2017 IEEE 56th Annual Conference on Decision and Control (CDC)
December 12-15, 2017, Melbourne, Australia

On the Average-Cost Optimality Equations and Convergence of Discounted-Cost Relative Value Functions for Inventory Control Problems with Quasiconvex Cost Functions

Eugene A. Feinberg and Yan Liang

Abstract— **Average-cost optimality inequalities imply the existence of stationary optimal policies for Markov Decision Processes with average costs per unit time, and these inequalities hold under broad natural conditions. Additional conditions are required for the validity of the average-cost optimality equations. Recently Feinberg and Liang [10, Theorem 3.2] showed that the equicontinuity of value functions for discounted costs is sufficient additional condition for the validity of average-cost optimality equations for problems with weakly continuous transition probabilities and with possibly unbounded one-step costs, and this condition holds for setup-cost inventory control problems with backorders and convex holding/backlog costs. This paper studies periodic-review setup-cost inventory control problem with backorders and with quasiconvex cost functions and general demands. It is shown that such problems satisfy the equicontinuity condition. Therefore, optimality inequalities hold in the form of equalities with a continuous average-cost relative value function for this problem. In addition, this implies that average-cost optimal (s, S) policies for the inventory control problem can be derived from the average-cost optimality equation. With the additional assumption on the monotonicity of the cost function, we establish the convergence of discounted-cost optimal ordering threshold s_α and convergence of discounted-cost relative value functions, when the discount factor converges to 1, to the corresponding optimal threshold and optimal relative value function for the average-cost problem.**

Keywords: *Markov processes, Stochastic systems, Optimal control.*

I. Introduction

For Markov Decision Processes (MDPs) with average costs per unit time, the existence of stationary optimal policies follows from the validity of the average-cost optimality inequality (ACOI). Feinberg et al. [5] established broad sufficient conditions for the validity of ACOIs for MDPs with weakly continuous transition probabilities and possibly noncompact action sets and unbounded one-step costs. In particular, these and even stronger conditions hold for the periodic-review setup-cost inventory control problem with backorders; see Feinberg [4] or Feinberg and Lewis [8]. Previously, Schäl [15] established sufficient conditions for the validity of ACOIs for MDPs with compact action sets and possibly unbounded costs. Cavazos-Cadena [2] provided an example in which the ACOI holds but the average-cost optimality equation (ACOE) does not. Feinberg and Liang [10, Theorem 3.2] provided sufficient conditions for the validity of ACOEs for MDPs with infinite state spaces, weakly continuous transition probabilities and possibly noncompact action sets and unbounded one-step costs. This paper shows that the setup-cost inventory control problems with quasiconvex cost functions and general demands satisfy these conditions and establishes the validity of the ACOEs for the setup-cost inventory control problems with backorders.

Sufficient conditions for the validity of the ACOEs for discrete-time MDPs with general state and action spaces with setwise continuous transition probabilities are described in Hernández-Lerma and Lasserre [11, Section 5.5]. Jaśkiewicz and Nowak [13] considered MDPs with Borel state space, compact action sets, weakly continuous transition probabilities and unbounded costs. The geometric ergodicity of transition probabilities is assumed in Jaśkiewicz and Nowak [13] to ensure the validity of the ACOEs. Costa and Dufour [3] studied the validity of ACOEs for MDPs with Borel state and action spaces, weakly continuous transition probabilities, which are positive Harris recurrent, and with possibly noncompact action sets and unbounded costs. Neither the geometric ergodicity nor positive Harris recurrent conditions hold for the periodic-review inventory control problem.

For the inventory control problems with quasiconvex cost functions, Veinott [17] proved the optimality of (s, S) policies for finite-horizon problems. Zheng [18] provided a simple proof of the optimality of (s, S) policies for inventory control problems with discrete demands under discounted-cost and average-cost criteria. Zheng [18] also established the validity of ACOEs for discrete demand. Huh et al. [12] considered the inventory control model with quasiconvex cost functions with bounded derivatives and with some additional linear restriction on costs and proved the optimality of (s, S) policy under the average-cost criterion.

Section II of this paper describes the general MDPs framework. In particular, it states Assumptions **W*** and **B** from Feinberg et al. [5], which guarantee the validity of the ACOIs. Section III provides the sufficient conditions from Feinberg and Liang [10, Theorem 3.2] for the validity of ACOEs for MDPs with weakly continuous transition probabilities. Section IV discusses several assumptions on the convexity or quasiconvexity of cost functions and their relations. By verifying conditions provided in Sections II and III, it is shown in Section IV, that ACOEs hold for the setup-cost inventory control problems with quasiconvex cost functions. The paper also establishes equicontinuity

This research was partially supported by NSF grants CMMI-1335296 and CMMI-1636193.

E. A. Feinberg and Y. Liang are with the Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11794, USA `eugene.feinberg@stonybrook.edu`, `yan.liang@stonybrook.edu`

of discounted-cost relative value functions and continuity of the average-cost relative value function and shows that an optimal (s,S) policy can be derived from the ACOEs. It also shows that at the level s there are at least two optimal decisions: do not order and order up to the level S. Section V establishes the convergence of discounted-cost optimal ordering thresholds s_α, when the discount factor α converges to 1. Section VI establishes the convergence of discounted-cost relative value functions, when the discount factor converges to 1.

II. General MDPs Framework

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 $c(x,a) : \mathbb{X}\times\mathbb{A} \to \overline{\mathbb{R}} = \mathbb{R}\cup\{+\infty\}$ be the one-step cost and $q(B|x,a)$ be the transition kernel representing the probability that the next state is in $B \in \mathcal{B}(\mathbb{X})$, given that the action a is chosen in the state x.

We recalled that a function $f : U \to \mathbb{R}\cup\{+\infty\}$ for a metric space $\mathbb{U}$, where U is a subset of a metric space $\mathbb{U}$, is called inf-compact, if for every $\lambda\in\mathbb{R}$ the level set $\{u\in\mathbb{U} : f(u)\le\lambda\}$ is compact.

Definition 1. (Feinberg et al. [6, Definition 1.1], Feinberg [4, Definition 2.1]) A function $f : \mathbb{X}\times\mathbb{A}\to\overline{\mathbb{R}}$ is called $\mathbb{K}$-inf-compact, if for every nonempty compact subset K of $\mathbb{X}$ the function $f : K\times\mathbb{A}\to\overline{\mathbb{R}}$ is inf-compact.

Let the one-step cost function c and transition probability q satisfy the following condition.

Assumption W* (Feinberg et al. [5], Feinberg and Lewis [8], or Feinberg [4]).

(i) c is $\mathbb{K}$-inf-compact and bounded below, and

(ii) the transition probability $q(\cdot|x,a)$ is weakly continuous in $(x,a)\in\mathbb{X}\times\mathbb{A}$, that is, for every bounded continuous function $f:\mathbb{X}\to\mathbb{R}$, the function $\tilde{f}(x,a) := \int_{\mathbb{X}} f(y)q(dy|x,a)$ is continuous on $\mathbb{X}\times\mathbb{A}$.

The decision process proceeds as follows: at each time epoch $t = 0,1,\dots$, 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 $t = 0,1,\dots$. Let Π be the set of all policies. A (randomized) decision rule at period $t = 0,1,\dots$ is a regular transition probability $\pi_t : H_t \to \mathbb{A}$, 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. Moreover, 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 stationary 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-field defined by Borel σ-field of $\mathbb{X}$ and $\mathbb{A}$; see Bertsekas and Shreve [1, pp. 140–141] or Hernández-Lerma and Lasserre [11, p. 178]. Let $\mathbb{E}_x^\pi$ be an expectation with respect to P_x^π.

For a finite-horizon $N = 0,1,\dots,$ let us define the expected total discounted costs,

$$v^\pi_{N,\alpha} := \mathbb{E}^\pi_x \sum_{t=0}^{N-1} \alpha^t c(x_t,a_t), \ \ x\in\mathbb{X}, \tag{1}$$

where $\alpha\in[0,1)$ is the discount factor and $v^\pi_{0,a}(x) = T(x)$ is the terminal cost. When $N=\infty$, equation (1) defines an infinite-horizon expected total discounted cost denoted by $v^\pi_\alpha(x)$. Let $v_\alpha := \inf_{\pi\in\Pi} v^\pi_\alpha(x)$, $x\in\mathbb{X}$. A policy π is called optimal for the respective criterion with discount factor α if $v^\pi_{N,\alpha}(x) = v_{N,\alpha}(x)$ or $v^\pi_\alpha(x) = v_\alpha(x)$ for all $x\in\mathbb{X}$.

The *average cost per unit time* is defined as

$$w^\pi(x) := \limsup_{N\to+\infty} \frac{1}{N} v^\pi_{N,1}(x), \ \ x\in\mathbb{X}. \tag{2}$$

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 $x\in\mathbb{X}$.

Let

$$\begin{aligned} m_\alpha &:= \inf_{x\in\mathbb{X}} v_\alpha(x), \quad u_\alpha(x) := v_\alpha(x) - m_\alpha, \\ \underline{w} &:= \liminf_{\alpha\uparrow 1}(1-\alpha)m_\alpha, \quad \bar{w} := \limsup_{\alpha\uparrow 1}(1-\alpha)m_\alpha \end{aligned} \tag{3}$$

The function u_α is called the discounted-cost relative value function. Assume that the following assumption holds in addition to Assumption **W***.

Assumption B. (i) $w^* := \inf_{x\in\mathbb{X}} w(x) < \infty$, and

(ii) $\sup_{\alpha\in[0,1)} u_\alpha(x) < \infty$, $x\in\mathbb{X}$.

As follows from Schäl [15, Lemma 1.2(a)], Assumption **B**(i) implies that $m_\alpha < +\infty$ for all $\alpha\in[0,1)$. Thus, all the quantities in (3) are defined. According to Feinberg et al. [5, Theorem 4], if Assumptions **W*** and **B** hold, then $\underline{w} = \bar{w}$. In addition, $\lim_{n\to+\infty}(1-\alpha_n)m_{\alpha_n} = \underline{w} = \bar{w}$ for each sequence $\{\alpha_n\}_{n=1,2,\dots}$ such that $\alpha_n\uparrow 1$ as $n\to+\infty$.

Define the following function on $\mathbb{X}$ for the sequence $\{\alpha_n\uparrow 1\}_{n=1,2,\dots}$:

$$\tilde{u}(x) \ := \ \liminf_{n\to+\infty, y\to x} u_{\alpha_n}(y). \tag{4}$$

In words, $\tilde{u}(x)$ is the largest number such that $\tilde{u}(x)\le \liminf_{n\to\infty} u_{\alpha_n}(y_n)$ for all sequences $\{y_n\to x\}$. Since $u_\alpha(x)$ is nonnegative by definition, then $\tilde{u}(x)$ is also nonnegative. The function $\tilde{u}$, defined in (4) for a sequence $\{\alpha_n\uparrow 1\}_{n=1,2,\dots}$ of nonnegative discount factors, is called an average-cost relative value function.

III. Sufficient Conditions for ACOEs

If Assumptions **W*** and **B** hold, then, according to Feinberg et al. [5, Corollary 2], there exists a stationary policy ϕ satisfying

$$\underline{w} + \tilde{u}(x) \ge c(x,\phi(x)) + \int_{\mathbb{X}} \tilde{u}(y)q(dy|x,\phi(x)), \tag{5}$$

for all $x \in \mathbb{X}$, where $\tilde{u}$ is defined in (4) for an arbitrary sequence of discount factors $\{\alpha_n \uparrow 1\}_{n=1,2,\ldots}$, and

$$w^{\phi}(x) = \underline{w} = \lim_{\alpha\uparrow 1}(1-\alpha)v_\alpha(x) = \bar{w} = w^*, \quad x \in \mathbb{X}. \quad (6)$$

These equalities imply that the stationary policy ϕ is average-cost optimal and $w^{\phi}(x)$ does not depend on x.

Inequality (5) is known as the ACOI. We remark that a weaker form of the ACOI with $\underline{w}$ substituted with $\bar{w}$ is also described in Feinberg et al. [5]. If Assumptions **W*** and **B** hold, let us define $w := \underline{w}$; see (6) for other equalities for w.

Recall the following definition of equicontinuity.

Definition 2. *A family $\mathcal{H}$ of real-valued functions on a metric space X is called equicontinuous at the point $x \in X$ if for each $\epsilon > 0$ there exists an open set G containing x such that*

$$|h(y) - h(x)| < \epsilon \ \text{ for all } y \in G \text{ and for all } h \in \mathcal{H}.$$

The family $\mathcal{H}$ is called equicontinuous (on X) if it is equicontinuous at all $x \in X$.

Consider the following equicontinuity condition (EC) on the discounted-cost relative value functions.

Assumption EC. (Feinberg and Liang [10]) There exists a sequence $\{\alpha_n \uparrow 1\}_{n=1,2,\ldots}$ of nonnegative discount factors such that

(i) the family of functions $\{u_{\alpha_n}\}_{n=1,2,\ldots}$ is equicontinuous, and

(ii) there exists a nonnegative measurable function $U(x)$, $x \in \mathbb{X}$, such that $U(x) \geq u_{\alpha_n}(x)$, $n = 1, 2, \ldots$, and $\int_{\mathbb{X}} U(y)q(dy|x,a) < \infty$ for all $x \in \mathbb{X}$ and $a \in \mathbb{A}$.

The following theorem provides sufficient conditions under which there exist a stationary policy ϕ and a function $\tilde{u}(\cdot)$ satisfying ACOEs for MDPs with weakly continuous transition probabilities.

Theorem 3. *Let Assumptions **W*** and **B** hold. Consider a sequence $\{\alpha_n \uparrow 1\}_{n=1,2,\ldots}$ of nonnegative discount factors. If Assumption EC is satisfied for the sequence $\{\alpha_n\}_{n=1,2,\ldots}$, then the following statements hold.*

(i) There exists a subsequence $\{\alpha_{n_k}\}_{k=1,2,\ldots}$ of sequence $\{\alpha_n\}_{n=1,2,\ldots}$ such that $\{u_{\alpha_{n_k}}(x)\}$ converges pointwise to $\tilde{u}(x)$, $x \in \mathbb{X}$, where $\tilde{u}(x)$ is defined in (4) for the sequence $\{\alpha_{n_k}\}_{k=1,2,\ldots}$. In addition, the function $\tilde{u}(x)$ is continuous.

(ii) There exists a stationary policy ϕ satisfying the ACOE with the nonnegative function $\tilde{u}$ defined for the subsequence $\{\alpha_{n_k}\}_{k=1,2,\ldots}$ mentioned in statement (i)*, that is, for all $x \in \mathbb{X}$,*

$$w + \tilde{u}(x) = c(x, \phi(x)) + \int_{\mathbb{X}} \tilde{u}(y)q(dy|x,\phi(x))$$
$$= \min_{a\in\mathbb{A}}[c(x,a) + \int_{\mathbb{X}} \tilde{u}(y)q(dy|x,a)], \quad (7)$$

and, since the left equation in (7) implies inequality (5), every stationary policy satisfying (7) is average-cost optimal.

IV. ACOEs for Inventory Control Problem

Let $\mathbb{R}$ denote the real line, $\mathbb{Z}$ denote the set of all integers, $\mathbb{R}_+ := [0, +\infty)$ and $\mathbb{N}_0 := \{0, 1, 2, \ldots\}$. Consider the stochastic periodic-review setup-cost inventory control problem with backorders. At times $t = 0, 1, \ldots$, a decision-maker views the current inventory of a single commodity and makes an ordering decision. Assuming zero lead times, the products are immediately available to meet demand. Demand is then realized, the decision-maker views the remaining inventory, and the process continues. The unmet demand is backlogged. The demand and the order quantity are assumed to be nonnegative. The state and action spaces are either (i) $\mathbb{X} = \mathbb{R}$ and $\mathbb{A} = \mathbb{R}_+$, or (ii) $\mathbb{X} = \mathbb{Z}$ and $\mathbb{A} = \mathbb{N}_0$. The inventory control problem is defined by the following parameters.

1) $\alpha \in [0, 1)$ is the discount factor;
2) $K \geq 0$ is a fixed ordering cost;
3) $\bar{c} > 0$ is the per unit ordering cost;
4) $\{D_t, t = 1, 2, \ldots\}$ is a sequence of i.i.d. nonnegative finite random variables representing the demand at periods $0, 1, \ldots$. We assume that $\mathbb{E}[D] < \infty$ and $P(D > 0) > 0$, where D is a random variable with the same distribution as D_1;
5) $h(x)$ is the holding/backlog cost per period if the inventory level is x. Assume that: (i) the function $\mathbb{E}[h(x - D)]$ is finite and continuous for all $x \in \mathbb{X}$; and (ii) $\mathbb{E}[h(x - D)] \to \infty$ as $|x| \to \infty$.

Without loss of generality, assume that h is nonnegative and $h(0) = 0$. The assumption $P(D > 0) > 0$ avoids the trivial case when there is no demand. If $P(D = 0) = 1$, then the optimality inequality does not hold because $w(x)$ depends on x; see Feinberg and Lewis [8] for details.

The dynamic of the system is defined by the equation

$$x_{t+1} = x_t + a_t - D_{t+1}, \quad t = 0, 1, 2, \ldots,$$

where x_t and a_t denote the current inventory level and the ordered amount at period t, respectively. Then, for $(x, a) \in \mathbb{X} \times \mathbb{A}$ the one-step cost is

$$c(x,a) = KI_{\{a>0\}} + \bar{c}a + \mathbb{E}[h(x + a - D)], \quad (8)$$

where $I_{\{a>0\}}$ is an indicator of event $\{a > 0\}$.

Recall the concept of (s, S) policies and quasiconvex functions. Suppose $f(x)$ is a continuous function such that $f(x) > K + \inf_{x\in\mathbb{X}} f(x)$ as $|x| \to \infty$. Let

$$S \in \operatorname*{argmin}_{x\in\mathbb{X}}\{f(x)\}, \quad (9)$$

$$s = \inf\{x \leq S : f(x) \leq K + f(S)\}. \quad (10)$$

Definition 4. *Let s_t and S_t be real numbers such that $s_t \leq S_t$, $t = 0, 1, \ldots$. A policy is called an (s_t, S_t) policy at step t if it orders up to the level S_t, if $x_t < s_t$, and does not order, if $x_t \geq s_t$. A Markov policy is called an (s_t, S_t) policy if it is an (s_t, S_t) policy at all steps $t = 0, 1, \ldots$. A policy is called an (s, S) policy if it is stationary and it is an (s, S) policy at all steps $t = 0, 1, \ldots$.*

Definition 5. *A function f is quasiconvex on a convex set X if for all x and $y \in X$ and $0 \leq \lambda \leq 1$,*

$$f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}.$$

Consider the following assumptions which guarantee the optimality of (s, S) policies for discounted-cost and average-cost criteria. Define

$$h_\alpha(x) = h(x) + (1-\alpha)\bar{c}x + \bar{c}\mathbb{E}[D]. \tag{11}$$

Assumption 1. *(i) The function*

$$\mathbb{E}[h_\alpha(x-D)] = \mathbb{E}[h(x-D)] + (1-\alpha)\bar{c}x + \alpha\bar{c}\mathbb{E}[D] \tag{12}$$

is quasiconvex for all $\alpha \in [0, 1]$;

(ii) There exists $\tilde{\alpha} \in [0, 1)$ such that

$$\lim_{x\to-\infty} \mathbb{E}[h_{\tilde{\alpha}}(x-D)] > K + \inf_{x\in\mathbb{X}}\{\mathbb{E}[h_{\tilde{\alpha}}(x-D)]\}. \tag{13}$$

Assumption 2. *The function $h(\cdot)$ is convex on $\mathbb{X}$.*

Note that since $\mathbb{E}[h(x-D)] \to \infty$ as $x \to \infty$ and $(1-\alpha)\bar{c} \geq 0$ for all $\alpha \in [0, 1]$, then (12) implies that $\mathbb{E}[h_\alpha(x-D)] \to \infty$ as $x \to \infty$ for all $\alpha \in [0, 1]$.

For $\alpha \in [0, 1]$, if

$$\lim_{x\to-\infty} \mathbb{E}[h_\alpha(x-D)] > \inf_{x\in\mathbb{X}} \mathbb{E}[h_\alpha(x-D)], \tag{14}$$

then define

$$x_\alpha^{\min} := \min\left\{\operatorname*{argmin}_{x\in\mathbb{X}}\{\mathbb{E}[h_\alpha(x-D)]\}\right\}, \tag{15}$$

where (14), the continuity of the function $\mathbb{E}[h_\alpha(x-D)]$ and $\mathbb{E}[h_\alpha(x-D)] \to \infty$ as $x \to \infty$ imply that $|x_\alpha^{\min}| < \infty$.

In addition to Assumption 1, consider the following assumption, which is used to establish the convergence of the discounted optimal lower thresholds and relative value functions in Sections V and VI, respectively.

Assumption 3. *For a given $\alpha \in [0, 1]$, the function $\mathbb{E}[h_\alpha(x-D)]$ is strictly decreasing on $(-\infty, x_\alpha^{\min}]$, where $x_\alpha^{\min}$ is defined in* (15)*.*

Consider the function h_α defined in (11). Let

$$\alpha^* := \sup\{\beta \in [0,1) : \lim_{x\to-\infty} \mathbb{E}[h_\beta(x-D)] \leq K + \inf_{x\in\mathbb{X}} \mathbb{E}[h_\beta(x-D)]\}, \tag{16}$$

where the supremum of an empty set is $-\infty$. If Assumption 1(ii) holds, then $\alpha^* < 1$. Note that $x_\alpha^{\min} \geq x_\beta^{\min}$ if $\alpha > \beta > \alpha^*$.

Proposition 6. *If Assumption 2 holds, then Assumption 1 holds with α^* defined as*

$$\alpha^* := 1 + \lim_{x\to-\infty} \frac{h(x)}{\bar{c}x}, \tag{17}$$

where the limit exists and $\alpha^* < 1$.

Proposition 7. *If Assumption 2 holds, then Assumption 3 holds.*

Lemma 8. *Let Assumption 1 hold. The described inventory control model with quasiconvex cost functions satisfies Assumptions **W*** and **B**.*

Since Assumptions **W*** and **B** hold for the MDP corresponding to the described inventory control problem, then the optimality equations for the infinite-horizon total discounted costs can be written as

$$v_\alpha(x) = \min\{\min_{a\geq 0}[K + G_\alpha(x+a)], G_\alpha(x)\} - \bar{c}x, \tag{18}$$

where

$$G_\alpha(x) := \bar{c}x + \mathbb{E}[h(x-D)] + \alpha\mathbb{E}[v_\alpha(x-D)]. \tag{19}$$

In addition, the function $c : \mathbb{X} \times \mathbb{A} \to \mathbb{R}$ is inf-compact. This property and the validity of Assumption **W*** imply that for each $\alpha \in [0, 1)$ the function v_α is inf-compact (Feinberg and Lewis [7, Proposition 3.1(iv)]) and therefore the set $X_\alpha := \{x \in \mathbb{X} | v_\alpha(x) = m_\alpha\}$, where m_α is defined in (3), is nonempty and compact. The validity of Assumptions **W*** and **B**(i) and the inf-compactness of c imply that there is a compact subset $\mathcal{K}$ of $\mathbb{X}$ such that $\mathbb{X}_\alpha \subseteq \mathcal{K}$ for all $\alpha \in [0, 1)$; Feinberg et. al. [5, Theorem 6]. Following Feinberg and Lewis [8], let us consider a bounded interval $[x_L^*, x_U^*] \subseteq \mathbb{X}$ such that

$$X_\alpha \subseteq [x_L^*, x_U^*] \qquad \text{for all } \alpha \in [0, 1). \tag{20}$$

The following theorem states the optimality of (s_α, S_α) policies for discounted-cost problems with quasiconvex cost functions.

Theorem 9. *If Assumption 1 holds, then for an infinite-horizon problem with discount factor $\alpha \in (\alpha^*, 1)$, an (s_α, S_α) policy is optimal, where the real numbers S_α and s_a are defined in* (9) *and* (10) *respectively with $f(x) := G_\alpha(x)$, $x \in \mathbb{X}$.*

If Assumption 2 holds, then Theorem 9 is proved in Feinberg and Liang [9, Theorem 4.4] with α^* defined in (17).

Since Assumptions **W*** and **B** hold, then the average-cost optimality inequality can be written as

$$w + \tilde{u}(x) \geq \min\{\min_{a\geq 0}[K + H(x+a)], H(x)\} - \bar{c}x, \tag{21}$$

where

$$H(x) := \bar{c}x + \mathbb{E}[h(x-D)] + \mathbb{E}[\tilde{u}(x-D)]. \tag{22}$$

The following theorem is proved in Feinberg and Lewis [8, Theorem 6.10(iii)] under Assumption 2. By verifying the validity of Assumptions **W*** and **B** under Assumption 1, the results can be generalized to inventory control problems with quasiconvex cost functions.

Theorem 10. *Let Assumption 1 hold. For each nonnegative $\alpha \in (\alpha^*, 1)$, consider an optimal (s'_α, S'_α) policy for the discounted-cost criterion with the discount factor α. Let $\{\alpha_n \uparrow 1\}_{n=1,2,\ldots}$ be a sequence of negative numbers with $\alpha_1 > \alpha^*$. Every sequence $\{(s'_{\alpha_n}, S'_{\alpha_n})\}_{n=1,2,\ldots}$ is bounded, and each its limit point (s^*, S^*) defines an average-cost optimal (s^*, S^*) policy. Furthermore, this policy satisfies* *the optimality inequality* (21)*, where the function $\tilde{u}$ is defined in* (4) *for an arbitrary subsequence $\{\alpha_{n_k}\}_{k=1,2,\dots}$ of $\{\alpha_n\}_{n=1,2,\dots}$ satisfying $(s^*,S^*)=\lim_{k\to\infty}(s'_{\alpha_{n_k}},S'_{\alpha_{n_k}})$.*

The following theorem states that the conditions and conclusions of Theorem 3, which describes sufficient conditions for the validity of average-cost optimality equations for MDPs, hold for the described inventory control problem. It also states some problem-specific results.

Theorem 11. *Let Assumption 1 hold. The MDP for the described inventory control problem satisfies the sufficient conditions stated in Theorem 3. Therefore, the conclusions of Theorem 3 hold for any sequence $\{\alpha_n\uparrow 1\}_{n=1,2,\dots}$ of nonnegative discount factors with $\alpha_1>\alpha^*$, that is, there exist a subsequence $\{\alpha_{n_k}\}_{k=1,2,\dots}$ of $\{\alpha_n\}_{n=1,2,\dots}$, a stationary policy φ, and a function $\tilde{u}$ defined in* (4) *for the subsequence $\{\alpha_{n_k}\}_{k=1,2,\dots}$ such that for all $x\in\mathbb{X}$*

$$\begin{aligned} w+\tilde{u}(x) &= KI_{\{\varphi(x)>0\}}+H(x+\varphi(x))-\bar{c}x \\ &= \min\{\min_{a\geq 0}[K+H(x+a)],H(x)\}-\bar{c}x, \quad (23)\end{aligned}$$

where the function H is defined in (22)*. In addition, the functions $\tilde{u}$ and H are continuous and inf-compact, and a stationary optimal policy φ satisfying* (23) *can be selected as an (s^*,S^*) policy described in Theorem 10. It also can be selected as an (s,S) policy with the real numbers S and s satisfying* (9) *and defined in* (10) *respectively for $f(x)=H(x)$, $x\in\mathbb{X}$.*

Let us formulate several auxiliary facts for the discounted-cost relative value functions. Consider the renewal process

$$\mathbf{N}(t):=\sup\{n=0,1,\dots\,|\,\mathbf{S}_n\leq t\},$$

where $t\in\mathbb{R}_+$, $\mathbf{S}_0=0$ and $\mathbf{S}_n=\sum_{j=1}^n D_j$ for $n=1,2,\dots$. Observe that since $P(D>0)>0$, then $\mathbb{E}[\mathbf{N}(t)]<+\infty$, $t\in\mathbb{R}_+$; see Resnick [14, Theorem 3.3.1].

Consider an arbitrary $\alpha\in[0,1)$ and a state x_α such that $u_\alpha(x_\alpha)=m_\alpha$, where m_α is defined in (3). Then, in view of (20), the inequalities $x_L^*\leq x_\alpha\leq x_U^*$ take place.

Define $E_y(x):=\mathbb{E}[h(x-\mathbf{S}_{\mathbf{N}(y)+1})]$ for $x\in\mathbb{X}$, $y\geq 0$. In view of Feinberg and Lewis [8, Lemma 6.2], $E_y(x)<+\infty$. Let $E(x):=h(x)+E_{x-x_L^*}(x)$ and

$$U(x):= \begin{cases} K+\bar{c}(x_U^*-x), & \text{if } x<x_L^*,\\ K+\bar{c}(x_U^*-x_L^*)+ & \\ (E(x)+\bar{c}\mathbb{E}[D])(1+\mathbb{E}[\mathbf{N}(x-x_L^*)]), & \text{if } x\geq x_L^*. \end{cases} \quad (24)$$

Lemma 12. *Let Assumption 1 hold. The following inequalities hold for $\alpha\in[0,1)$:*

(i) $u_\alpha(x)\leq U(x)<+\infty$ for all $x\in\mathbb{X}$;

(ii) If $x_*,x\in\mathbb{X}$ and $x_*\leq x$, then $C(x_*,x):=\sup_{y\in[x_*,x]}U(y)<+\infty$;

(iii) $\mathbb{E}[U(x-D)]<+\infty$ for all $x\in\mathbb{X}$.

The following theorem shows that the equicontinuity conditions stated in Theorem 3 holds for the inventory model with holding/backlog costs satisfying quasiconvexity assumptions.

Theorem 13. *Let Assumption 1 hold. Consider α^* defined in* (16)*. Then for each $\beta\in(\alpha^*,1)$, the family of functions $\{u_\alpha\}_{\alpha\in[\beta,1)}$ is equicontinuous on $\mathbb{X}$.*

According to Lemma 12 and Theorem 13, Assumption **EC** holds for the inventory control problem satisfying the quasiconvexity assumption. Furthermore, the continuity of average-cost relative value functions implies the following corollary.

Corollary 14. *Let Assumption 1 hold, the state space $\mathbb{X}=\mathbb{R}$, and the action space $\mathbb{A}=\mathbb{R}_+$. For the (s,S) policy defined in Theorems 11, consider the stationary policy φ coinciding with this policy at all $x\in\mathbb{X}$, except $x=s$, and with $\varphi(s)=S-s$. Then the stationary policy φ also satisfies the optimality equation* (23)*, and therefore the policy φ is average-cost optimal.*

V. Convergence of Optimal Ordering Thresholds s_α

In this section, we establish for problems with quasiconvex costs the convergence of discounted-cost optimal ordering thresholds $s_\alpha\to s$ as $\alpha\uparrow 1$, where s an optimal average-cost threshold.

Theorem 15. *Let Assumptions 1 and 3 hold. Then the limit*

$$s^*:=\lim_{\alpha\uparrow 1}s_\alpha \quad (25)$$

exists and $s^*\leq x_1^{\min}$, where $x_1^{\min}$ is defined in (15)*.*

We also state several auxiliary facts. Consider

$$\overline{v}_\alpha(x):=v_\alpha(x)+\bar{c}x, \qquad x\in\mathbb{X}. \quad (26)$$

Then the optimality equations (18) can be written as $\overline{v}_\alpha(x)=\min\{\min_{a\geq 0}[K+G_\alpha(x+a)],G_\alpha(x)\}$, where $G_\alpha(x)=\mathbb{E}[h_\alpha(x-D)]+\alpha\mathbb{E}[\overline{v}_\alpha(x-D)]$.

For $x\in\mathbb{X}$, define

$$\overline{m}_\alpha:=\min_{x\in\mathbb{X}}\{\overline{v}_\alpha(x)\} \quad \text{and} \quad \overline{u}_\alpha(x):=\overline{v}_\alpha(x)-\overline{m}_\alpha. \quad (27)$$

For $\alpha\in(\alpha^*,1)$ define the set of all possible optimal discounted lower thresholds

$$\mathcal{G}_\alpha:=\{x\leq S_\alpha : G_\alpha(y)=K+G_\alpha(S_\alpha) \text{ for all } y\in[s_\alpha,x]\}, \quad (28)$$

where S_α satisfies (9) and s_α is defined in (10) with $f:=G_\alpha$. Note that $s_\alpha\in\mathcal{G}_\alpha$ and $y\geq s_\alpha$ for all $y\in\mathcal{G}_\alpha$.

Lemma 16. *Let Assumption 1 hold. Then $(1-\alpha)(\overline{m}_\alpha+K)=\mathbb{E}[h_\alpha(y-D)]$ for all $\alpha\in(\alpha^*,1)$ and $y\in\mathcal{G}_\alpha$.*

Lemma 17. *Let Assumption 1 hold. Then $y\leq x_\alpha^{\min}\leq x_1^{\min}$ for all $\alpha\in(\alpha^*,1)$ and $y\in\mathcal{G}_\alpha$.*

Lemma 18. *Let Assumption 1 hold. Then,*

$$\lim_{\alpha\uparrow 1}(1-\alpha)\overline{m}_\alpha=\lim_{\alpha\uparrow 1}\mathbb{E}[h_\alpha(s_\alpha-D)]=w. \quad (29)$$

The following theorem establishes the uniqueness of possible optimal lower thresholds for the inventory model with convex cost functions under the discounted criterion.

Theorem 19. *Let Assumption 1 hold and Assumption 3 hold for the discount factor $\alpha \in (\alpha^*, 1)$, where α^* is defined in* (16)*. Then $\mathcal{G}_\alpha = \{s_\alpha\}$, where $\mathcal{G}_\alpha$ and s_α are defined in* (28) *and* (10) *with $f := G_\alpha$, respectively.*

VI. Convergence of Relative Value Functions for Inventory Control

This section establishes the convergence of discounted relative value functions to the average-cost relative value function for the setup-cost inventory model when the discount factor tends to 1.

Let us define

$$u(x) := \liminf_{\alpha\uparrow 1, y\to x} u_\alpha(y). \tag{30}$$

The following theorem states the convergence of discounted relative value functions, when the discount factor converges to 1, to the average-cost relative value function u.

Theorem 20. *Let Assumption 1 hold and Assumption 3 hold for $\alpha = 1$. Then,*

$$\lim_{\alpha\uparrow 1} u_\alpha(x) = u(x), \qquad x \in \mathbb{X}, \tag{31}$$

and the function u is continuous.

Theorem 20 implies that (22) can be written as $H(x) := \bar{c}x + \mathbb{E}[h(x-D)] + \mathbb{E}[u(x-D)]$.

We also state several auxiliary facts.

Lemma 21. *Let Assumption 1 hold. Then:*

(i) for each $\beta \in (\alpha^*, 1)$, where α^* is defined in (16)*, the family of functions $\{\overline{u}_\alpha\}_{\alpha\in[\beta,1)}$ is equicontinuous on $\mathbb{X}$;*

(ii) $\sup_{\alpha\in(\alpha^*,1)} \overline{u}_\alpha(x) < +\infty$ for all $x \in \mathbb{X}$.

Lemma 22. *Let Assumption 1 hold and Assumption 3 hold for $\alpha = 1$. Then there exists the limit $\overline{u}(x) := \lim_{\alpha\uparrow 1} \overline{u}_\alpha(x)$ for all $x \in \mathbb{X}$, where the function $\overline{u}$ is continuous on $\mathbb{X}$.*

In view of (3), (26) and (27), $u_\alpha(x) = \overline{u}_\alpha(x) + \overline{m}_\alpha - m_\alpha - \bar{c}x$, for all $x \in \mathbb{X}$.

Corollary 23. *Let Assumption 1 hold and Assumption 3 hold for $\alpha = 1$. Then the conclusions in Theorem 11 hold with $\tilde{u} = u$ defined in* (31) *and s^* defined in* (25)*, that is, the functions $\tilde{u}$ and the thresholds s^* defined in* (4) *and Theorem 10, respectively, are the same for all sequences $\{\alpha_n \uparrow 1\}_{n=1,2,\dots}$.*

Define the set of all possible optimal average-cost lower thresholds

$$\mathcal{G} := \{x \le S : H(y) = K + H(S) \text{ for all } y \in [s, x]\}, \tag{32}$$

where $S = \min\{\operatorname{argmin}_x\{H(x)\}\}$ and s is defined in (10) with $f := H$. Note that $s \in \mathcal{G}$ and $y \ge s$ for all $y \in \mathcal{G}$.

The following theorem establishes the uniqueness of possible optimal lower thresholds for the inventory model with holding/backlog costs satisfying quasiconvexity assumptions under the average cost criterion.

Theorem 24. *Let Assumption 1 hold and Assumption 3 hold for $\alpha = 1$. Then $\mathcal{G} = \{s^*\}$, where $\mathcal{G}$ and s^* are defined in* (32) *and* (25)*, respectively.*

The following corollary states that all the results of this paper hold for inventory models with convex holding/backlog costs.

Corollary 25. *The conclusions of lemmas, theorems, and corollaries in Sections V and VI hold under Assumption 2.*

References

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

[2] Cavazos-Cadena R. (1991). A counterexample on the optimality equation in markov decision chains with the average cost criterion. *System and Control Letters,* 16(5), 387–392.

[3] Costa O.L.V., & Dufour F. (2012). Average control of Markov decision processes with Feller transition probabilities and general action spaces. *Journal of Mathematical Analysis and Applications,* 396(1), 58–69.

[4] Feinberg E. A. (2016). 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), INFORMS, Cantonsville, MD.

[5] Feinberg E. A., Kasyanov P. O., & Zadoianchuk N. V. (2012). Average cost Markov decision processes with weakly continuous transition probability. *Math. Oper. Res.,* 37(4), 591–607.

[6] Feinberg E. A., Kasyanov P. O., & Zadoianchuk N. V. (2013). Berges theorem for noncompact image sets. *J. Math. Anal. Appl.,* 397(1), 255–259.

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

[8] Feinberg E. A. & Lewis M. E. (2017). On the convergence of optimal actions for Markov decision processes and the optimality of (s, S) inventory policies. *Naval Research Logistic,* DOI:10.1002/nav.21750.

[9] Feinberg E. A. & Liang Y. (2017). Structure of optimal policies to periodic-review inventory models with convex costs and backorders for all values of discount factors. *Annals of Operations Research,* DOI:10.1007/s10479-017-2548-6.

[10] Feinberg E. A. & Liang Y. (2017). On the optimality equation for average cost Markov decision processes and its validity for inventory control. *Annals of Operations Research,* DOI:10.1007/s10479-017-2561-9.

[11] Hernández-Lerma O., & Lasserre J. B. (1996). *Discrete-Time Markov Control Processes: Basic Optimality Creteria.* Springer-Verlag, New York.

[12] Huh W. T., Janakiraman G., & Nagarajan M. (2011) Average Cost Single-Stage Inventory Models: An Analysis Using a Vanishing Discount Approach. *Operations Research,* 59(1), 143–155.

[13] Jaśkiewicz A., & Nowak A. S. (2006). On the optimality equation for average cost Markov control processes with Feller transition probabilities. *Journal of Mathematical Analysis and Applications,* 316(2), 495–509.

[14] Resnick S.I. (1992). *Adventures in stochastic processes.* Birkhauser, Boston.

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

[16] Simchi-Levi D., Chen X. & Bramel J. (2005) *The Logic of Logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management.* Springer-Verlag, New York.

[17] Veinott A. F. (1966). On the Optimality of (s, S) Inventory Policies: New Condition and a New Proof. *J. SIAM Appl. Math.,* 14(5), 1067–1083.

[18] Zheng Y. (1991). A simple proof for optimality of (s, S) policies in infinite-horizon inventory systems. *J. Appl. Prob.,* 28(4), 802–810.