

Optimality Inequalities for Average Cost Markov Decision Processes and the Optimality of (s, S) Policies

Eugene A. Feinberg

Department of Applied Mathematics and Statistics

State University of New York at Stony Brook

Stony Brook, NY 11794-3600

efeinberg@notes.cc.sunysb.edu

631-632-7189

Mark E. Lewis

Department of Industrial and Operations Engineering

University of Michigan

1205 Beal Avenue, Ann Arbor, MI 48109-2117

melewis@engin.umich.edu

734-763-0519

submitted April 15, 2005

Keywords: Markov decision process, average cost per unit time, optimality inequality, optimal policy, inventory control

MSC 2000 Subject Classifications: Primary: 90C40; Secondary: 90B05

OR/MS subject classifications: Primary: Dynamic Programming/optimal control/Markov/Infinite state; Secondary: Inventory/production/Uncertainty/Stochastic

Abstract

For general state and action space Markov decision processes, we present sufficient conditions for convergence of both the optimal discounted cost value function and policies to the corresponding objects for the average costs per unit time. We extend Schäl's [24] assumptions, guaranteeing the existence of a solution to the average cost optimality inequalities for compact action sets, to non-compact action sets. Since a stationary policy satisfying the optimality inequalities is average cost optimal, this paper provides sufficient conditions for the existence of stationary optimal policies for the average cost criterion. Inventory models are natural candidates for the application of our results. In particular, we provide straightforward proofs of the optimality of (s, S) policies for classic inventory control problems with generally distributed non-negative demand and the convergence of optimal thresholds for discounted costs to optimal thresholds for average costs per unit time as the discount factor tends to 1.

1 Introduction

In a discrete-time Markov decision process (MDP) the usual method to study the average cost criterion is to find a solution to the average cost optimality equations. A policy that achieves the minimum in this system of equations is then average cost optimal. When the state and action spaces are infinite, one may be required to

replace the equations with inequalities, yet the conclusions are the same; a policy that achieves the minimum in the inequalities is average cost optimal. Schäl [24] provides two groups of general conditions that imply the existence of a solution to the average cost optimality *inequalities* (ACOI). The first group, referred to as Assumptions **(W)** in Schäl [24], require weak continuity of the transition probabilities. The second group, Assumptions **(S)**, require setwise continuity of the transition probabilities. In either case, for each state a compact action set was assumed in [24]. The purpose of this paper is to relax the compact action space assumptions in Schäl [24] so that the results can be applied to problems with noncompact action sets; in particular to those related to inventory control. As was noted in [11], typical inventory control models (with general demand distributions) require weak continuity; setwise continuity is not enough. On the other hand, when the demand distribution is restricted to be continuous or the inventory is restricted to be integer, we show that setwise continuity does suffice.

The books by Sennott [25] and Hernández-Lerma and Lasserre [15] deal with countable and general state MDPs, respectively. Hernández-Lerma and Lasserre [15], Chapter 5, and Fernández-Gaucherand [12] present results for non-compact action sets but assume setwise continuity. Moreover, Section 5.7 in Hernández-Lerma and Lasserre [15] provides conditions for the existence of stationary optimal policies for an MDP with weakly continuous transition probabilities but the derivation is done directly; without deriving the optimality equations or inequalities. We are interested not only in the existence of optimal policies but in the validity of the optimality inequalities. This is an important step since these inequalities can often be used to prove structural properties of optimal stationary policies and to prove convergence of discounted cost optimal policies to average cost optimal policies. We recall that, according to the example constructed by Cavazos-Cadena [4], optimality inequalities may hold for an MDP for which optimality equalities do not hold. In addition, optimality inequalities imply the existence of optimal policies [24, Proposition 1.3].

The inventory control literature is far too expansive to attempt a complete literature review. The reader is pointed to the survey article by Porteus [21]. There is also a treatment of both Markov decision processes and their relationship to inventory control in [16]. In the case of inventory control, under the average cost criterion the optimality of (s, S) policies was proved by Iglehart [18] and Veinott and Wagner [27] in the continuous and discrete demand cases, respectively. The latter proof was simplified significantly

by Zheng [28] where the author proved in the discrete case the existence of a solution to the average cost optimality equations by construction, instead of taking (undiscounted) limits of the finite horizon problem. A more recent paper in average cost inventory models is the work of Beyer and Sethi [3]. The authors reconsider the continuous demand model of Iglehart [18] and verify several assumptions that apparently were not stated in the original work. They also make the observation that most of the work following Iglehart's paper showing that (s, S) policies are optimal in the average cost case assume some bounds on the inventory position after ordering. Without such assumptions, the optimality of (s, S) policies for average cost inventory control problems follows from Chen and Simchi-Levi [7] where methods specific to inventory control were used.

The rest of the paper is organized as follows. In Section 2 we discuss the general Markov decision process framework. Section 3 explains some related results from Hernández-Lerma and Lasserre [15] and Schäl [24]. Section 4 contains the main theoretical contributions of the paper, Theorem 4.7, that provides two sets of assumptions that lead to: (i) the validity of the optimality inequalities for average cost MDPs, (ii) the existence of optimal policies for average cost MDPs, and (iii) the convergence of discounted cost optimal values and policies to average cost per unit time optimal values and policies. In Section 5 we discuss the relevance of the setwise and weak continuity assumptions to inventory control. We formulate the inventory control problem in Section 6 and show that it satisfies the weak continuity assumptions of Theorem 4.7. Theorem 6.9 states the optimality of (s_α, S_α) policies for discounted cost problems when the discount factor α is close to 1, the optimality of (s, S) policies for average costs per unit time, and the convergence of the thresholds s_α and S_α to the thresholds s and S , respectively as α tends to 1. The paper is concluded in Section 7.

2 Model Definition

Consider a discrete-time Markov decision process with the state space \mathbb{X} and action space \mathbb{A} . Assume that both \mathbb{X} and \mathbb{A} are Borel subsets of Polish (complete separable metric) spaces. For each $x \in \mathbb{X}$ the nonempty

Borel subset $A(x)$ represents the set of actions available at x . Define the graph of \mathbb{A} by

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

and assume that: (i) $Gr(\mathbb{A})$ is a measurable subset of $\mathbb{X} \times \mathbb{A}$, and (ii) there exists a measurable mapping ϕ from \mathbb{X} to \mathbb{A} such that $\phi(x) \in A(x)$ for all $x \in \mathbb{X}$. The one step cost, $c(x, a)$, for choosing action $a \in A(x)$ in state x is presumed a non-negative (or equivalently, bounded below) measurable function on $Gr(\mathbb{A})$. Let $q(B|x, a)$, also measurable on $Gr(\mathbb{A})$, be the transition kernel representing the probability that $B \subseteq \mathbb{X}$ is entered next, given that action a is chosen in state x . This means that $q(\cdot|x, a)$ is a probability measure on \mathbb{X} for each pair $(x, a) \in \mathbb{X} \times \mathbb{A}$, and $q(B|\cdot, \cdot)$ is a Borel function on $\mathbb{X} \times \mathbb{A}$ for any Borel subset $B \subseteq \mathbb{X}$.

The decision process proceeds as follows: at time n the current state of the system, x , is observed. A decision-maker decides which action, a , to choose, the cost $c(x, a)$ is accrued, the system moves to the next state according to $q(\cdot|x, a)$, and the process continues. Let $H_n = (\mathbb{X} \times \mathbb{A})^n \times \mathbb{X}$ be the set of histories for $n = 0, 1, \dots$. A (randomized) decision rule at epoch $n = 0, 1, \dots$ is a regular transition probability π_n from H_n to \mathbb{A} concentrated on $A(x_n)$. In other words, (i) $\pi_n(\cdot|h_n)$ is a probability distribution on \mathbb{A} such that $\pi_n(A(x_n)|h_n) = 1$, where $h_n = (x_0, a_0, x_1, \dots, a_{n-1}, x_n)$ and (ii) for any measurable subset $B \subseteq \mathbb{A}$, the function $\pi_n(B|\cdot)$ is measurable on H_n . A policy π is a sequence (π_0, π_1, \dots) of decision rules. Moreover, π is called non-randomized if each probability measure $\pi_n(\cdot|h_n)$ is concentrated at one point. A non-randomized policy is called Markov if all decisions depend only on the current state and time. A Markov policy is called stationary if all decisions depend only on the current state. Thus, a Markov policy ϕ is defined by a sequence ϕ_0, ϕ_1, \dots of measurable mappings $\phi_n : \mathbb{X} \rightarrow \mathbb{A}$ such that $\phi_n(x) \in A(x)$ for all $x \in \mathbb{X}$. A stationary policy ϕ is defined by a measurable mapping $\phi : \mathbb{X} \rightarrow \mathbb{A}$ such that $\phi(x) \in A(x)$ for all $x \in \mathbb{X}$.

The Ionescu–Tulcea theorem (cf. p. 140-141 of [2] or p. 178 of [15]) yields that an initial state x and a policy π define a unique probability distribution \mathbb{P}_x^π on the set of all trajectories $H_\infty = (\mathbb{X} \times \mathbb{A})^\infty$ endowed with the product σ -field defined by Borel σ -fields of \mathbb{X} and \mathbb{A} . Let \mathbb{E}_x^π be the expectation with respect to this

distribution. For a finite horizon $N = 0, 1, \dots$ define the expected total discounted costs

$$v_{N,\alpha}^\pi(x) := \mathbb{E}_x^\pi \sum_{n=0}^{N-1} \alpha^n c(x_n, a_n), \quad (2.1)$$

where $\alpha \in [0, 1)$ and $v_{0,\alpha}^\pi = 0$. When $N = \infty$, (2.1) defines the infinite horizon expected total discounted cost of π denoted $v_\alpha^\pi(x)$. The average costs per unit time are defined

$$w^\pi(x) := \limsup_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}_x^\pi \sum_{n=0}^{N-1} c(x_n, x_n). \quad (2.2)$$

For each function $G^\pi(x) = v_{N,\alpha}^\pi(x)$, $v_\alpha^\pi(x)$, or $w(x)$, define the optimal cost

$$G(x) := \inf_{\pi \in \Pi} G^\pi(x), \quad (2.3)$$

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 $x \in \mathbb{X}$.

It is well-known (see e.g. [2, Proposition 8.2]) that $v_{n,\alpha}(x)$ satisfies the following *optimality equations*,

$$v_{n+1,\alpha}(x) = \inf_{a \in A(x)} \{c(x, a) + \alpha \int v_{n,\alpha}(y) q(dy|x, a)\}, \quad x \in \mathbb{X}, n = 0, 1, \dots \quad (2.4)$$

In addition, a Markov policy ϕ_α , defined at the first N steps by the mappings $\phi_0, \dots, \phi_{N-1}$ that satisfies the following equations for all $x \in X$ and all $n = 1, \dots, N$

$$v_{n,\alpha}(x) = c(x, \phi_{N-n}(x)) + \alpha \int v_{n-1,\alpha}(y) q(dy|x, \phi_{N-n,\alpha}(x)), \quad x \in \mathbb{X}, \quad (2.5)$$

is optimal for the horizon N ; see e.g. [2, Lemma 8.7].

It is also well-known (see e.g. [2, Proposition 9.8]) that $v_\alpha(x)$ satisfies the following *discounted cost optimality equations* (DCOE),

$$v_\alpha(x) = \inf_{a \in A(x)} \{c(x, a) + \alpha \int v_\alpha(y) q(dy|x, a)\}, \quad x \in \mathbb{X}. \quad (2.6)$$

If a stationary policy ϕ_α satisfies

$$v_\alpha(x) = c(x, \phi_\alpha(x)) + \alpha \int v_\alpha(y)q(dy|x, \phi_\alpha(x)), \quad x \in \mathbb{X}, \quad (2.7)$$

then ϕ_α is optimal; [2, Proposition 9.8]. According to [24, Proposition 2.1], each of the conditions **(W)** and **(S)**, explained in the following section, imply the existence of a stationary policy that satisfies (2.7).

Since $c \geq 0$, for each $\pi \in \Pi$ and $x \in \mathbb{X}$, $v_{n,\alpha}^\pi(x)$ is non-decreasing in n and bounded above by $v_\alpha^\pi(x)$. Thus, $v_{n,\alpha}(x)$ is also non-decreasing in n and bounded above by $v_\alpha(x)$. Let $v_{\infty,\alpha}$ be the monotone limit of $v_{n,\alpha}$. It should be clear that for each $x \in \mathbb{X}$, $v_{\infty,\alpha}(x) \leq v_\alpha(x)$. In general, it is possible that $v_{\infty,\alpha}(x) < v_\alpha(x)$; [10, Example 6.6]. However, each of the conditions **(W)** and **(S)**, are sufficient for the validity of the equality $v_{\infty,\alpha} = v_\alpha$; see [23] or Proposition 4.1 below.

A little more subtle is the average cost case. For the remainder of the paper, assume that the following condition holds.

Assumption (G): There exists a policy π and an initial state x such that

$$w^\pi(x) < \infty. \quad (2.8)$$

Note that this is equivalent to the General Assumption of Schäl [24]: $\inf_{x \in \mathbb{X}} \inf_{\phi \in \Pi} w^\phi(x) < \infty$. Define the following quantities

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

and

$$w^* = \liminf_{\alpha \rightarrow 1} (1 - \alpha)m_\alpha. \quad (2.9)$$

Assumption **(G)** implies that $w^* < \infty$; Schäl [24, Lemma 1.2]. According to Schäl [24, Proposition 1.3], if there exists a measurable function $u : \mathbb{X} \rightarrow [0, \infty)$ and a stationary policy ϕ such that

$$w^* + u(x) \geq c(x, \phi(x)) + \int u(y)q(dy|x, \phi(x)), \quad x \in \mathbb{X}, \quad (2.10)$$

then ϕ is average cost optimal and $w(x) = w^*$ for all $x \in \mathbb{X}$. The following condition plays an important role for the validity of (2.10).

Assumptions (B): Assumption (G) holds and $\sup_{\alpha < 1} u_\alpha(x) < \infty$ for all $x \in \mathbb{X}$.

We note that the second part of Assumptions (B) is Condition (B) in Schäl [24]. Thus, under Assumption (G), which is assumed throughout [24], Assumptions (B) are equivalent to Condition (B) in [24].

3 Known Results

In this section, we discuss some closely related results. Schäl [24] studies problems with compact action sets when transition probabilities satisfy either Assumptions (W) below, when the transition probabilities are weakly continuous, or Assumptions (S) below, when the transition probabilities are setwise continuous. Each of these conditions together with (B) yield the existence of an optimal policy satisfying (2.10) and the convergence along a subsequence of both the optimal discounted cost values and policies to those in the average cost case. Hernández-Lerma and Lasserre [15, section 5.4] does not require a compact action space, but deal only with MDPs whose transition probabilities are setwise continuous. We note that [15, Section 4.2] contains some results for weak continuity but apparently no particular results are available for the optimality inequalities for MDPs with non-compact action sets and weakly continuous transition probabilities.

Weak continuity (or continuity with respect to weak convergence) of q in (x, a) means that

$$\int f(z)q(dz|x^k, a^k) \rightarrow \int f(z)q(dz|x, a) \quad (3.1)$$

for any sequence $\{(x^k, a^k), k \geq 0\}$ converging to (x, a) , where $(x^k, a^k), (x, a) \in Gr(\mathbb{A})$, and for any bounded continuous function f . Recall that setwise continuity (or continuity with respect to setwise convergence) of q in (x, a) means that $q(B|x^k, a^k) \rightarrow q(B|x, a)$ as $(x^k, a^k) \rightarrow (x, a)$ for any Borel subset B of \mathbb{X} , where $(x^k, a^k), (x, a) \in Gr(\mathbb{A})$. Similarly, setwise continuity of q in a means that $q(B|x, a^k) \rightarrow q(B|x, a)$ as $a^k \rightarrow a$, where $a^k, a \in A(x)$, for any Borel subset B of \mathbb{X} and for any state $x \in \mathbb{X}$. We remark that: (i) an equivalent definition of setwise continuity in (x, a) is that (3.1) holds for all bounded measurable functions

f , and (ii) setwise continuity is a stronger assumption than weak continuity; see e.g. Hernández-Lerma and Lasserre [15, page 186].

Let $\mathbf{C}(\mathbb{A})$ denote the set of all non-empty compact subsets of \mathbb{A} and $\mathbb{P}(\mathbb{X})$ be the set of all probability measures on \mathbb{X} . The following two groups of assumptions are from Schäl [24].

Assumptions (W):

0. \mathbb{X} is locally compact with a countable base.
1. $A(x) \in \mathbf{C}(\mathbb{A})$ for $x \in \mathbb{X}$.
2. c is lower semi-continuous on $Gr(\mathbb{A})$.
3. $A : \mathbb{X} \rightarrow \mathbf{C}(\mathbb{A})$ is upper semi-continuous.
4. $q : Gr(\mathbb{A}) \rightarrow \mathbb{P}(\mathbb{X})$ is continuous with respect to weak convergence on $\mathbb{P}(\mathbb{X})$.

Assumptions (S):

1. $A(x) \in \mathbf{C}(\mathbb{A})$ for $x \in \mathbb{X}$.
2. $c(x, \cdot) : A(x) \rightarrow [0, \infty]$ is lower semi-continuous in a .
3. $q(x, \cdot) : A(x) \rightarrow \mathbb{P}(\mathbb{X})$ is continuous with respect to setwise convergence on $\mathbb{P}(\mathbb{X})$.

We next state the main results proved in Schäl [24, Proposition 3.5, Theorem 3.8].

Theorem 3.1 *Suppose (B) and either (W) or (S) hold. There exists a function $u : \mathbb{X} \rightarrow [0, \infty)$ and a stationary policy ϕ satisfying (2.10). Thus, $w^\phi(x) = w(x) = w^*$. Furthermore $w^* = \lim_{\alpha \rightarrow 1} (1 - \alpha)m_\alpha = \lim_{\alpha \rightarrow 1} (1 - \alpha)v_\alpha(x)$, $x \in \mathbb{X}$. Also, for any discount factor α fix an optimal policy ϕ_α . Then for any sequence of discount factors $\alpha(k) \rightarrow 1$ the following statements hold.*

1. Under (W),

(a) u can be defined as the lower semi-continuous function

$$u(x) = \liminf_{k \rightarrow \infty, y \rightarrow x} u_{\alpha(k)}(y); \tag{3.2}$$

(b) For each $x \in \mathbb{X}$ there exists a sequence $x_m \rightarrow x$ and a subsequence α_m of the sequence $\alpha(k)$ such that the mapping $\phi(x) = \lim_{m \rightarrow \infty} \phi_{\alpha_m}(x_m)$ defines a stationary policy that satisfies (2.10) with u defined in (3.2).

2. Under (S),

(a) u can be defined as the measurable function

$$u(x) = \liminf_{k \rightarrow \infty} u_{\alpha(k)}(x); \quad (3.3)$$

(b) For each $x \in \mathbb{X}$ there exists a subsequence α_m of the sequence $\alpha(k)$ such that the mapping $\phi(x) = \lim_{m \rightarrow \infty} \phi_{\alpha_m}(x)$ defines a stationary policy satisfying (2.10) with u defined in (3.3).

Since the lower semi-continuity and measurability of u in statements 1(a) and 2(a) of Theorem 3.1 were stated, but not proved in Schäl [24], we verify them here. Under (W), consider a sequence $x_n \rightarrow x$. For each $n = 1, 2, \dots$ consider an integer $k(n) \geq n$ and a point $y_{k(n)}$ in the neighborhood of x_n with the radius n^{-1} such that $u_{\alpha(k(n))}(y_{k(n)}) \leq u(x_n) + n^{-1}$. Then

$$\liminf_{n \rightarrow \infty} u(x_n) \geq \liminf_{n \rightarrow \infty} u_{\alpha(k(n))}(y_{k(n)}) \geq u(x),$$

where the last inequality follows from (3.2). Under (S), (3.3) defines u as a lower limit of a sequence of measurable functions. Therefore, by Shiryaev [26, p. 173], u is measurable.

Recall that lower semi-continuity requires that all of the level sets are closed. Consider the following stronger property.

Definition 3.2 A real-valued function f defined on a metric space \mathbb{Y} is called **inf-compact** if for all $\lambda \in \mathbb{R}$ the sets $D(\lambda) = \{y \in \mathbb{Y} \mid f(y) \leq \lambda\}$ are compact.

We formulate Assumptions 4.2.1 from Hernández-Lerma and Lasserre [15].

Assumptions (HL):

1. c is inf-compact on $Gr(\mathbb{A})$.

2. $q : Gr(\mathbb{A}) \rightarrow \mathbb{P}(\mathbb{X})$ is setwise continuous.

We remark that if (2.10) holds for a stationary policy ϕ and for a non-negative function u then obviously

$$w^* + u(x) \geq \inf_{a \in A(x)} \{c(x, a) + \int u(y)q(dy|x, a)\}, \quad x \in \mathbb{X}. \quad (3.4)$$

and each of the Assumptions **(W)**, **(S)**, and **(HL)** implies by the Arsenin-Kunugui theorem [19, Theorem 35.46] the existence of a policy ϕ such that the minimum in (3.4) is achieved when $a = \phi(x)$. Thus, infimum can be replaced with minimum in (3.4). Next we formulate the results from [15] relevant to Theorem 3.1.

Theorem 3.3 (see [15, Theorems 5.4.3 and 5.4.6]) *Let **(B)** and **(HL)** hold. Then there exists a function $u : \mathbb{X} \rightarrow [0, \infty)$ and a stationary policy ϕ satisfying (2.10). Therefore, $w^\phi(x) = w(x) = w^*$. Consider a sequence $\alpha(n) \rightarrow 1$ such that $\lim_{n \rightarrow \infty} (1 - \alpha(n))v_{\alpha(n)}(x)$ exists for some $x \in \mathbb{X}$ (such a sequence exists in light of the Lemma on p. 88 of [15]). Then for all $x \in \mathbb{X}$ this limit exists and equals w^* . Furthermore, the function u can be defined*

$$u(x) = \liminf_{n \rightarrow \infty} u_{\alpha(n)}(x). \quad (3.5)$$

In the next section we extend Assumptions **(S)** and **(W)** to MDPs with non-compact action sets.

4 Main Structural Results

We begin by stating assumptions for non-compact action sets; Assumptions **(Wu)** and **(Su)** below. These assumptions are similar to Assumptions **(W)** and **(S)**, respectively. The letter ‘‘u’’ signifies unbounded action sets.

Assumptions **(Wu)**:

0. \mathbb{X} is locally compact with a countable base.
1. c is inf-compact on $Gr(\mathbb{A})$.
2. $q : Gr(\mathbb{A}) \rightarrow \mathbb{P}(\mathbb{X})$ is continuous with respect to weak convergence on $\mathbb{P}(\mathbb{X})$.

Assumptions **(Su)**:

1. c is inf-compact on $A(x)$.
2. $q(x, \cdot) : A(x) \rightarrow \mathbb{P}(\mathbb{X})$ is continuous with respect to setwise convergence on $\mathbb{P}(\mathbb{X})$.

Recall that a stationary policy is optimal for the expected total discounted cost criterion if and only if (2.7) holds; see e.g. [8, Section 6.3]. The following proposition states that Assumptions **(Wu)** and **(Su)** imply the existence of stationary optimal policies for the expected total discounted cost criterion. This proposition is similar to [24, Proposition 2.1] for models with compact action sets.

Proposition 4.1 *Under **(Wu)** or **(Su)***

- (i) *For any $N = 1, 2, \dots$ there exists a Markov optimal policy $(\phi_0, \dots, \phi_{N-1})$ satisfying (2.5).*
- (ii) *For infinite horizon expected total discounted costs, there exists a stationary optimal policy satisfying (2.6).*
- (iii) *For infinite horizon expected total discounted costs there exists a stationary optimal policy ϕ_α satisfying (2.7).*
- (iv) *The functions $v_{n,\alpha}$, $n = 1, 2, \dots$, and v_α are inf-compact on \mathbb{X} .*
- (v) *$v_{n,\alpha}(x) \uparrow v_\alpha(x)$ for all $x \in \mathbb{X}$.*

Proof. We prove these results under **(Wu)**. Similar arguments hold under **(Su)**. Note that the weak continuity Assumption **(Wu2)** is equivalent to lower semi-continuity in (x, a) of the function

$$p(x, a) = \int f(y)q(dy|x, a)$$

for all lower semi-continuous, non-negative functions f ; [15, Proposition C.4].

Since $v_{0,\alpha}(x) = 0$ for all x , it is lower semi-continuous. Let $v_{n,\alpha}$ be lower semi-continuous for some n . Note that for all $\lambda \in \mathbb{R}$,

$$\begin{aligned} B_n(x, \lambda) &:= \{a \in A(x) \mid c(x, a) + \alpha \int v_{n,\alpha}(y)q(dy|x, a) \leq \lambda\} \\ &\subseteq \{a \in A(x) \mid c(x, a) \leq \lambda\} =: D(x, \lambda). \end{aligned}$$

Since $c(x, a)$ and $v_{n,\alpha}$ are lower semi-continuous, $B_n(x, \lambda)$ is a closed set. Thus, since the definition of inf-compactness implies $D(x, \lambda)$ is compact, $B_n(x, \lambda)$ is compact; a closed subset of a compact set is compact. That is to say, $J_n(x, a) = c(x, a) + \alpha \int v_{n,\alpha}(y)q(dy|x, a)$ is inf-compact. In addition, $v_{n+1,\alpha} = \min_{a \in A(x)} J_n(x, a)$ is inf-compact because $\{x \in \mathbb{X} \mid v_{n+1,\alpha}(x) \leq \lambda\}$ is the projection of a compact set.

The compactness of $B_n(x, \lambda)$ implies that $v_{n,\alpha} \uparrow v_\alpha$ (see [2, Proposition 9.17]). Therefore, the sets $\{x \in \mathbb{X} \mid v_\alpha(x) \leq \lambda\} = \bigcap_{n=1}^{\infty} \{x \in \mathbb{X} \mid v_{n,\alpha}(x) \leq \lambda\}$ are compact; v_α is inf-compact. This coupled with the weak continuity of q implies that $c(x, a) + \alpha \int v_\alpha(y)q(dy|x, a)$ is lower semi-continuous. Repeating the same argument as above yields the inf-compactness of $J(x, a)$ and (2.6) holds with infimum replaced by minimum. Moreover, the Arsenin-Kunugui theorem [19, Theorem 35.46] implies the existence of a Markov policy satisfying (2.5) and the existence of a stationary policy satisfying (2.7). ■

We remark that Proposition 9.17 in Bertsekas and Shreve [2] states the existence of a stationary optimal policy in addition to the convergence of $v_{n,\alpha}$ to v_α . In the proof of Proposition 4.1 we used the latter but we did not use the former because [2] used a more general concept of measurability than discussed in this paper. In fact, the above proof implies that the assumptions of [2, Proposition 9.17] imply the existence of a Borel measurable stationary optimal policy. This observation is not stated in [2] so we provided the proof here.

The next definition is an extension of the definition of a locally bounded function [13, p. 113].

Definition 4.2 *A real-valued function f defined on a metric space \mathbb{Y} is called **locally bounded above** at $x \in \mathbb{Y}$ if there exists an open set $B(x)$ containing x such that*

$$\sup_{y \in B(x)} f(y) < \infty.$$

*The function f is called **locally bounded above** if it is locally bounded for each $x \in \mathbb{Y}$.*

For a point $x \in \mathbb{Y}$, let $\mathcal{O}(x) := \{B \subseteq \mathbb{Y} \mid x \in B, B \text{ is open}\}$ be the set of open sets containing x . Consider the function

$$\bar{f}(x) := \inf_{B \in \mathcal{O}(x)} \sup_{y \in B} f(y).$$

Obviously, $f(x) \leq \bar{f}(x)$ and, if f is locally bounded at x , $\bar{f}(x) < \infty$.

Lemma 4.3 *Suppose f is a locally bounded above function on a complete separable metric space \mathbb{Y} . Then the function $\bar{f}(x)$ is upper semi-continuous.*

Proof. Fix $\epsilon > 0$ and let $B_\epsilon(x)$ be an open set containing x such that $\bar{f}(x) \geq \sup_{y \in B_\epsilon} f(y) - \epsilon$. Let $x_n \rightarrow x$. Choose $k(\epsilon)$ such that $x_n \in B_\epsilon$ for all $n \geq k(\epsilon)$. Then for $n \geq k(\epsilon)$

$$\bar{f}(x_n) \leq \sup_{y \in B_\epsilon} f(y) \leq \bar{f}(x) + \epsilon.$$

Thus, $\bar{f}(x) \geq \limsup_{n \rightarrow \infty} \bar{f}(x_n) - \epsilon$. Since $\epsilon > 0$ is arbitrary, $\bar{f}(x) \geq \limsup_{n \rightarrow \infty} \bar{f}(x_n)$. ■

For $\alpha \in [0, 1)$ define $r_\alpha(x) = \sup_{\alpha \leq \beta < 1} u_\beta(x)$. Assumptions **(B)** are equivalent to the validity of the following pair of assumptions: (i) Assumption **(G)**, and (ii) for any $x \in \mathbb{X}$ there exists $\alpha \in [0, 1)$ with $r_\alpha(x) < \infty$. Thus, Assumptions **(B)** are equivalent to [15, Condition 5.4.5] and implies that $r_\alpha(x) < \infty$ for all $\alpha \in [0, 1)$ and for $x \in \mathbb{X}$.

Let $\Gamma_\alpha(x) := \sup_{\alpha \leq \beta < 1} \{v_\beta(x) - \beta m_\beta\}$, where $\alpha \in [0, 1)$. According to [24, Lemma 1.2], Assumption **(G)** implies that $\limsup_{\alpha \rightarrow 1} (1 - \alpha)m_\alpha < \infty$. Thus, if Assumptions **(B)** hold, there exists $\alpha^* \in [0, 1)$ such that for all $\alpha \in [\alpha^*, 1)$

$$\Gamma_\alpha(x) \leq r_\alpha(x) + \sup_{\alpha \leq \beta < 1} \{(1 - \beta)m_\beta\} < \infty. \quad (4.1)$$

The following condition strengthens **(B)**.

Assumptions (LB). Assumption **(G)** holds and there exists $\alpha_0 \in [0, 1)$ such that the function $r_{\alpha_0}(x)$ is locally bounded above on \mathbb{X} .

We observe that the function $\Gamma_\alpha(x)$ decreases in α . Thus, if **(LB)** hold then the function Γ_α is locally bounded above for any $\alpha \in [\alpha_0, 1)$. We select an arbitrary $\alpha^* \in [\alpha_0, 1)$, such that (4.1) holds when $\alpha = \alpha^*$, and denote $\Gamma = \Gamma_{\alpha^*}$.

We say that an MDP is a *submodel* of another MDP if the only difference between these MDPs is that the sets of available actions of the former MDP are subsets of available actions of the latter. Define the

following subsets of the set of available actions in state x

$$\tilde{A}(x) := \{a \in A(x) \mid c(x, a) \leq \bar{\Gamma}(x)\}.$$

Let Assumptions **(LB)** hold and $\alpha \in [\alpha_0, 1)$. According to Lemma 4.3, the function $\bar{\Gamma}$ is upper semi-continuous and is therefore measurable. This implies that the graph of the mapping $\tilde{A} : \mathbb{X} \rightarrow \mathbb{A}$ is Borel, and the set $\{\mathbb{X}, \mathbb{A}, \tilde{A}(x), q, c\}$ defines a submodel of the original MDP $\{\mathbb{X}, \mathbb{A}, A(x), q, c\}$. Recall that a set-valued mapping $F : \mathbb{X} \rightarrow \mathbb{A}$ is called *closed* at x if $a_n \in F(x_n)$ and $(x_n, a_n) \rightarrow (x, a)$ imply $a \in F(x)$. A set-valued mapping $F : \mathbb{X} \rightarrow \mathbb{A}$ is called *upper semi-continuous* at 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 $y \in U(x)$. A set-valued mapping is called closed (upper semi-continuous) if it is closed (upper semi-continuous) at all $x \in \mathbb{X}$; see Definitions 4.4 and 4.5 in [20] or Definitions 1 and 2 in [17, pp. 21-23]. Moreover, a closed set-valued mapping F is upper semi-continuous if all sets, $F(x)$, are subsets of a compact set; see [20, Lemma 4.4] or the comments prior to Proposition 2 of [17, p. 23]). We make use of this fact in the proof of the following lemma.

Lemma 4.4 *If Assumptions **(LB)** and **(Wu)** hold then the mapping \tilde{A} is upper semi-continuous.*

Proof. We show first that the mapping $\tilde{A}(x)$ is closed. Suppose $a_n \in \tilde{A}(x_n)$ and $(a_n, x_n) \rightarrow (a, x)$. As previously noted, Lemma 4.3 yields that the function $\bar{\Gamma}(x)$ is upper semi-continuous. The definition of \tilde{A} implies $c(x_n, a_n) \leq \bar{\Gamma}(x_n)$. Thus,

$$c(x, a) \leq \liminf_{n \rightarrow \infty} c(x_n, a_n) \leq \limsup_{n \rightarrow \infty} c(x_n, a_n) \leq \limsup_{n \rightarrow \infty} \bar{\Gamma}(x_n) \leq \bar{\Gamma}(x),$$

where the first and fourth inequalities follow from the lower semi-continuity of c and from the upper semi-continuity of $\bar{\Gamma}$, respectively. So, $a \in \tilde{A}(x)$ as desired, and the mapping \tilde{A} is closed.

To show that \tilde{A} is upper semi-continuous, it is sufficient to show that any $x \in \mathbb{X}$ has a neighborhood U such that $\tilde{A}(U)$ is a subset of a compact set. Fix $x \in \mathbb{X}$. Since $\bar{\Gamma}$ is an upper semi-continuous function, there exists a neighborhood U of x such that $\bar{\Gamma}(y) \leq \bar{\Gamma}(x) + 1$ for all $y \in U$. The set $B = \{(y, a) \in \mathbb{X} \times \mathbb{A} \mid c(y, a) \leq \bar{\Gamma}(x) + 1\}$ is compact. Therefore, its projection on \mathbb{A} , say $B_{\mathbb{A}}$, is also compact. For each

$y \in U$ we have

$$\tilde{A}(y) = \{a \in A(y) \mid c(y, a) \leq \bar{\Gamma}(y)\} \subseteq \{a \in A(y) \mid c(y, a) \leq \bar{\Gamma}(x) + 1\} \subseteq B_{\mathbb{A}},$$

and the proof is complete. \blacksquare

Definition 4.5 We say that a submodel with action sets $A'(x)$, $x \in \mathbb{X}$, represents the original MDP if there exists $\tilde{\alpha} \in [0, 1)$ such that, under the α -discounted cost criterion with $\alpha \in [\tilde{\alpha}, 1)$, any stationary optimal policy, say $\phi_\alpha(x)$, for the original MDP belongs to this submodel, i.e. $\phi_\alpha(x) \in A'(x)$ for all $x \in \mathbb{X}$.

Proposition 4.6 If the original MDP satisfies Assumptions **(LB)** and **(Wu)** or **(Su)** then the submodel $\{\mathbb{X}, \mathbb{A}, \tilde{A}(x), q, c\}$ represents the original MDP and this submodel satisfies Assumptions **(B)** and **(W)** or **(S)**, respectively.

Proof. In view of Proposition 4.1, there exist stationary α -discounted cost optimal policies. Moreover, a stationary policy ϕ_α is α -discounted cost optimal if and only if

$$(1 - \alpha)m_\alpha + u_\alpha(x) = \min_{a \in A(x)} \{c(x, a) + \alpha \int u_\alpha(y)q(dy|x, a)\} = c(x, \phi_\alpha(x)) + \alpha \int u_\alpha(y)q(dy|x, \phi_\alpha(x)). \quad (4.2)$$

Let $\alpha \in [\alpha^*, 1)$, where α^* is defined in the paragraph following Assumption **(LB)**. Since $u_\alpha \geq 0$ we have that $c(x, \phi_\alpha(x)) \leq (1 - \alpha)m_\alpha + u_\alpha(x) = v_\alpha(x) - \alpha m_\alpha \leq \Gamma_{\alpha^*}(x) \leq \bar{\Gamma}(x)$. Therefore, $\phi_\alpha(x) \in \tilde{A}(x)$ for all $x \in \mathbb{X}$. The stationary policy ϕ_α belongs to the submodel and, thus, is optimal for the submodel. In light of this fact, the value functions v_α for the original MDP and for the submodel are equal. Thus, the submodel $\{\mathbb{X}, \mathbb{A}, \tilde{A}(x), q, c\}$ satisfies Assumption **(LB)** which implies that it satisfies Assumptions **(B)**.

Each of the Assumptions **(Wu1)** and **(Su1)** implies that **(W1)**, which is the same as **(S1)**, holds for the submodel. In addition, the following assumptions for the original MDP imply the corresponding assumptions for the submodel: **(Wu1)** implies **(W2)**, **(Wu2)** implies **(W4)**, **(Su1)** implies **(S2)**, and **(Su2)** implies **(S3)**. In view of Lemma 4.4, **(Wu1)** and **(LB)** imply **(W3)**. \blacksquare

Proposition 4.6 and Theorem 3.1 imply the main result of this section.

Theorem 4.7 *The statement of Theorem 3.1 remains valid with the Assumptions **(B)**, **(W)**, and **(S)** substituted with Assumptions **(LB)**, **(Wu)**, and **(Su)** respectively. In addition, the function u is inf-compact under **(Wu)** and is measurable under **(Su)**.*

We notice that Theorem 3.1 states that the function u is lower semi-continuous under **(Wu)**. Therefore, for any finite constant λ , the set $D_u(\lambda) = \{x \in \mathbb{X} \mid u(x) \leq \lambda\}$ is closed. Since c is inf-compact under **(Wu)**, the set $D_c(\lambda^*) = \{(x, a) \in \mathbb{X} \times \mathbb{A} \mid c(x, a) \leq \lambda^*\}$, where $\lambda^* = \lambda + w^*$, is compact. Therefore, its projection $D_c^{\mathbb{X}}(\lambda^*)$ on \mathbb{X} is compact as well. Since $D_u(\lambda) \subseteq D_c^{\mathbb{X}}(\lambda^*)$, the set $D_u(\lambda)$ is compact and u is inf-compact under **(Wu)**. In the next section, we discuss when setwise continuity holds in inventory control problems.

5 Relevance of Weak and Setwise Continuity to Inventory Control

Consider the typical dynamic state equation for inventory control models

$$x_{n+1} = x_n + a_n - D_{n+1}, \quad n = 0, 1, 2, \dots, \quad (5.1)$$

where x_n is the inventory at the end of period n , a_n is the decision how much should be ordered, and D_n is the demand during period n . The demand is assumed to be i.i.d. Let $q(dx_{n+1} \mid x_n, a_n)$ be the probability distribution of x_{n+1} for given x_n and a_n . As was mentioned in [11], (5.1) implies the weak continuity Assumptions **(Wu2)** and **(W4)**, while the setwise continuity Assumptions **(Su2)** and **(S3)** and the stronger version **(HL2)** do not hold. Indeed, let

$$x_{n+1}^k = x_n^k + a_n^k - D_{n+1}, \quad n = 0, 1, 2, \dots, \quad (5.2)$$

where $x_n^k \rightarrow x_n$ and $a_n^k \rightarrow a_n$ almost surely. Then $x_{n+1}^k \rightarrow x_{n+1}$ for each value of D_{n+1} and therefore $x_{n+1}^k \rightarrow x_{n+1}$ almost surely. The almost sure convergence $x_{n+1}^k \rightarrow x_{n+1}$ implies weak convergence; see e.g. Shiryaev [26, page 256]. Since convergence in the state implies convergence (along the actions) in the transition probabilities, Assumption **(Wu2)** holds. The following example illustrates that the setwise continuity assumptions of q may not hold. Let $D_n = 1$ (deterministically), $a_n^k = a_n + \frac{1}{k}$ and $x_n^k = x_n$. Then $q(B \mid x_n, a_n) = 1$ for $B = (-\infty, x_n + a_n - 1]$ and $q(B \mid x_n, a_n^k) = 0$ for all $k = 1, 2, \dots$

Since weak continuity holds for inventory control problems of the form (5.1) and setwise continuity may not hold, we concentrate on Assumptions **(Wu)** in the next section. The natural question is: when are transition probabilities setwise continuous for inventory control problems? The answer is that it holds for two particular cases often considered in the inventory control literature: (i) when inventory is integer and (ii) when demand is continuous.

For problems with integer inventory and integer demand the sets $\mathbb{X} = \mathbb{Z}$ and $\mathbb{A} = \mathbb{Z}^+$, the sets of all integers and non-negative integers, respectively. Any function on \mathbb{Z} is continuous and therefore the notions of weak and setwise continuity coincide. In order to analyze the continuous demand case, consider the notion of the distance in variation (also often called in total variation) between two probability measures P and Q on \mathbb{X} ,

$$\|P - Q\| = \sup_{A \in \mathcal{B}(\mathbb{X})} |P(A) - Q(A)|,$$

where $\mathcal{B}(\mathbb{X})$ is the Borel σ -field on \mathbb{X} . Since $|P(A) - Q(A)| \leq \|P - Q\|$, convergence in total variation implies setwise convergence.

The following lemma applied to (5.1) implies that the setwise continuity assumptions **(HL2)** and **(Su2)** hold when D_n are continuous random variables.

Lemma 5.1 *Let a random variable ξ have a density $f(x)$ with respect to Lebesgue integration. Consider random variables $\eta^k = y^k + \xi$, where y^k is a convergent sequence of real numbers, $y^k \rightarrow y^0$. Then the probability distributions of the random variables η^k converge in variation to the probability distribution of η^0 .*

Proof. See Appendix.

We end this section by mentioning that the above conclusions that setwise continuity holds in inventory control when either the demand distribution is continuous or the inventory is integer extends to other classic models. Consider the dynamic equations for inventory control problems with lost sales,

$$x_{n+1} = (x_n + a_n - D_{n+1})^+, \quad n = 0, 1, 2, \dots, \quad (5.3)$$

where $c^+ = \max\{0, c\}$ for a number c . Observe that the above counter-example where weak continuity holds and setwise continuity does not still applies. Moreover, the results of Lemma 5.1 that imply that setwise continuity holds in the continuous case remains valid.

6 Optimality of (s, S) Policies for Inventory Control Problems

In this section we consider a classic inventory control model with fixed ordering costs. We show that Assumptions **(LB)** and **(Wu)** hold so that the results of Theorem 4.7 yield the existence of non-randomized stationary optimal policies for average costs per unit time. We then use that theorem to prove the optimality of (s, S) policies for the average cost criterion. Before describing the model we state a technical lemma and a definition that will be used in the analysis. Consider the following assumptions.

Assumptions (C):

1. The constant $\bar{w} := \limsup_{\alpha \rightarrow 1} (1 - \alpha)m_\alpha$ is finite.
2. For any finite real number N there exists a compact subset $K_N \subseteq \mathbb{X}$ such that $c(x) \geq N$ for all $x \in \mathbb{X} \setminus K_N$ where $c(x) = \inf_{a \in A(x)} c(x, a)$.

Note that **(C1)** actually follows from Assumption **(G)** (see [24, Lemma 1.2]), but is included here for completeness. Let $M(\alpha) := \{x \in \mathbb{X} | v_\alpha(x) = m_\alpha\}$. The next result is similar to [11, Lemma 6.1] and, as mentioned in [11], to [5, Lemma 4] and [24, Lemma 6.6].

Lemma 6.1 *Let Assumptions **(C1)** and **(C2)** hold. Then there exists $\alpha_0 < 1$ and a compact subset $\mathcal{K} \subseteq \mathbb{X}$ such that $M(\alpha) \subseteq \mathcal{K}$ for all $\alpha \in [\alpha_0, 1)$.*

Proof. Consider $N > \bar{w} + 1 = \limsup_{\alpha \rightarrow 1} (1 - \alpha)m_\alpha + 1$. Therefore, there exists $\alpha_0 < 1$ such that $N/(1 - \alpha) > m_\alpha + (1 - \alpha)^{-1}$ for all $\alpha \in [\alpha_0, 1)$. This formula is identical to (6.2) in [11] and the rest of the proof coincides with the the proof of Lemma 6.1 in [11] following (6.2) there. ■

The model has the following decision-making scenario: 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. Assume the unmet demand is backlogged and the cost of inventory held or backlogged (negative inventory) is modeled as a convex function. The demand and the order quantity are assumed to be non-negative. The dynamics of the system are defined by (5.1). Let

- $\alpha \in (0, 1)$ be the discount factor,
- $K \geq 0$ be a fixed ordering cost,
- $c > 0$ be the per unit ordering cost,
- $h(\cdot)$ denote the holding/backordering cost per period; convex, non-negative, takes finite values, and $h(x) \rightarrow \infty$ as $|x| \rightarrow \infty$,
- $\{D_n, n \geq 0\}$ be a sequence of i.i.d. random variables where $D_n \geq 0$ (almost surely) represents demand in the n^{th} period. We assume that $\mathbb{E} h(x - D) < \infty$ for all $x \in \mathbb{R}$ and $P(D > 0) > 0$, where D is a random variable with the same distribution as D_n .

Without loss of generality, assume that $h(0) = 0$. The fact that $P(D > 0) > 0$ avoids the trivial case. For example, if $D = 0$ almost surely then the policy that never orders when the inventory level is non-negative and orders up to zero when the inventory level is negative, is optimal under the average cost criterion. Note that the finiteness of $\mathbb{E} h(x - D)$ and the assumed properties of the function h imply that $\mathbb{E} |D| < \infty$.

The cost function for the model is

$$c(x, a) = K1_{\{a \neq 0\}} + ca + \mathbb{E} h(x + a - D).$$

Note that the problem is posed with $\mathbb{X} = \mathbb{R}$. However, if the demand and action sets are integer or with probability 1 on a lattice, the model can be restated with $\mathbb{X} = \mathbb{Z}$; see Remark 6.10.

Consider the policy ϕ that orders up to the level 0 if the inventory level is non-positive and does nothing otherwise. Then for $x \leq 0$

$$w^\phi(x) \leq K + c\mathbb{E} D + \mathbb{E} h(-D) < \infty.$$

That is, Assumption **(G)** holds.

Moreover, since $h(x) \rightarrow \infty$ when $|x| \rightarrow \infty$, **(C2)** holds. This coupled with the observation that **(C1)** holds (via Assumption **(G)**) implies that the results of Lemma 6.1 hold for the inventory control model. Consider the renewal process

$$N(t) := \sup\{n | \mathbf{S}_n \leq t\}. \quad (6.1)$$

where $\mathbf{S}_0 = 0$ and $\mathbf{S}_n = \sum_{j=1}^n D_j$ for $n > 0$. Observe that $\mathbb{E} N(t) < \infty$ for each $0 \leq t < \infty$; Resnick [22, Theorem 3.3.1]. Thus, Wald's identity yields that for any $0 \leq y < \infty$

$$\mathbb{E} \sum_{j=1}^{N(y)+1} D_j = \mathbb{E}(N(y) + 1) \mathbb{E} D_1 < \infty. \quad (6.2)$$

We next state a useful lemma.

Lemma 6.2 *For fixed initial state x*

$$E_y(x) := \mathbb{E} h(x - \mathbf{S}_{N(y)+1}) < \infty, \quad (6.3)$$

where $0 \leq y < \infty$.

Proof. Define

$$h^*(x) := \begin{cases} h(x) & \text{for } x \leq 0, \\ 0 & \text{for } x > 0. \end{cases}$$

Observe that it suffices to show that

$$E_y^*(x) := \mathbb{E} h^*(x - \mathbf{S}_{N(y)+1}) < \infty. \quad (6.4)$$

Indeed, for $Z = x - \mathbf{S}_{N(y)+1}$,

$$E_y(x) = \mathbb{E} 1\{Z \leq 0\} h^*(Z) + \mathbb{E} 1\{Z > 0\} h(Z) \leq E_y^*(x) + h(x).$$

To show that $E_y^*(x)$ we shall prove the inequality

$$\mathbb{E} h^*(x - \mathbf{S}_{N(y)+1}) \leq (1 + \mathbb{E} N(y)) \mathbb{E} h^*(x - y - D_1). \quad (6.5)$$

If (6.5) holds, the assumptions on h imply (6.4). Define the function $f(z) = h^*(x - y - z)$. This function is nondecreasing and convex. Since f is convex, its derivative exists almost everywhere. Denote the excess of $N(y)$ by $R(y) := \mathbf{S}_{N(y)+1} - y$. According to [14, p. 59]

$$P\{R(y) > t\} = 1 - F(y + t) + \int_0^y (1 - F(y + t - s))dU(s),$$

where $U(s) = \mathbb{E} N(s)$ is the renewal function. Thus,

$$\mathbb{E} h^*(x - \mathbf{S}_{N(y)+1}) = \mathbb{E} h^*(x - y - R(y)) = \mathbb{E} f(R(y)) = \int_0^\infty f'(t)P\{R(y) > t\}dt = J_1 + J_2, \quad (6.6)$$

where $J_1 = \int_0^\infty f'(t)(1 - F(y + t))dt$, $J_2 = \int_0^\infty f'(t) \left(\int_0^y (1 - F(y + t - s))dU(s) \right) dt$, and the third equality in (6.6) holds according to [9, p. 263]. Note that since F is non-decreasing,

$$J_1 \leq \int_0^\infty f'(t)(1 - F(t))dt = \mathbb{E} f(D_1) = \mathbb{E} h^*(x - y - D_1), \quad (6.7)$$

where the first equality follows from [9, p. 263]. Similarly, by applying Fubini's theorem

$$\begin{aligned} J_2 &= \int_0^y \left(\int_0^\infty f'(t)(1 - F(y + t - s))dt \right) dU(s) \\ &\leq \int_0^y \left(\int_0^\infty f'(t)(1 - F(t))dt \right) dU(s) = \mathbb{E} f(D_1) \mathbb{E} U(y) = \mathbb{E} h^*(x - y - D_1) \mathbb{E} N(y). \end{aligned} \quad (6.8)$$

Combining (6.6)-(6.8) yields (6.5) and the lemma is proven. ■

The following result states that the inventory control problem satisfies the remaining assumptions guaranteeing the existence of stationary optimal policies and the validity of the optimality inequalities.

Proposition 6.3 *In the inventory control model, Assumptions (Wu) and (LB) hold. Therefore, the results of Proposition 4.1 and Theorem 4.7 hold.*

Proof. Note that the cost function is inf-compact since $c(x, a) \rightarrow \infty$ as $a \rightarrow \infty$ or $|x| \rightarrow \infty$ and is convex in a on $(0, \infty)$ and thus continuous in a (except $a = 0$ where it is lower semi-continuous). The fact that the

transition kernel is continuous with respect to weak convergence was discussed in Section 5. Thus, **(Wu1)** and **(Wu2)** hold.

To show **(LB)** we need to show the local boundedness of r_{α_0} . Let x_α be any inventory level such that $v_\alpha(x_\alpha) = m_\alpha$. Fix the initial state x and let α_0 be defined as in Lemma 6.1. That is to say that for $\alpha \in [\alpha_0, 1)$, x_α is trapped on the compact set $\mathcal{K} := [x_L, x_U]$. Since increasing x_U only expands \mathcal{K} , without loss of generality assume that $x_U > x$. For any $\alpha \in [\alpha_0, 1)$ consider two cases: $x \leq x_\alpha$ and $x > x_\alpha$. For $x \leq x_\alpha$, suppose ϕ is a stationary policy that immediately orders up to level x_α plus orders whatever amount a stationary optimal policy for the discount factor α would order in x_α . From then on it proceeds to follow the optimal policy. We have the following sequence of inequalities

$$v_\alpha(x) - m_\alpha \leq v_\alpha^\phi(x) - m_\alpha \leq K + c(x_U - x). \quad (6.9)$$

Suppose now that $x > x_\alpha$ and that ϕ does not order until the total demand is greater than $x - x_L$. In this case, the difference in costs between a process (Process 1) starting in x that follows ϕ and one that starts in x_α (and follows the optimal policy) can be broken into 3 parts; the holding costs accrued before the inventory of Process 1 moves below x_L , the holding cost accrued in the step that takes the inventory position below x_L and the ordering costs accrued to move the position to x_α .

Since h is convex, $\max\{h(x_L), h(x_U)\} \geq h(y)$ for all $y \in [x_L, x_U]$ so that the expected total discounted holding costs accrued before the inventory position falls below x_L is bounded by $\mathbb{E}N(x_U - x_L) \max\{h(x_L), h(x_U)\}$. The inventory position immediately prior to the order being placed is then $x - \mathbf{S}_{N(x-x_L)+1}$. Since $x_U > x$ and h is convex, the expected total discounted holding cost is bounded by

$$E(x) = \max\{E_{x_U-x_L}(x), h(x_U)\},$$

where $E_{x_U-x_L}(x)$ is defined in (6.3) and the finiteness of $E(x)$ follows from Lemma 6.3. The expected discounted order cost is bounded by

$$K + c(x_\alpha - [x - \mathbb{E}(N(x - x_L) + 1) \mathbb{E}D_1]) \leq K + c(x_U + \mathbb{E}(N(x_U - x_L) + 1) \mathbb{E}D_1).$$

Combining these upper bounds yields

$$\begin{aligned} v_\alpha(x) - m_\alpha &\leq \mathbb{E} N(x_U - x_L) \max\{h(x_L), h(x_U)\} + E(x) \\ &\quad + K + c(x_U + \mathbb{E}(N(x_U - x_L) + 1) \mathbb{E} D_1). \end{aligned} \quad (6.10)$$

Note that the right hand side of (6.10) is continuous in x . Indeed, the first term is constant in x , the second term is a maximum of a constant and a convex function defined on the real line, and the last term is constant. Combining (6.9) and (6.10) we see that $u_\alpha(x)$ is bounded above by a finite continuous in x function. This implies **(LB)**. ■

Recall from the proof of Proposition 4.1 that the functions

$$J_n(x, a) := K1_{\{a \neq 0\}} + ca + \mathbb{E}[h(x + a - D) + \alpha v_{n,\alpha}(x + a - D)], \quad (6.11)$$

$$J(x, a) := K1_{\{a \neq 0\}} + ca + \mathbb{E}[h(x + a - D) + \alpha v_\alpha(x + a - D)], \quad (6.12)$$

$n = 0, 1, \dots$, are inf-compact and the optimality equations (2.4) and (2.6) can be written with minimums instead of infimums. Thus,

$$v_{n+1,\alpha}(x) = \min_{0 \leq a < \infty} \{J_n(x, a)\}, \quad (6.13)$$

$$v_\alpha(x) = \min_{0 \leq a < \infty} \{J(x, a)\}. \quad (6.14)$$

Similarly, from Theorem 4.7 the ACOI are

$$w + u(x) \geq \min_{0 \leq a < \infty} \{K1_{\{a \neq 0\}} + ca + \mathbb{E} h(x + a - D) + \mathbb{E} u(x + a - D)\}. \quad (6.15)$$

The sets of equations (6.14) and (6.15) can be rewritten

$$v_\alpha(x) = \min\{\min_{a>0}[K + G_\alpha(x + a)], G_\alpha(x)\} - cx, \quad (6.16)$$

$$w + u(x) \geq \min\{\min_{a>0}[K + H(x + a)], H(x)\} - cx, \quad (6.17)$$

where

$$G_\alpha(x) := cx + \mathbb{E} h(x - D) + \alpha \mathbb{E} v_\alpha(x - D), \quad (6.18)$$

$$H(x) := cx + \mathbb{E} h(x - D) + \mathbb{E} u(x - D).$$

Recall the following classic definition.

Definition 6.4 A real-valued function f is called K -convex, $K \geq 0$, if for any $x \leq y$ and for any $\lambda \in [0, 1]$,

$$f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y) + \lambda K.$$

The next result is a version of Bertsekas [1, Lemma 4.2.1(c-d)] for lower semi-continuous K -convex functions.

Proposition 6.5 The following results hold:

1. If $g(y)$ is K -convex and D is a random variable, then $\mathbb{E} g(y - D)$ is also K -convex provided $\mathbb{E} |g(y - D)| < \infty$ for all y .
2. Suppose g is a lower semi-continuous K -convex function such that $g(x) \rightarrow \infty$ as $|x| \rightarrow \infty$. Let

$$S \in \operatorname{argmin}_{x \in \mathbb{R}} \{g(x)\}, \quad (6.19)$$

$$s = \inf \{x \leq S \mid g(x) \leq K + g(S)\}. \quad (6.20)$$

Then

- (a) $g(S) \leq g(x)$ for all $x \in \mathbb{R}$,
- (b) $g(S) + K < g(x)$ for all $x < s$,
- (c) $g(x)$ is decreasing on $(-\infty, s)$,
- (d) $g(x) \leq g(S) + K$ for all x such that $s \leq x \leq S$,
- (e) $g(x) \leq g(z) + K$ for all $S < x \leq z$.

Proof. We prove only 2(d). The others follow in the same way as Lemma 4.2.1 in Bertsekas [1]. If $x = s$ or S the result is trivial. Suppose $s < x < S$. By the definition of lower semi-continuity there exists $\delta > 0$ such that $g(x + \delta) > g(x) - \epsilon \left(\frac{\delta}{S-x} \right)$ for arbitrary $\epsilon > 0$. However, K-convexity implies

$$K + g(S) \geq g(x) + \frac{S-x}{\delta} [g(x + \delta) - g(x)] > g(x) - \epsilon.$$

Since ϵ is arbitrary the result follows. ■

Consider the discounted cost problem and suppose G_α is K-convex, lower semi-continuous and approaches infinity as $|x| \rightarrow \infty$. If we define S_α and s_α by (6.19) and (6.20) with g replaced by G_α , Proposition 6.5 parts 2(b) and (c), along with the DCOE imply that it is optimal to order up to S_α when $x < s_\alpha$. Parts 2(d) and (e) imply that it is optimal not to order when $s_\alpha \leq x$.

Lower semi-continuity of G_α (recall (6.18)) follows from the convexity of h , the lower semi-continuity of v_α , and the weak continuity of the transition probabilities. In order to show that G_α is K-convex, note that v_α is K-convex since it is a limit K-convex functions $v_{n,\alpha}$; see Bertsekas [1, Section 4.2]. The next result along with the first result of Proposition 6.5 completes the proof that G_α is K-convex.

Proposition 6.6 $\mathbb{E} v_\alpha(x - D) < \infty$ for each $x \in \mathbb{R}$ and $\alpha \in [0, 1)$.

Proof. For $x \leq 0$ suppose that the policy ϕ orders up to zero. Then

$$v_\alpha(x) \leq v_\alpha^\phi(x) \leq K - cx + \frac{\alpha(K + c\mathbb{E}D + \mathbb{E}h(-D))}{1 - \alpha}.$$

Letting $B := \frac{\alpha(K + c\mathbb{E}D + \mathbb{E}h(-D))}{1 - \alpha}$ we have $\mathbb{E} v_\alpha(x - D) \leq K - c\mathbb{E}(x - D) + B < \infty$. For $x > 0$, let $M = \sup_{y \in [0, x]} \{v_\alpha(y)\}$. To see that M is finite, we apply Proposition 6.3 and consider α_0 such that r_{α_0} is locally bounded. Thus since v_α is non-decreasing in α for $\alpha \leq \alpha_0$,

$$v_\alpha(x) - m_{\alpha_0} \leq v_{\alpha_0}(x) - m_{\alpha_0} \leq r_{\alpha_0}(x) \leq \bar{r}_{\alpha_0}(x),$$

where \bar{r}_{α_0} is as defined in Lemma 4.3 for $f = r_{\alpha_0}$. Taking the supremum over $[0, x]$ yields

$$\sup_{y \in [0, x]} \{v_\alpha(y)\} \leq \sup_{y \in [0, x]} \{\bar{r}_{\alpha_0}(y)\} + m_{\alpha_0} < \infty,$$

where the finiteness of the right hand side follows from the upper semi-continuity of \bar{r}_{α_0} shown in Lemma 4.3. A similar argument holds for $\alpha \geq \alpha_0$ since $v_\alpha - m_\alpha \leq r_{\alpha_0}$. Thus, for $x > 0$

$$\mathbb{E} v_\alpha(x - D) \leq K + c \mathbb{E} D + B + M < \infty$$

as desired. ■

Since v_α is non-negative we have that $G_\alpha(x) \rightarrow \infty$ as $x \rightarrow \infty$. However, it is possible that $G_\alpha(x)$ does not tend to ∞ as $x \rightarrow -\infty$. In what follows we provide conditions under which convergence to ∞ (in both directions) is guaranteed.

Lemma 6.7 *Define $G_{N,\alpha}(x)$ by (6.18) with v_α replaced with $v_{N,\alpha}$. Consider the following cases for the convergence of $G_{N,\alpha}(x)$ and $G_\alpha(x)$.*

1. *Suppose there exists $z < y$ such that*

$$\frac{\mathbb{E}[h(y - D) - h(z - D)]}{y - z} < -c. \tag{6.21}$$

Then $G_\alpha(x)$ and $G_{N,\alpha}(x) \rightarrow \infty$ as $|x| \rightarrow \infty$ for all $\alpha \in [0, 1)$ and for all $N \geq 0$.

2. *There exists $\alpha^* \in [0, 1)$ and $k \geq 0$ such that for each $\alpha \in [\alpha^*, 1)$ and for each $N \geq k$, $G_{N,\alpha}(x) \rightarrow \infty$ as $|x| \rightarrow \infty$, and therefore, $G_\alpha(x) \rightarrow \infty$ as $|x| \rightarrow \infty$.*

Proof. We prove the first assertion for the finite horizon by induction. Obviously, $G_{N,\alpha}(x) \rightarrow \infty$ as $x \rightarrow \infty$. We show that the result continues to hold when $x \rightarrow -\infty$. Suppose $z < y$ satisfy (6.21). Rearrange terms to get

$$cy + \mathbb{E} h(y - D) < cz + \mathbb{E} h(z - D).$$

Thus, $G_{\alpha,0}(z) > G_{\alpha,0}(y)$. Since $G_{\alpha,0}$ is convex $G_{\alpha,0}(x) \rightarrow \infty$ as $x \rightarrow -\infty$ and the result holds for $N = 0$. Assume that it holds for N . Since $v_{N,\alpha}$ is lower semi-continuous and q is weakly continuous, $G_{\alpha,N}(x)$ is lower semi-continuous. This together with the inductive hypothesis implies the existence of a minimum of

$G_{\alpha,N}(x)$, say $S_{\alpha,N}$. Thus, there exists L_N such that $v_{\alpha,N+1}(x) = K + G_{\alpha,N}(S_{\alpha,N}) - cx$ for all $x \leq L_N$. That is, $v_{\alpha,N+1}(x) \rightarrow \infty$ as $x \rightarrow -\infty$. Since

$$G_{\alpha,N+1}(x) = G_{\alpha,0}(x) + \mathbb{E} v_{\alpha,N+1}(x - D)$$

the result holds for all N . Since $G_{\alpha,N}$ is non-decreasing in N , letting $N \rightarrow \infty$ yields the result for $G_\alpha(x)$.

To prove the second assertion, consider the special case of $K = 0$. Denote the corresponding functions of K with a superscript K . For example G_α^0 corresponds to G_α with $K = 0$. Since the function $v_\alpha^K(x)$ is non-decreasing in K , $G_\alpha^K(x) \geq G_\alpha^0(x)$ for all x and $K \geq 0$. Therefore, it is sufficient to prove the existence of $\alpha^* \in [0, 1)$ such that $G_\alpha^0(x) \rightarrow \infty$ as $x \rightarrow -\infty$ and $\alpha \in [\alpha^*, 1)$.

Since the functions v_α^K and G_α^K are K -convex, the functions v_α^0 and G_α^0 are convex. We show by contradiction that there exists $\alpha^* \in [0, 1)$ such that G_α^0 is decreasing on an interval $(-\infty, M_\alpha]$ for some $M_\alpha > -\infty$ when $\alpha \in [\alpha^*, 1)$. Suppose this is not the case. For $K = 0$, (6.16) can be written

$$v_\alpha^0(x) = \inf_{a \geq 0} \{G_\alpha^0(x + a)\} - cx. \quad (6.22)$$

If a constant M_α does not exist for some $\alpha \in (0, 1)$ then the convexity and nonnegativity of $G_\alpha^0(x)$ imply that the policy that never orders is optimal for the discount factor α . If there is no α^* with the described properties, Theorem 4.7 implies that this policy is average cost optimal as well. On the other hand, the average cost for this policy is ∞ while the average cost of the policy that orders up to the level 0 is $\mathbb{E} h(-D) + c \mathbb{E} D < \infty$: a contradiction. Since G_α^0 is convex and becomes (strictly) decreasing as x approaches $-\infty$ for $\alpha \in [\alpha^*, 1)$, $G_\alpha^0(x) \rightarrow \infty$ as $x \rightarrow -\infty$ when $\alpha \in [\alpha^*, 1)$.

To show that the result holds for the N -horizon problem for N sufficiently large, suppose α is such that $G_\alpha^0(x) \rightarrow \infty$ as $x \rightarrow -\infty$. Choose $x < 0$ such that $G_\alpha^0(x) > G_\alpha^0(0)$. Since $G_{N,\alpha}^0(x) \uparrow G_\alpha^0(x)$, there exists k such that $N \geq k$ implies $G_{N,\alpha}^0(x) > G_\alpha^0(0) \geq G_{N,\alpha}^0(0)$. The convexity of $G_{N,\alpha}^0$ now implies $G_{N,\alpha}^0(x) \rightarrow \infty$ as $x \rightarrow -\infty$. Since $G_{N,\alpha}^0 \leq G_{N,\alpha}$, the result follows. ■

Definition 6.8 Let s_n and S_n be real numbers such that $s_n \leq S_n$, $n = 0, 1, \dots$. Suppose x_n denotes the current inventory level at decision epoch n . A policy is called an (s_n, S_n) policy at step n if it orders up to

the level S_n if $x_n < s_n$, does not order when $x_n > s_n$ and either orders up to S_n or does not order when $x_n = s_n$. A policy is called an (s_n, S_n) policy if it is an (s_n, S_n) policy at all steps $n = 0, 1, \dots$. If $s_n = s$ and $S_n = S$ for all n , the policy is called an (s, S) policy.

Let $S_{N,\alpha}$ and $s_{N,\alpha}$ be defined by (6.19) and (6.20) with g replaced by $G_{N,\alpha}$. Similarly, consider S_α and s_α defined by (6.19) and (6.20) with g replaced by G_α . The following theorem is the main result of this section.

Theorem 6.9 *The following hold for the inventory control problem.*

1. *Consider the N -horizon expected total discounted cost criterion. If the assumption of Statement 1 of Lemma 6.7 holds then the $(s_{N-n,\alpha}, S_{N-n,\alpha})$ policy, $n = 0, \dots, N-1$, is optimal. Otherwise, for $\alpha \in [\alpha^*, 1)$ and $N \geq k$, where the existences of α^* and k are stated in Statement 2 of Lemma 6.7, there exists an optimal Markov policy that is an $(s_{N-n,\alpha}, S_{N-n,\alpha})$ policy at steps $n = 0, \dots, N-k-1$.*
2. *Consider the infinite-horizon expected total discounted cost criterion. If the assumption of Statement 1 of Lemma 6.7 holds then the (s_α, S_α) policy is optimal for any $\alpha \in [0, 1)$. If this assumption does not hold, the (s_α, S_α) policy is optimal when $\alpha \in [\alpha^*, 1)$. In addition, any $(\tilde{s}_\alpha, \tilde{S}_\alpha)$ policy is optimal where \tilde{s}_α and \tilde{S}_α are any limit points of the sequences $s_{N,\alpha}$ and $S_{N,\alpha}$ respectively, $N = 0, 1, \dots$.*
3. *Consider the infinite-horizon average cost per unit time criterion. For each $\alpha \in [\alpha^*, 1)$, in view of Statement 2, there exists an optimal (s'_α, S'_α) policy for the discounted cost criterion. For example, it is possible to select $s'_\alpha = s_\alpha$ and $S'_\alpha = S_\alpha$. For any selection of s'_α and S'_α , the inequalities $-\infty < \liminf_{\alpha \rightarrow 1} s'_\alpha < \limsup_{\alpha \rightarrow 1} S'_\alpha < \infty$ hold. Therefore, there exists a subsequence $\alpha_m \uparrow 1$ and two finite numbers s and S such that $s'_{\alpha_m} \rightarrow s$ and $S'_{\alpha_m} \rightarrow S$. This (s, S) policy is average cost optimal.*

Proof. We first prove the existence of optimal (s, S) policies in the finite and infinite horizon expected total discounted cost cases. We observe that $G_{N,\alpha}$ are K -convex lower semi-continuous functions; see the arguments in the paragraph preceding Proposition 6.6. Lemma 6.7 implies that $G_{N,\alpha}(x)$ approaches ∞ as

$|x| \rightarrow \infty$. The proof for the finite horizon case follows directly from Proposition 6.5. Similar arguments hold in the infinite horizon case since G_α satisfies the hypotheses of Proposition 6.5.

The existence of a limit point of $(s_{N,\alpha}, S_{N,\alpha})$ such that the resulting $(\tilde{s}_\alpha, \tilde{S}_\alpha)$ policy is optimal follows, for example, from the remarks on p. 149 of [1]. For a discount factor $\alpha \in [\alpha^*, 1)$ we fix a stationary optimal policy ϕ_α that is the (s'_α, S'_α) policy. According to Theorem 4.7, for any sequence of discount factors $\alpha(k) \rightarrow 1$ and for any $x \in \mathbb{X}$ there is a subsequence α_m of $\alpha(k)$ and a sequence of states $x_m \rightarrow x$ such that the discounted cost optimal policy $\phi_m(x_m) \rightarrow \phi(x)$ as $\alpha_m \rightarrow 1$, where ϕ is an average cost optimal policy. Moreover, in view of Theorem 4.7, this policy satisfies (6.15) with u defined in (3.2). Select $\alpha(k) \geq \alpha^*$ in a way that there are limits $s = \lim_{k \rightarrow \infty} s'_{\alpha(k)}$ and $S = \lim_{k \rightarrow \infty} S'_{\alpha(k)}$. Note that in the limiting (s, S) policy both s and S are finite; otherwise w^* is infinite and Assumption (G) does not hold. Thus, in view of Theorem 4.7, this (s, S) policy is optimal for average costs per unit time. ■

Remark 6.10 *For the inventory control problem, we have considered an MDP with $\mathbb{X} = \mathbb{R}$ and $A(x) = \mathbb{R}^+ = [0, \infty)$ for each $x \in \mathbb{X}$. However, if the demand takes only integer values, for many problems it is natural to consider $\mathbb{X} = \mathbb{Z}$ and $A(x) = \mathbb{Z}^+$. Therefore, if the demand is integer, we have two MDPs for the inventory control problems: an MDP with $\mathbb{X} = \mathbb{R}$ and an MDP with $\mathbb{X} = \mathbb{Z}$. Though the first MDP yields potentially lower costs, its implementation may not be reasonable for some applications because it may prescribe to order up to a non-integer inventory level. However, all of the results of this paper hold for the second representation, when the state space is integer, with a minor modification that the action sets are integer as well. In addition, in this case, for (s_n, S_N) and (s, S) policies, it is possible to fix a version whether we order at s (s_n) up to the level S (S_n) or do not order at s (s_n). One more note is that the case when the possible demand is proportional to some number $d \in \mathbb{R}$ is similar to the integer demand case. One can consider an MDP with an integer state space for this case as well.*

7 Conclusions

In this paper we presented two sets of conditions that lead to the convergence of the optimal discounted cost value function and policies to those in the average cost case. The results of Schäl [24] play an integral

role as the basis for the analysis but we do not assume that the action sets are compact. We establish that one of the sets of the sufficient conditions hold for classic inventory control problems and therefore stationary optimal policies exist for average cost inventory control models. In addition, optimal discounted cost policies converge to optimal undiscounted cost policies. This convergence implies the optimality of (s, S) policies for average cost inventory control policies.

We believe that additional studies of MDPs will lead to straightforward proofs of additional properties of inventory control problems: (a) the value functions u are continuous and K -convex, (b) the optimality equalities hold, and (c) both versions of (s, S) policies are optimal when the state space is \mathbb{R} , namely the policy that does not order and the policy that orders up to the level S when the current inventory level is s . Other natural research directions are to investigate problems with possibly negative demand and to study other inventory control models including models that combine pricing and inventory decisions; see [6, 7] and the references therein.

A Appendix: Proof of Lemma 5.1

Without loss of generality set $y^0 = 0$. Then $\eta^0 = \xi$ and η^k has the density $f(x - y^k)$. Let P^k be the probability distribution of η^k on $(-\infty, \infty)$. According to Shiryaev [26, page 362],

$$\|P^k - P^0\| = \int_{-\infty}^{\infty} |f(x) - f(x - y^k)| dx. \quad (\text{A.1})$$

Fix $\epsilon > 0$ and consider $K > 0$ such that

$$\int_{-K}^K f(x) dx \geq 1 - \frac{\epsilon}{8}.$$

Consider integers k large enough for $|y^k| \leq 1$ and select any number $K^* \geq K + 1$. Then

$$0 \leq \int_{-\infty}^{\infty} |f(x) - f(x - y^k)| dx - \int_{-K^*}^{K^*} |f(x) - f(x - y^k)| dx \leq \frac{\epsilon}{4}. \quad (\text{A.2})$$

Fix $G > 0$ such that

$$\int_{-\infty}^{\infty} f(x)I\{f(x) > G\}dx \leq \frac{\epsilon}{8} \quad (\text{A.3})$$

and define the bounded function $f_G(x) = f(x)I\{f(x) \leq G\} + GI\{f(x) > G\}$. Then, in view of (A.3),

$$\int_{-K^*}^{K^*} |f(x) - f(x - y^k)|dx \leq \int_{-K^*}^{K^*} |f_G(x) - f_G(x - y^k)|dx + \frac{\epsilon}{4}. \quad (\text{A.4})$$

From (A.1), (A.2), and (A.4), we have that

$$\|P^k - P^0\| \leq \int_{-K^*}^{K^*} |f_G(x) - f_G(x - y^k)|dx + \frac{\epsilon}{2}. \quad (\text{A.5})$$

Lusin's theorem [19, page 108] state that for any real-valued measurable function f and for any $\varepsilon > 0$ there exists a continuous function g such that the Lebesgue measure of the set $\{f(x) \neq g(x)\}$ is not greater than ε . Consider the density function f and select a continuous function g such that Lusin's theorem holds for $\varepsilon = \epsilon/(8G)$.

Define the non-negative bounded continuous function $g_G(x) = g(x)I\{0 \leq g(x) \leq G\} + GI\{g(x) > G\}$. Since $\{f_G(x) \neq g_G(x)\} \subseteq \{f(x) \neq g(x)\}$,

$$\begin{aligned} \int_{-K^*}^{K^*} |f_G(x) - f_G(x - y^k)|dx &\leq \int_{-K^*}^{K^*} |g_G(x) - g_G(x - y^k)|dx + \int_{-K^*}^{K^*} |g_G(x) - f_G(x)|dx \\ &+ \int_{-K^*}^{K^*} |g_G(x - y^k) - f_G(x - y^k)|dx \leq \int_{-K^*}^{K^*} |g_G(x) - g_G(x - y^k)|dx + \frac{\epsilon}{4}. \end{aligned} \quad (\text{A.6})$$

Since the function g_G is continuous, it is uniformly continuous on the interval $[-(K^* + 1), (K^* + 1)]$. Thus, there exists an integer N such that $|g_G(x) - g_G(x - y_k)| \leq \epsilon/(8K^*)$ when $k \geq N$ and $-K^* \leq x \leq K^*$.

This bound and (A.6) imply that for $k \geq N$

$$\int_{-K^*}^{K^*} |f_G(x) - f_G(x - y^k)|dx \leq \frac{\epsilon}{2}.$$

The last inequality and (A.5) imply that $\|P^k - P^0\| \leq \epsilon$ when $k \geq N$. ■

Acknowledgements The authors thank Irwin Schochetman for providing several references on upper semi-continuity of a set-valued function. Research of the first author was partially supported by NSF grant DMI-0300121.

References

- [1] D.P. Bertsekas, *Dynamic Programming and Optimal Control*, second edition, vol. 1, Athena Scientific, Belmont, MA, 2000.
- [2] D.P. Bertsekas and S.E. Shreve, *Stochastic Optimal Control: the Discrete-Time Case*, Athena Scientific, Belmont, MA, 1996.
- [3] D. Beyer and S.P. Sethi, *The classical average-cost inventory models of Iglehart and Veinott-Wagner revisited*, *Journal of Optimization and Applications* **101** (1999), no. 3, 523–555.
- [4] R. Cavazos-Cadena, *A counterexample on the optimality equation in markov decision chains with the average cost criterion*, *Systems & Control Letters* **16** (1991), 387–392.
- [5] R. Cavazos-Cadena and L.I. Sennott, *Comparing recent assumptions for the existence of average optimal stationary policies*, *Operations Research Letters* **11** (1992), 33–37.
- [6] X. Chen and D. Simchi-Levi, *Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case*, *Operations Research* **52** (2004), 887–896.
- [7] X. Chen and D. Simchi-Levi, *Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The infinite horizon case*, *Mathematics of Operations Research* **29** (2004), 698–723.
- [8] E.A. Dynkin and A.A. Yushkevich, *Controlled Markov Processes*, Springer-Verlag, New York, 1979.
- [9] E.A. Feinberg, *A generalization of ‘expectation equals reciprocal of intensity’ to non-stationary exponential distributions*, *J. Appl. Prob.* **31** (1994), 262–267.

- [10] E.A. Feinberg, *Total reward criteria*, Handbook of Markov Decision Processes (E.A. Feinberg and A. Shwartz, eds.), Kluwer, Boston, 2002, pp. 173–207.
- [11] E.A. Feinberg and M.E. Lewis, *Optimality of four-threshold policies in inventory systems with customer returns and borrowing/storage options*, Probability in the Engineering and Informational Sciences **19** (2004), 45–71.
- [12] E. Fernández-Gaucherand, A. Arapostathis, and S.I. Marcus, *Convex stochastic control problems*, Proceedings of the 31st Conference on Decision and Control, IEEE, December 1992, pp. 2179–2180.
- [13] H.S. Gaskill and P.P. Narayanaswami, *Elements of Real Analysis*, Prentice-Hall, Upper Saddle River, New Jersey, 1998.
- [14] A. Gut, *Stopped Random Walks*, Springer-Verlag, New York, 1988.
- [15] O. Hernández-Lerma and J.B. Lasserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, Springer, New York, 1996.
- [16] D.P. Heyman and M.J. Sobel, *Stochastic Models in Operations Research*, vol. II, McGraw-Hill, New York, 1984.
- [17] W. Hildenbrand, *Core and Equilibria of a Large Economy*, Princeton Studies in Mathematical Economics, Princeton University Press, Princeton, 1974.
- [18] D.L. Iglehart, *Dynamic programming and stationary analysis of inventory problems*, Office of Naval Research Monographs on Mathematical Methods in Logistics, ch. 1, pp. 1–31, Stanford University Press, Stanford, California, 1963.
- [19] A.S. Kechris, *Classical Descriptive Set Theory*, Springer-Verlag, New York, 1994.
- [20] H. Nikaido, *Convex Structures and Economic Theory*, Mathematics in Science and Engineering, vol. 51, Academic Press, New York, 1968.

- [21] E. Porteus, *Stochastic Inventory Theory*, Handbooks in Operations Research and Management Science (Daniel Heyman and Matthew Sobel, eds.), vol. 2, Elsevier Science Publishers, Amsterdam, 1990, pp. 605–652.
- [22] S.I. Resnick, *Adventures in Stochastic Processes*, Birkhäuser, Boston, 1992.
- [23] M. Schäl, *Conditions for optimality in dynamic programming and for the limit of n -stage optimal policies to be optimal*, *Z. Wahrsch. Verw. Gebiete* **32** (1975), 179–196.
- [24] M. Schäl, *Average optimality in dynamic programming with general state space*, *Mathematics of Operations Research* **18** (1993), no. 1, 163–172.
- [25] L.I. Sennott, *Stochastic Dynamic Programming and the Control of Queueing Systems*, John Wiley and Sons, Inc., New York, 1999.
- [26] A.N. Shiryaev, *Probability*, second edition, Springer, New York, 1996.
- [27] A.F. Veinott, Jr. and H.M. Wagner, *Computing optimal (s,S) policies*, *Management Science* **11** (1965), no. 5, 525–552.
- [28] Y. Zheng, *A simple proof for the optimality of (s,S) policies in infinite horizon inventory systems*, *Journal of Applied Probability* **28** (1991), 802–810.