the validity of ACOI (2.5) with a relative value function u defined in (3.19).

Corollary 3.7. *Let Assumptions W* hold, Assumption* $\underline{\mathrm{B}}_{\{\alpha_n\}}$ *hold for a sequence* $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$*, and the sequence of functions* $\{u_{\alpha_n}\}_{n\in\mathbb{N}^*}$ *be lower semi-equicontinuous. Then the conclusions of Theorem* 3.2 *hold for the function* u *defined in* (3.19) *for this sequence* $\{\alpha_n\}_{n\in\mathbb{N}^*}$*.*

Proof. Since the sequence of functions $\{u_{\alpha_n}\}_{n\in\mathbb{N}^*}$ is lower semi-equicontinuous, the functions u defined in (3.7) and in (3.19) coincide in view of [16, Theorem 3.1(i)]. □

Consider the following version of the equicontinuity condition (EC) on the discounted relative value functions from [16].

Assumption $\mathrm{EC}_{\{\alpha_n\}}$**.** The sequence of discount factors $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$ satisfies the following properties:

(i) the sequence of functions $\{u_{\alpha_n}\}_{n\in\mathbb{N}^*}$ is equicontinuous;
(ii) there exists a nonnegative measurable function $U(x)$, $x\in\mathbb{X}$, such that $U(x)\ge u_{\alpha_n}(x)$, $n\in\mathbb{N}^*$, and $\int_{\mathbb{X}}U(y)q(dy|x,a)<+\infty$ for all $x\in\mathbb{X}$ and $a\in\mathbb{A}$.

Under each of the Assumptions W* or [25, Assumption 4.2.1], which is stronger than Assumption S*, and under Assumptions $\underline{\mathrm{B}}_{\{\alpha_n\}}$ and $\mathrm{EC}_{\{\alpha_n\}}$, there exists a deterministic policy ϕ satisfying the average-cost optimality equation (ACOE)

$$\begin{aligned} w^*+u(x) =& c(x,\phi(x))+\int_{\mathbb{X}}u(y)q(dy|x,\phi(x)) \\ =& \min_{a\in\mathbb{A}}\left[c(x,a)+\int_{\mathbb{X}}u(y)q(dy|x,a)\right],\ x\in\mathbb{X}, \end{aligned} \tag{3.20}$$

with u defined in (3.7) for the sequence $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$, and the function u is continuous; see Feinberg and Liang [22, Theorem 3.2] for Assumption W* and Hernández-Lerma and Lasserre [25, Theorem 5.5.4]. We remark that the quantity w^* in (3.20) can be replaced with any other quantity in (3.4). In addition, since the first equation in (3.20) implies inequality (3.3), every deterministic policy ϕ satisfying (3.20) is average-cost optimal. Observe that in these cases the function u is continuous (see [22, Theorem 3.2] for Assumption W* and [25, Theorem 5.5.4]), while under conditions of Theorems 3.2 and 3.3 the corresponding functions u may not be continuous; see Examples 7.1 and 7.2 from [16]. Below we provide more general conditions for the validity of the ACOEs. In particular, under these conditions the relative value functions u may not be continuous.

Now, we introduce Assumption $\mathrm{LEC}_{\{\alpha_n\}}$, which is weaker than Assumption $\mathrm{EC}_{\{\alpha_n\}}$. Indeed, Assumption $\mathrm{EC}_{\{\alpha_n\}}$(i) is obviously stronger than $\mathrm{LEC}_{\{\alpha_n\}}$(i). In view of the Ascoli theorem (see [25, p. 96] or [27, p. 179]), $\mathrm{EC}_{\{\alpha_n\}}$(i) and the first claim in $\mathrm{EC}_{\{\alpha_n\}}$(ii) imply $\mathrm{LEC}_{\{\alpha_n\}}$(ii). The second claim in $\mathrm{EC}_{\{\alpha_n\}}$(ii) implies $\mathrm{LEC}_{\{\alpha_n\}}$(iii). It is shown in Theorem 3.8 that the ACOEs hold under Assumptions W*, $\underline{\mathrm{B}}_{\{\alpha_n\}}$, and $\mathrm{LEC}_{\{\alpha_n\}}$.

Assumption $\mathrm{LEC}_{\{\alpha_n\}}$**.** The sequence of discount factors $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$ satisfies the following properties:

(i) the sequence of functions $\{u_{\alpha_n}\}_{n\in\mathbb{N}^*}$ is lower semi-equicontinuous;
(ii) $\lim_{n\to\infty}u_{\alpha_n}(x)$ exists for each $x\in\mathbb{X}$;
(iii) for each $x\in\mathbb{X}$ and $a\in\mathbb{A}$ the sequence $\{u_{\alpha_n}\}_{n\in\mathbb{N}^*}$ is asymptotically uniformly integrable with respect to the probability measure $q(\,\cdot\,|x,a)$, that is,

$$\lim_{K\to+\infty}\limsup_{n\in\mathbb{N}^*}\int_{\mathbb{X}}u_{\alpha_n}(z)q(dz|x,a)=0,$$

which, according to [15, Theorem 2.2], is equivalent to the existence of $N\in\mathbb{N}^*$ such that the sequence $\{u_{\alpha_N},u_{\alpha_{N+1}},\dots\}$ is uniformly integrable with respect to the probability measure $q(\,\cdot\,|x,a)$.

Theorem 3.8. *Let Assumption W* hold, and let Assumption* $\underline{\mathrm{B}}_{\{\alpha_n\}}$ *hold for a sequence* $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$. *If Assumption* $\mathrm{LEC}_{\{\alpha_n\}}$ *is satisfied for the sequence* $\{\alpha_n\}_{n\in\mathbb{N}^*}$, *then there exists a deterministic policy* ϕ *such that ACOE* (3.20) *hold with the function* $u(x)$ *defined in* (3.19).

Proof. Since Assumptions W* and $\underline{\mathrm{B}}_{\{\alpha_n\}}$ hold, and $\{u_{\alpha_n}\}_{n\in\mathbb{N}^*}$ is lower semi-equicontinuous, then Corollary 3.7 implies the existence of a deterministic policy ϕ satisfying (3.3) with u

defined in (3.19)

$$c(x,\phi(x))+\int_{\mathbb{X}}u(y)q(dy|x,\phi(x))\le w^*+u(x),\quad x\in\mathbb{X}. \tag{3.21}$$

To prove the ACOE, it remains to prove the opposite inequality to (3.21). According to Feinberg et al. [17, Theorem 2(iv)], for each $n\in\mathbb{N}^*$ and $x\in\mathbb{X}$ the discounted-cost optimality equation is $v_{\alpha_n}(x)=\min_{a\in\mathbb{A}}[c(x,a)+\alpha_n\int_{\mathbb{X}}v_{\alpha_n}(y)q(dy|x,a)]$, which, by subtracting $\alpha_n m_{\alpha_n}$ from both sides and by replacing α_n with 1, implies that for all $a\in\mathbb{A}$

$$(1-\alpha_n)m_{\alpha_n}+u_{\alpha_n}(x)\le c(x,a)+\int_{\mathbb{X}}u_{\alpha_n}(y)q(dy|x,a),\qquad x\in\mathbb{X}. \tag{3.22}$$

Let $n\to\infty$. In view of (3.4), Assumptions $\mathrm{LEC}_{\{\alpha_n\}}$(ii, iii), and Fatou's lemma [31, p. 211], (3.22) imply that for all $a\in\mathbb{A}$

$$w^*+u(x)\le c(x,a)+\int_{\mathbb{X}}u(y)q(dy|x,a),\qquad x\in\mathbb{X}. \tag{3.23}$$

We remark that the integral in (3.22) converges to the integral in (3.23) since the sequence $\{u_{\alpha_n}\}_{n\in\mathbb{N}^*}$ converges pointwise to u and is u.i.; see [16, Theorem 2.1]. Then, (3.23) implies

$$w^*+u(x)\le\min_{a\in\mathbb{A}}[c(x,a)+\int_{\mathbb{X}}u(y)q(dy|x,a)]\le c(x,\phi(x))+\int_{\mathbb{X}}u(y)q(dy|x,\phi(x)),\quad x\in\mathbb{X}. \tag{3.24}$$

Thus, (3.21) and (3.24) imply (3.20). □

In the following example, Assumptions W*, $\underline{\mathrm{B}}_{\{\alpha_n\}}$, and $\mathrm{LEC}_{\{\alpha_n\}}$ hold. Hence the ACOEs hold. However, Assumption $\mathrm{EC}_{\{\alpha_n\}}$ does not hold. Therefore, Assumption $\mathrm{LEC}_{\{\alpha_n\}}$ is more general than Assumption $\mathrm{EC}_{\{\alpha_n\}}$.

Example 3.9. ([16, Example 7.1]) Consider $\mathbb{X}=[0,1]$ equipped with the Euclidean metric and $\mathbb{A}=\{a^{(1)}\}$. The transition probabilities are $q(0|x,a^{(1)})=1$ for all $x\in\mathbb{X}$. The cost function is $c(x,a^{(1)})=\mathbf{I}\{x\ne 0\}$, $x\in\mathbb{X}$. Then the discounted-cost value is $v_\alpha(x)=u_\alpha(x)=\mathbf{I}\{x\ne0\}$, $\alpha\in[0,1)$ and $x\in\mathbb{X}$, and the average-cost value is $w^*=w(x)=0$, $x\in\mathbb{X}$. It is straightforward to see that Assumptions W* and $\underline{\mathrm{B}}_{\{\alpha_n\}}$ hold. In addition, since the function $u(x)=\mathbf{I}\{x\ne0\}$ is lower semi-continuous, but it is not continuous, the sequence of functions $\{u_{\alpha_n}\}_{n\in\mathbb{N}^*}$ is lower semi-equicontinuous, but it is not equicontinuous for each sequence $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$. Therefore, Assumption $\mathrm{LEC}_{\{\alpha_n\}}$ holds since $0\le u_{\alpha_n}(x)\le 1$, $x\in\mathbb{X}$, and Assumption $\mathrm{EC}_{\{\alpha_n\}}$ does not hold. The (3.20) holds with $w^*=0$, $u(x)=\mathbf{I}\{x\ne0\}$, and $\phi(x)=a^{(1)}$, $x\in\mathbb{X}$. □

The following theorem states the validity of the ACOE under Assumptions S*, $\underline{\mathrm{B}}_{\{\alpha_n\}}$, and $\mathrm{LEC}_{\{\alpha_n\}}$(ii,iii).

Theorem 3.10. *Let Assumption S* hold, and let Assumption $\underline{\mathrm{B}}_{\{\alpha_n\}}$ hold for a sequence $\{\alpha_n\uparrow 1\}_{n\in\mathbb{N}^*}$. If Assumptions $\mathrm{LEC}_{\{\alpha_n\}}$(ii,iii) are satisfied for the sequence $\{\alpha_n\}_{n\in\mathbb{N}^*}$, then there exists a deterministic policy ϕ such that ACOE (3.20) holds with the function $u(x)$ defined in (3.19).*

Proof. According to Theorem 3.3, if Assumptions S* and $\underline{\mathrm{B}}_{\{\alpha_n\}}$ hold, then we have that: (i) equalities in (3.4) hold; (ii) there exists a deterministic policy ϕ satisfying ACOI (3.21) with the function u defined in (3.19); and (iii) for each $n\in\mathbb{N}^*$ and $x\in\mathbb{X}$ the discounted-cost optimality equation is $v_{\alpha_n}(x)=\min_{a\in\mathbb{A}}[c(x,a)+\alpha_n\int_{\mathbb{X}}v_{\alpha_n}(y)q(dy|x,a)]$. Therefore, the same arguments as in the proof of Theorem 3.8 starting from (3.22) imply the validity of (3.20) with u defined in (3.19). $\square$

Observe that the MDP described in Example 3.9 also satisfies Assumptions S*, $\underline{\mathrm{B}}_{\{\alpha_n\}}$, and $\mathrm{LEC}_{\{\alpha_n\}}$(ii,iii). We provide Example 3.11, in which Assumptions S*, $\underline{\mathrm{B}}_{\{\alpha_n\}}$, and $\mathrm{LEC}_{\{\alpha_n\}}$(ii,iii) hold. Hence, the ACOEs hold. However, Assumptions W*, $\mathrm{LEC}_{\{\alpha_n\}}$(i), and $\mathrm{EC}_{\{\alpha_n\}}$ do not hold.

Example 3.11. ([16, Example 7.2]) Let $\mathbb{X}=[0,1]$ and $\mathbb{A}=\{a^{(1)}\}$. The transition probabilities are $q(0|x,a^{(1)})=1$ for all $x\in\mathbb{X}$. The cost function is $c(x,a^{(1)})=D(x)$, where D is the Dirichlet function defined as

$$D(x)=\begin{cases}0, & \text{if } x \text{ is rational},\\ 1, & \text{if } x \text{ is irrational},\end{cases} \qquad x\in\mathbb{X}.$$

Since there is only one available action, Assumption S* holds. The discounted-cost value is $v_\alpha(x)=u_\alpha(x)=D(x)=u(x)$, $\alpha\in[0,1)$ and $x\in\mathbb{X}$, and the average-cost value is $w^*=w(x)=0$, $x\in\mathbb{X}$. Then Assumptions $\underline{\mathbf{B}}_{\{\alpha_n\}}$ and $\mathrm{LEC}_{\{\alpha_n\}}$(ii,iii) hold. Hence, the ACOEs (3.20) hold with $w^*=0$, $u(x)=D(x)$, and $\phi(x)=a^{(1)}$, $x\in\mathbb{X}$. Thus, the average-cost relative function u is not lower semi-continuous. However, since the function $c(x,a^{(1)})=D(x)$ is not lower semi-continuous, Assumption W* does not hold. Since the function $u(x)=u_\alpha(x)=D(x)$ is not lower semi-continuous, Assumptions $\mathrm{LEC}_{\{\alpha_n\}}$(i) and $\mathrm{EC}_{\{\alpha_n\}}$ do not hold either. $\square$

We recall that, in view of Theorem 3.1, $w^*=\overline{w}$ under assumptions of this theorem. The following theorem provides sufficient conditions for $w^*=\overline{w}=\underline{w}$. While ACOE (3.20) is a stronger fact than WACOE (2.3), Corollary 3.12 provides sufficient conditions for the optimality equality which is stronger than ACOE 2.5.

Corollary 3.12. *Let assumptions of either Theorem 3.8 or Theorem 3.10 hold. If, in addition* $\lim_{n\to\infty}(1-\alpha_n)m_{\alpha_n}=\underline{w}$, *then* $w^*=\overline{w}=\underline{w}$ *and ACOE (3.20) holds with* w^* *substituted with* $\underline{w}$.

Proof. The proof follows from the arguments provided after the formulation of Theorem 3.1. $\square$

We remark that [14, Example 4.1] satisfies the assumptions of Corollary 3.12, but it does not satisfy Assumption B.

Acknowledgments

This research was partially supported by SUNY System Administration under SUNY Research Seed Grant Award 231087 and by the U.S. Office of Naval Research (ONR) under Grants N000142412608 and N000142412646.

References

[1] A. Arapostathis, V. Borkar, E. Fernandez-Gaucherand, M. Ghosh, and S. Marcus Discrete-time controlled Markov processes with average cost criterion: A survey, SIAM J. Control Optim. 31 (1993) 282–244.

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

[3] J. Bather, Optimal decision procedures for finite Markov chains, Part II: Eommunicating systems, Adv. Appl. Prob. 5 (1973) 521–540.

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

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

[6] V. S. Borkar, Convex analytic methods in Markov decision processes, In: E. A. Feinberg, A. Shwartz (eds.), Handbook of Markov Decision Processes: Methods and Applications, pp. 347–375, Kluwer, Boston, 2002.
[7] R. Cavazos-Cadena, A counterexample on the optimality equation in Markov decision chains with the average cost criterion, Systems and Control Lett. 16 (1991) 387–392.
[8] R. Ya. Chitashvili, A controlled finite Markov chain with an arbitrary set of decisions, Theory Prob. Appl. 20 (1975) 839–847.
[9] C, Derman, On sequential decisions and Markov chains, Management Sci. 9 (1962) 16–24.
[10] E. B. Dynkin and A. A. Yushkevich, Controlled Markov Processes, Springer-Verlag, New York, 1979.
[11] E. A. Feinberg, On controlled finite state Markov processes with compact control sets, Theory Prob. Appl. 20 (1975), 856–862.
[12] E. A. Feinberg, On controlled finite state Markov processes with compact control sets, Theory Prob. Appl. 25 (1980) 70–81.
[13] E. A. Feinberg, Optimality Conditions for Inventory Control, In: A. Gupta & A. Capponi (eds.), Tutorials in operations research, Optimization challenges in complex, networked, and risky systems, pp. 14–44, Cantonsville, MD, INFORMS, 2016.
[14] E. A. Feinberg and P. O. Kasyanov, MDPs with setwise continuous transition probabilities, Oper. Res. Lett. 49 (2021) 734-740.
[15] E. A. Feinberg, P. O. Kasyanov, and Y. Liang, Fatou's lemma for weakly converging measures under the uniform integrability condition, Theory Prob. Appl. 65 (2020) 270-291.
[16] E. A. Feinberg, P. O. Kasyanov, and Y. Liang, Fatou's lemma in its classical form and Lebesgue's convergence theorems for varying measures with applications to Markov decision processes, Theory Prob. Appl. 64 (2020) 615-631.
[17] E. A. Feinberg, P. O. Kasyanov, and N. V. Zadoianchuk, Average cost Markov decision processes with weakly continuous transition probabilities, Math. Oper. Res. 37 (2012) 591–607.
[18] E. A. Feinberg, P. O. Kasyanov, and N. V. Zadoianchuk, Berge's theorem for noncompact image sets, J. Math. Anal. Appl. 397 (2013) 255–259.
[19] E. A. Feinberg, P. O. Kasyanov, and M. Z. Zgurovsky, Convergence of probability measures and Markov decision models with incomplete information, Proc. Steklov Inst. Math. 287 (2014) 96–117.
[20] E. A. Feinberg, P. O. Kasyanov, and M. Z. Zgurovsky, Partially observable total-cost Markov decision processes with weakly continuous transition probabilities, Math. Oper. Res. 41 (2016), 656–681.
[21] E. A. Feinberg and M. E. Lewis, On the convergence of optimal actions for Markov decision processes and the optimality of (*s,S*) inventory policies, Naval Res. Logist. 65 (2018) 619–637.
[22] E. A. Feinberg and Y. Liang, On the optimality equation for average cost Markov decision processes and its validity for inventory control, Ann. Oper. Res. 317 (2022) 569–586.
[23] X Guo, Y. Huang, and Y. Zhang, On average optimality for non-stationary Markov decision processes in Borel spaces, Math. Oper. Res. 50 (2005) 2433–3282.
[24] O. Hernández-Lerma, Average optimality in dynamic programming on Borel spaces — Unbounded costs and controls, Sys. Control Lett. 17 (1991) 237–242.
[25] O. Hernández-Lerma and J. B. Lasserre, Discrete-Time Markov Control Processes: Basic Optimality Criteria, Springer-Verlag, New York, 1996.
[26] S. Ross, Introduction to Stochastic Dynamic Programming, Academic Press, New York, 1983.
[27] H. L. Royden, Real Analysis, 2nd ed., Macmillan, New York, 1968.
[28] M. Schäl, Average optimality in dynamic programming with general state space, Math. Oper. Res. 18 (1993) 163–172.
[29] L. I. Sennott, Stochastic Dynamic Programming and the Control of Queueing Systems, John Wiley & Sons, New York, 1998.
[30] L. I. Sennott, Average reward optimization theory for denumerable state systems, In: E. A. Feinberg, A. Shwartz (eds.), Handbook of Markov Decision Processes: Methods and Applications, pp. 153–173, Kluwer, Boston, 2002.
[31] A. N. Shiryaev, Probability, 2nd ed., Springer-Verlag, New York, 1996.
[32] O. V. Viskov and A. N. Shiryaev, On controls leading to optimal stationary regimes, Proceedings of the Steklov Institute of Mathematics, 71 (1964) 35–45. (In Russian)