Adv. Appl. Prob. **50,** 204–230 (2018)
*doi:*10.1017/apr.2018.10

VANISHING DISCOUNT APPROXIMATIONS IN CONTROLLED MARKOV CHAINS WITH RISK-SENSITIVE AVERAGE CRITERION

ROLANDO CAVAZOS-CADENA,* *Universidad Autónoma Agraria Antonio Narro*
DANIEL HERNÁNDEZ-HERNÁNDEZ,** *Centro de Investigación en Matemáticas*

Abstract

This work concerns Markov decision chains on a finite state space. The decision-maker has a constant and nonnull risk sensitivity coefficient, and the performance of a control policy is measured by two different indices, namely, the discounted and average criteria. Motivated by well-known results for the risk-neutral case, the problem of approximating the optimal risk-sensitive average cost in terms of the optimal risk-sensitive discounted value functions is addressed. Under suitable communication assumptions, it is shown that, as the discount factor increases to 1, appropriate normalizations of the optimal discounted value functions converge to the optimal average cost, and to the functional part of the solution of the risk-sensitive average cost optimality equation.

Keywords: Exponential utility; certainty equivalent; vanishing discount method; Hölder's inequality; convex function

2010 Mathematics Subject Classification: Primary 60J19
Secondary 90C39

1. Introduction

This work deals with discrete-time Markov decision chains on a finite state space. It is assumed that the controller has a (nonnull and) constant risk sensitivity coefficient, and two performance criteria for a control policy are considered, namely, the risk-sensitive average cost and the risk-sensitive discounted index. Besides standard continuity-compactness requirements, the basic framework of the paper is determined by the following property: the state process is communicating under each stationary policy, a condition that is necessary to ensure that the optimal risk-sensitive average cost is constant. In this context, the following problem is studied.

- To obtain convergent approximations for the optimal risk-sensitive average cost in terms of the family of optimal risk-sensitive discounted value functions.

This problem has a well-known solution in the risk-neutral case, which corresponds to a null risk sensitivity coefficient; see, for example, [1], [17], [22], and [27].

For the case of a nonnull risk sensitivity coefficient, the above problem was studied in [8], [9], and [13]. In [13] controlled Markov chains on a Borel space were analyzed under a geometric ergodicity assumption that is not generally satisfied in the present framework, and a solution

Received 13 June 2017; revision received 14 October 2017.
* Postal address: Departamento de Estadística y Cálculo, Universidad Autónoma Agraria Antonio Narro, Boulevard Antonio Narro 1923, Buenavista, Saltillo, Coah, 25315, México.
** Postal address: Centro de Investigación en Matemáticas, Apartado Postal 402, Guanajuato, Gto, 36000, México. Email address: dher@cimat.mx

was given under the condition that the magnitude of the risk sensitivity coefficient is small enough. In the other two aforementioned works such a condition was not imposed, but the results were concerned with uncontrolled Markov chains. The results of this paper extend the conclusions in those papers.

The risk-sensitive average criterion has been intensively studied, and the theory can be traced back, at least, to the seminal paper by Howard and Matheson [18], where the optimal average cost was characterized via an optimality equation rendering an optimal stationary policy. More recently, models with a finite or denumerable state space were considered in [6], [11], [15], [16], and [24], whereas Markov decision processes on a general state space were studied in [12]–[14] and [19]. The risk-sensitive average criterion is connected with the theory of large deviations [2], [20], and mathematical finance [3], [21], [25]. Markov decision models with a notion of risk more general than the one used in this paper were studied, for example, in [4] and [23]. In the context of dynamic games, a recent paper by Bäuerle and Rieder [5] deals with the characterization of risk-sensitive optimal strategies for a zero-sum game.

The main results of this work can be briefly described as follows. For a nonnull risk sensitivity coefficient, it is shown that, as the discount factor increases to 1, appropriate normalizations of the optimal discounted value functions converge to the optimal risk-sensitive average cost, and to the functional part of the solution of the average cost optimality equation.

If the classical risk-neutral normalization to approximate the average cost is applied in the present risk-sensitive context, then the normalized optimal discounted value function converges to the mean value of the optimal average cost over the interval joining zero and the risk sensitivity parameter.

The conditions under which these conclusions are obtained depend on the sign of the risk sensitivity coefficient. As will be explained in Section 3 below, this difference can be traced back to the fact that, in the risk-seeking case (corresponding to a negative risk sensitivity coefficient), a convexity property associated with the optimal discounted value functions holds, but such a feature cannot be generally ensured for positive risk sensitivity coefficients.

The organization of the paper is as follows. In Section 2 the decision model is formally introduced, the risk-sensitive average and discounted criteria are formulated, and the basic continuity-compactness and communication conditions are stated as Assumptions 2.1 and 2.2, respectively. In Section 3 the main conclusions of the paper are formulated in Theorems 3.1–3.3, the strategy that will be used to establish those results is outlined, and the technical reason for which an additional condition is required to analyze the case of a positive risk sensitivity coefficient is discussed. From this point onwards, the remainder of the paper is dedicated to proving those theorems. Thus, in Section 4 we present some general technical instruments that will be used to prove Theorems 3.1 and 3.2 in Sections 6 and 7, respectively, whereas Theorem 3.3 is established in Section 7. After some brief comments in Section 8, the exposition concludes with Appendix A in which we present the proof of some technical results which are used in the derivation of the main results.

Definition 1.1. (*Notation.*) Given a real-valued function f, the corresponding supremum norm is given by $\|f\| := \sup\{|f(y)|: y \text{ belongs to the domain of } f\}$, whereas $\mathcal{B}(S) := \{f: S \to \mathbb{R}: \|f\| < \infty\}$ denotes the space of all bounded functions defined on the nonempty set S, and $\mathcal{B}(S)$ is always endowed with the supremum norm. On the other hand, given an event A, the corresponding indicator function is denoted by $\mathbf{1}(A)$, and any relation between random variables holds almost surely with respect to the underlying probability measure. Finally, $\mathbb{N}$ denotes the set of all nonnegative integers.

2. The model

Throughout the remainder of the paper $\mathcal{M} = (S, A, \{A(x)\}, C, P)$ denotes a Markov decision chain, a model for a dynamical system evolving under the influence of a decision-maker (controller). The nonempty and finite set S is the state space and is endowed with the discrete topology, the metric space A is the action set, and, for each $x \in S$, $A(x) \subset A$ is the nonempty class of admissible actions at x. On the other hand, $C: \mathbb{K} \to \mathbb{R}$ is the cost function, where $\mathbb{K} := \{(x, a) \mid a \in A(x),\ x \in S\}$ is the set of admissible pairs, whereas $P = [p_{x,y}(\cdot)]$ is the controlled transition law. The interpretation of this model is as follows. At each time $t \in \mathbb{N}$, the state of the system is observed, say $X_t = x$, and the controller applies an action $A_t = a \in A(x)$. As a consequence of that intervention, a cost $C(x, a)$ is incurred and, regardless of the previous states and actions, at time $t + 1$ the state $X_{t+1} = y$ will be observed with probability $p_{x,y}(a)$, where $\sum_{y \in S} p_{x,y}(a) = 1$; this is the Markov property of the decision process.

Assumption 2.1. (i) *For each $x \in S$, $A(x)$ is a compact subset of A.*

(ii) *For every $x, y \in S$, the mappings $a \mapsto C(x, a)$ and $a \mapsto p_{x,y}(a)$ are continuous in $a \in A(x)$.*

Definition 2.1. (*Policies.*) The class $\mathcal{P}$ of decision policies consists of the (measurable) rules π for choosing actions which, at each time $t \in \mathbb{N}$, may depend on the states observed up to t and on the actions applied before t. Given $\pi \in \mathcal{P}$ and the initial state $X_0 = x$, the distribution $\mathbb{P}_x^\pi$ of the state-action process $\{(X_t, A_t)\}_{t \in \mathbb{N}}$ is uniquely determined, and $\mathbb{E}_x^\pi$ stands for the corresponding expectation operator; for details, see, for example, [22]. Define $\mathbb{F} := \prod_{x \in S} A(x)$, which is a compact metric space and consists of all functions $f: S \to A$ such that $f(x) \in A(x)$ for all $x \in S$. The class $\mathbb{M}$ of Markov policies consists of the sequences $\pi = (f_n)_{n \in \mathbb{N}}$, where $f_n \in \mathbb{F}$ for every $n \in \mathbb{N}$. Under $\pi = (f_n) \in \mathbb{M}$, the equality $A_n = f_n(X_n)$ is always valid, and if $f_n = f \in \mathbb{F}$ for every n, the policy is referred to as stationary and is naturally identified with f; with this convention, $\mathbb{F} \subset \mathbb{M} \subset \mathcal{P}$. Note that under a stationary policy $f \in \mathbb{F}$, the state process $\{X_t\}$ is a Markov chain with time-invariant transition matrix $[p_{x,y}(f(x))]_{x,y \in S}$.

Definition 2.2. (*Risk sensitivity and average criterion.*) Throughout the remainder of this paper, it is supposed that the controller has a constant risk sensitivity $\lambda \in \mathbb{R}$ so that a random cost W is assessed via the expectation of $U_\lambda(W)$, where the strictly increasing (dis-)utility function $U_\lambda(\cdot)$ is defined as follows. For each $w \in \mathbb{R}$,

$$U_\lambda(w) := \begin{cases} \operatorname{sign}(\lambda) \mathrm{e}^{\lambda w} & \text{if } \lambda \neq 0, \\ w & \text{if } \lambda = 0; \end{cases} \tag{2.1}$$

note that

$$U_\lambda(c + w) = \mathrm{e}^{\lambda c} U_\lambda(w), \qquad \lambda \neq 0, \qquad c, w \in \mathbb{R}. \tag{2.2}$$

If the decision-maker can choose between two random costs W_0 and W_1, she/he will prefer to pay W_0 when $\mathbb{E}[U_\lambda(W_1)] > \mathbb{E}[U_\lambda(W_0)]$, and will be indifferent between both costs when $\mathbb{E}[U_\lambda(W_1)] = \mathbb{E}[U_\lambda(W_0)]$. The certainty equivalent of a random cost W with respect to U_λ is the unique real number $\mathcal{E}[\lambda, W]$ satisfying $U_\lambda(\mathcal{E}[\lambda, W]) = \mathbb{E}[U_\lambda(W)]$, and then the controller will gladly pay the fixed amount $\mathcal{E}[\lambda, W]$ to avoid W. Note that, by (2.1),

$$\mathcal{E}[0, W] = \mathbb{E}[W], \qquad \mathcal{E}[\lambda, W] = \frac{1}{\lambda} \log(\mathbb{E}[\mathrm{e}^{\lambda W}]), \quad \lambda \neq 0, \tag{2.3}$$

and that

$$\mathbb{P}[|W| \le b] = 1 \quad \Longrightarrow \quad |\mathcal{E}[\lambda, W]| \le b, \qquad \lambda \in \mathbb{R}. \tag{2.4}$$

By Jensen's inequality, $\mathcal{E}[\lambda, W] \ge \mathbb{E}[W] = \mathcal{E}[0, W] \ge \mathcal{E}[-\lambda, W]$ for every $\lambda > 0$. The controller is referred to as risk averse (respectively, risk seeking) if $\lambda > 0$ (respectively, $\lambda < 0$). When $\lambda = 0$, the controller is risk neutral. Now suppose that the system is driven by a policy $\pi \in \mathcal{P}$ starting at $X_0 = x \in S$. Given $n \in \mathbb{N} \setminus \{0\}$, the total (random) cost incurred before time n is $\sum_{k=0}^{n-1} C(X_t, A_t)$, and the corresponding certainty equivalent is

$$J_n(\lambda, \pi, x) = \begin{cases} \lambda^{-1} \log\left(\mathbb{E}_x^\pi \left[\exp\left(\lambda \sum_{k=0}^{n-1} C(X_k, A_k) \right) \right] \right), & \lambda \ne 0, \\ \mathbb{E}_x^\pi \left[\sum_{k=0}^{n-1} C(X_k, A_k) \right], & \lambda = 0; \end{cases} \tag{2.5}$$

see (2.3). Thus, the controller is willing to pay $J_n(\lambda, \pi, x)$ to avoid the random cost associated with the first n state-action pairs (X_k, A_k), $0 \le k < n$, which represents an average of $J_n(\lambda, \pi, x)/n$ per action applied. The largest limit point of these averages is the (long-run) λ-sensitive average cost at state $x \in S$ under policy π:

$$J(\lambda, \pi, x) := \limsup_{n \to \infty} \frac{1}{n} J_n(\lambda, \pi, x). \tag{2.6}$$

The optimal λ-sensitive average cost function $J^*(\lambda, \cdot)$ is given by

$$J^*(\lambda, x) := \inf_{\pi \in \mathcal{P}} J(\lambda, \pi, x), \qquad x \in S, \tag{2.7}$$

and $\pi^* \in \mathcal{P}$ is λ-average optimal if $J(\lambda, \pi^*, x) = J^*(\lambda, x)$ for every $x \in S$. From (2.4)–(2.7), it is not difficult to see that

$$\|J^*(\lambda, \cdot)\| \le \|C\|, \qquad \lambda \in \mathbb{R}. \tag{2.8}$$

The above average criterion will be analyzed under the following condition on the transition law, ensuring that $J^*(\lambda, \cdot)$ is constant and is characterized in terms of a single optimality equation.

Assumption 2.2. *For each $f \in \mathbb{F}$, the corresponding Markov chain is communicating, that is, given $y, w \in S$, there exist a positive integer $N(f, y, w) \equiv N$ and states $y_k \in S$, $1 \le k \le N$, such that $y_0 = y$, $y_N = w$, and $p_{y_{i-1}, y_i}(f(y_{i-1})) > 0$ for every positive integer $i \le N$.*

For each $w \in S$, let T_w be the first return time to state w, which is given by

$$T_w := \min\{n \ge 1 \mid X_n = w\},$$

where, as usual, the minimum of the empty set is ∞. Note that

$$[T_w = n] \in \sigma(X_k, k \le n), \qquad w \in S, n \in \mathbb{N}. \tag{2.9}$$

Lemma 2.1. *Let $z \in S$ and $\lambda \in \mathbb{R}$ be arbitrary but fixed. Under Assumptions 2.1 and 2.2, the following assertions hold.*

(i) *The optimal λ-sensitive average cost function is constant, say $J^*(\lambda, \cdot) = g^*(\lambda)$.*

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

(ii) *Given the initial state $x \in S$, let $h^*(\lambda, x)$ be the minimum certainty equivalent of the total relative cost $\sum_{k=0}^{T_z-1}[C(X_k, A_k) - g^*(\lambda)]$ incurred before the first visit to state z in a positive time so that*

$$U_\lambda(h^*(\lambda, x)) = \inf_{\pi \in \mathcal{P}} \mathbb{E}_x^\pi\left[U_\lambda\left(\sum_{k=0}^{T_z-1}[C(X_k, A_k) - g^*(\lambda)]\right)\right], \qquad x \in S.$$

With this notation,

$$h^*(\lambda, z) = 0$$

and the following optimality equation holds:

$$U_\lambda(g^*(\lambda) + h^*(\lambda, x)) = \inf_{a \in A(x)}\left[\sum_{y \in S} p_{x,y}(a) U_\lambda(C(x, a) + h^*(\lambda, y))\right], \qquad x \in S. \tag{2.10}$$

(iii) *If $h\colon S \to \mathbb{R}$ and $g \in \mathbb{R}$ are such that (2.10) is satisfied when $h^*(\lambda, \cdot)$ and $g^*(\lambda)$ are replaced by $h(\cdot)$ and g, respectively, then $g = g^*(\lambda)$ and $h(\cdot) - h(z) = h^*(\lambda, \cdot)$.*

(iv) *There exists a policy $f^\lambda \in \mathbb{F}$ such that, for every $x \in S$, the action $f^\lambda(x)$ minimizes the term within brackets in (2.10), and such a stationary policy f^λ is λ-average optimal. Moreover,*

$$g^*(\lambda) = \lim_{n\to\infty} \frac{1}{n} J_n(\lambda, f^\lambda, x), \qquad x \in S.$$

For the risk-neutral case $\lambda = 0$, this is a classical result and its proof can be found in [17] or [22]. For the $\lambda \neq 0$ case, a verification of the above lemma can be found, for example, in [7]. Throughout the remainder of this paper, the state $z \in S$ is fixed, and the pair $(g^*(\lambda), h^*(\lambda, \cdot))$ is as described in Lemma 2.1.

Remark 2.1. Under Assumptions 2.1 and 2.2, the following simultaneous Doeblin condition holds:

$$\sup_{\pi \in \mathcal{P}} \mathbb{E}_x^\pi[T_y] =: B < \infty, \qquad x, y \in S; \tag{2.11}$$

see, for example, [26].

Definition 2.3. (*Discounted criteria.*) Let $\alpha \in (0, 1)$ be a discount factor so that the value at time 0 of the cost $C(X_t, A_t)$ to be paid at time t is $\alpha^t C(X_t, A_t)$. Then, all of the costs $\{C(X_t, A_t)\}_{t \in \mathbb{N}}$ incurred during the evolution of the system are worth $\sum_{t=0}^\infty \alpha^t C(X_t, A_t)$ at the beginning of the decision processes. Given the initial state $X_0 = x$ and the policy π used to drive the system, the corresponding λ-certainty equivalent is

$$V(\lambda, \alpha, \pi, x) := \begin{cases} \dfrac{1}{\lambda}\log\left(\mathbb{E}_x^\pi\left[\exp\left(\lambda\sum_{k=0}^\infty \alpha^k C(X_k, A_k)\right)\right]\right), & \lambda \neq 0, \\ \mathbb{E}_x^\pi\left[\sum_{k=0}^\infty \alpha^k C(X_k, A_K)\right], & \lambda = 0, \end{cases} \tag{2.12}$$

and the optimal (λ-sensitive) α-discounted value function is given by

$$V^*(\lambda, \alpha, x) := \inf_{\pi \in \mathcal{P}} V(\lambda, \alpha, \pi, x), \qquad x \in S; \tag{2.13}$$

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

if $\pi^* \in \mathcal{P}$ satisfies $V(\lambda, \alpha, \pi^*, x) = V^*(\lambda, \alpha, x)$ for all $x \in S$ then π^* is λ-optimal with respect to the α-discounted criterion. The optimality equations satisfied by the collection $\{V^*(\lambda, \alpha, \cdot)\}$ are given in the following lemma.

Lemma 2.2. *Under Assumptions 2.1 and 2.2, the following assertions hold.*

(i) *The family $\{V^*(\lambda, \alpha, \cdot)\}_{\lambda \in \mathbb{R},\, \alpha \in (0,1)}$ of optimal discounted value functions satisfies the following optimality equations. For each $x \in S$, $\lambda \in \mathbb{R}$, and $\alpha \in (0, 1)$,*

$$U_\lambda(V^*(\lambda, \alpha, x)) = \inf_{a \in A(x)} \Bigg[\sum_{y \in S} p_{x,y}(a) U_\lambda(C(x, a) + \alpha V^*(\lambda\alpha, \alpha, y)) \Bigg]. \tag{2.14}$$

(ii) *For each $\lambda \in \mathbb{R}$ and $\alpha \in (0, 1)$, there exists a policy $f^{\lambda,\alpha} \in \mathbb{F}$ such that, for every $x \in S$, the action $f^{\lambda,\alpha}(x)$ minimizes the term within the brackets on the right-hand side of (2.14). Moreover,*

$$V^*(\lambda, \alpha, x) = V(\lambda, \alpha, \pi^{\lambda,\alpha}, x), \qquad x \in S, \tag{2.15}$$

where the Markov policy $\pi^{\lambda,\alpha}$ is given by

$$\pi^{\lambda,\alpha} := (f^{\lambda,\alpha}, f^{\lambda\alpha,\alpha}, f^{\lambda\alpha^2,\alpha}, f^{\lambda\alpha^3,\alpha}, \dots). \tag{2.16}$$

A proof of this lemma can be found in [17] for the risk-neutral case, whereas the risk-sensitive case $\lambda \neq 0$ was verified in [4] and [13]. Note that the stationary policy $f^{\lambda,\alpha}$ in Lemma 2.2(ii) satisfies, for every $x \in S$ and $\lambda \in \mathbb{R}$,

$$U_\lambda(V^*(\lambda, \alpha, x)) = \sum_{y \in S} p_{x,y}(f^{\lambda,\alpha}(x)) U_\lambda(C(x, f^{\lambda,\alpha}(x)) + \alpha V^*(\lambda\alpha, \alpha, y)), \tag{2.17}$$

a relation that via (2.1) implies that

$$\mathrm{e}^{\lambda V^*(\lambda,\alpha,x)} = \mathrm{e}^{\lambda C(x, f^{\lambda,\alpha}(x))} \sum_{y \in S} p_{x,y}(f^{\lambda,\alpha}(x)) \mathrm{e}^{\lambda\alpha V^*(\lambda\alpha,\alpha,y)}, \qquad \lambda \neq 0. \tag{2.18}$$

2.1. The problem

In the risk-neutral case, under Assumptions 2.1 and 2.2, the following relation holds between the family of optimal discounted value functions $\{V^*(0, \alpha, \cdot)\}_{\alpha \in (0,1)}$ and the optimal average cost $g^*(0)$:

$$\lim_{\alpha \nearrow 1} (1 - \alpha) V^*(0, \alpha, x) = g^*(0), \qquad x \in S. \tag{2.19}$$

Moreover,

$$\lim_{\alpha \nearrow 1} [V^*(0, \alpha, x) - V^*(0, \alpha, z)] = h^*(0, x), \qquad x \in S, \tag{2.20}$$

where $h^*(0, \cdot)$ is the functional part of the solution to the optimality equation (2.10) specified in Lemma 2.1; see, for example, [1] and [22].

Given $\lambda \neq 0$, the main objective of the paper is to approximate the optimal λ-sensitive average cost $g^*(\lambda)$, as well as the function $h^*(\lambda, \cdot)$ in Lemma 2.1, via the family of optimal discounted value functions $\{V(\lambda, \alpha, \cdot)\}$.

The results on this problem, which are stated in the following section, extend the conclusions obtained in [13], where it was assumed that the absolute value of the risk sensitivity coefficient λ is small enough, and in [8], where the case of uncontrolled Markov chains was analyzed.

3. Main results

In this section the main approximation results of this paper will be stated. To begin with, it is convenient to introduce the following notation.

Definition 3.1. Let $z \in S$ be the fixed state in Lemma 2.1. For $\lambda \in \mathbb{R}$, $\alpha \in (0, 1)$, and $x \in S$, set

$$g_\alpha(\lambda, x) := V^*(\lambda, \alpha, x) - \alpha V^*(\lambda\alpha, \alpha, x) \quad \text{and} \quad h_\alpha(\lambda, x) := V^*(\lambda, \alpha, x) - V^*(\lambda, \alpha, z).$$

Note that $(1 - \alpha)V^*(0, \alpha, x) = g_\alpha(0, x)$ and $V^*(0, \alpha, x) - V^*(0, \alpha, z) = h_\alpha(0, x)$ so that (2.19) and (2.20) can be written as

$$\lim_{\alpha \nearrow 1} g_\alpha(0, x) = g^*(0) \quad \text{and} \quad \lim_{\alpha \nearrow 1} h_\alpha(0, x) = h^*(0, x), \qquad x \in S,$$

and the remainder of the paper is dedicated to extending these conclusions to the risk-sensitive context. The technical details are different for the risk-seeking and the risk-averse cases and require distinct conditions, as such the corresponding results are stated separately in the theorems stated below. To continue, let $\lambda \in \mathbb{R}$ and $\alpha \in (0, 1)$ be arbitrary, and note that the equalities

$$h_\alpha(\lambda, z) = 0 \quad \text{and} \quad h_\alpha(\lambda, \cdot) - \alpha h_\alpha(\lambda\alpha, \cdot) = g_\alpha(\lambda, \cdot) - g_\alpha(\lambda, z) \tag{3.1}$$

always hold. Next, observe that multiplying both sides of (2.14) by $\mathrm{e}^{-\lambda\alpha V^*(\lambda\alpha,\alpha,z)}$, direct calculations using the homogeneity property in (2.2) and the above definition yield that

$$U_\lambda(g_\alpha(\lambda, z)+h_\alpha(\lambda, x)) = \inf_{a\in A(x)}\left[\sum_{y\in S} p_{x,y}(a)U_\lambda(C(x, a)+\alpha h_\alpha(\lambda\alpha, y))\right], \qquad x \in S, \tag{3.2}$$

and it follows that, if $|h_\alpha(\lambda, \cdot) - \alpha h_\alpha(\lambda\alpha, \cdot)|$ is 'small' when α is 'close' to 1, then the pair $(g_\alpha(\lambda, z), h_\alpha(\lambda, (\cdot))$ is 'an approximate solution' of the optimality equation (2.10), and in such a case it might be expected that such a pair is 'approximately equal' to $(g^*(\lambda), h^*(\lambda, \cdot))$. In the following theorems we state conditions under which this intuitive argument can be formalized.

Theorem 3.1. *Under Assumptions 2.1 and 2.2, the following assertions are valid.*

(i) *For each $x \in S$ and $\lambda < 0$,*

$$\lim_{\alpha \nearrow 1} g_\alpha(\lambda, x) = g^*(\lambda) \tag{3.3}$$

and

$$\lim_{\alpha \nearrow 1} h_\alpha(\lambda, x) = h^*(\lambda, x), \tag{3.4}$$

where $(g^*(\lambda), h^*(\lambda, \cdot))$ is the solution of the optimality equation described in Lemma 2.1.

(ii) *The convergence (3.3) is uniform on compact intervals of $(-\infty, 0)$, that is, given $x \in S$ and a compact set $K \subset (-\infty, 0)$, $\sup_{\lambda\in K} |g_\alpha(\lambda, x) - g^*(\lambda)| \to 0$ as $\alpha \nearrow 1$.*

The key fact that will be used below to establish this theorem is that, for each $\alpha \in (0, 1)$ and $x \in S$, the mapping $\lambda \mapsto \lambda V^*(\lambda, \alpha, x)$ is convex in $\lambda \in (-\infty, 0)$. In general, this property cannot be ensured on the interval $(0, \infty)$ and, to extend the conclusions in Theorem 3.1 to the risk-averse case $\lambda > 0$, the following additional condition will be used.

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

Assumption 3.1. *For each* $(x, a) \in \mathbb{K}$, $p_{x,x}(a) > 0$.

Theorem 3.2. *Under Assumptions 2.1, 2.2, and 3.1, the following assertions hold.*

(i) *For each* $x \in S$ *and* $\lambda > 0$, *the convergences (3.3) and (3.4) are valid.*

(ii) *The convergence (3.3) is uniform on compact intervals of* $(0, \infty)$.

The proof of these two theorems will be presented in Sections 5 and 6, respectively, after establishing some basic tools in Section 4. On the other hand, due to (2.19), for $\lambda \neq 0$, it is interesting to investigate the behavior of $(1-\alpha)V^*(\lambda, \alpha, x)$ as $\alpha \nearrow 1$. The following result will be obtained from the uniform convergence conclusions in Theorems 3.1 and 3.2.

Theorem 3.3. *Suppose that Assumptions 2.1 and 2.2 hold. In this case, we have the following.*

(i) *For each* $\lambda < 0$,

$$\lim_{\alpha \nearrow 1} (1-\alpha)V^*(\lambda, \alpha, x) = \frac{1}{\lambda}\int_0^\lambda g^*(s)\,\mathrm{d}s. \tag{3.5}$$

(ii) *If, additionally, Assumption 3.1 is valid then the above convergence also holds for each* $\lambda > 0$.

Remark 3.1. (i) Theorems 3.1–3.3 extend the results of [8] and [9], where uncontrolled models were studied under the assumption that the underlying Markov chain is communicating. The proof of Theorem 3.1, which is based on a convexity property of the optimal discounted value functions, is motivated by the approach used in those papers. On the other hand, within the finite state context of this work, Theorems 3.1 and 3.2 extend a result of [13], where the convergence (3.3) was obtained assuming that $|\lambda|$ is small enough.

(ii) Let $\lambda \neq 0$ be fixed. Suppose that the convergences (3.3) and (3.4) hold and, for each $\alpha \in (0, 1)$, let $f^{\lambda,\alpha}$ be the stationary policy in Lemma 2.2(ii). In this case, any limit point $\tilde f$ of the family $\{f^{\lambda,\alpha}\}_{\alpha\in(0,1)}$ as $\alpha \nearrow 1$ is λ-average optimal. To verify this assertion, first note that direct calculations using Definition 3.1 and (2.2) yield that (2.17) is equivalent to

$$\begin{aligned} &U_\lambda(g_\alpha(\lambda, z) + h_\alpha(\lambda, x)) \\ &\quad = \sum_{y\in S} p_{x,y}(f^{\lambda,\alpha}(x))U_\lambda(C(x, f^{\lambda,\alpha}(x)) + \alpha h_\alpha(\lambda\alpha, y)), \qquad x \in S. \end{aligned} \tag{3.6}$$

Now, select a sequence $\{\alpha_n\} \subset (0, 1)$ increasing to 1 such that

$$\lim_{n\to\infty} f^{\lambda,\alpha_n}(\cdot) = \tilde f(\cdot)$$

and note that the second equality in (3.1), (3.3), and (3.4) together yield that

$$\lim_{n\to\infty} \alpha h_{\alpha_n}(\lambda\alpha, \cdot) = h^*(\lambda, \cdot).$$

Replacing α by α_n in (3.6), and taking the limit as n goes to ∞ on both sides of the resulting inequality, we see that (3.3) and (3.4) yield, via Assumption 2.1,

$$U_\lambda(g^*(\lambda) + h^*(\lambda, x)) = \sum_{y\in S} p_{x,y}(\tilde f(x))U_\lambda(C(x, \tilde f(x)) + h^*(\lambda, y)) \quad \text{for every state } x.$$

Thus, $\tilde f(x)$ minimizes the right-hand side of the optimality equation (2.10), and then $\tilde f$ is λ-optimal by Lemma 2.1(iii).

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

The remainder of this paper is devoted to the presentation of the proofs of the three theorems stated above. Since the arguments are rather technical, it is convenient to provide a brief description of the approach to be followed, as well as an outline of the organization of the subsequent material. Roughly, the essential objective of the analysis below is to establish that, for every $\lambda \neq 0$,

$$\lim_{\alpha \nearrow 1} g_\alpha(\lambda, x) = \lim_{\alpha \nearrow 1} g_\alpha(\lambda\alpha, x), \quad \text{where the limits do not depend on } x, \tag{3.7}$$

a fact that allows us to obtain the convergences (3.3) and (3.4) in a fairly direct way. To achieve this goal, the exposition has been organized as follows. Section 4 contains some general results that will be useful in the analysis of both the risk-seeking and the risk-averse cases, including bounds for the discounted approximations $g_\alpha(\lambda, \cdot)$ and $h_\alpha(\lambda, \cdot)$, relations between the functions $g_\alpha(\lambda, \cdot)$ and $g_\alpha(\lambda\alpha, \cdot)$ via stochastic matrices, and a continuity property of a dynamic programming operator. Next, Theorem 3.1 will be proved in Section 5. The argument relies on the fact that the mapping $\lambda \mapsto \lambda V^*(\lambda, \alpha, x)$ is always convex on $(-\infty, 0)$, a property that immediately yields that $g_\alpha(\cdot, x)$ is increasing on the negative axis and then, via the relations established in Section 4, (3.7) is directly derived. The proof of Theorem 3.2 is presented in Section 6 and, due to the fact that the convexity of the mapping $\lambda \mapsto \lambda V^*(\lambda, \alpha, x)$ cannot be generally ensured on the interval $(0, \infty)$, the argumentation is substantially more elaborated than the one used to prove Theorem 3.1, involving the mappings $\lambda \mapsto g_\alpha(\lambda\alpha^k, x)$ for every integer k, as well as a subtle application of Assumption 3.1. Indeed, the reason to introduce Assumption 3.1 in this paper is that, for the risk-averse case $\lambda > 0$, we have not been able to establish (3.7) based solely on Assumptions 2.1 and 2.2. Finally, Theorem 3.3 will be proved in Section 7 using the uniform convergence results in Theorems 3.1(ii) and 3.2(ii).

4. Auxiliary tools

This section contains basic technical instruments that will be used to prove the main conclusions of the paper. The first objective is to establish the following boundedness result.

Lemma 4.1. *Suppose that Assumptions 2.1 and 2.2 hold. In this context, the following assertions hold.*

(i) *For each $x \in S$,*

$$|g_\alpha(\lambda, x)| \le \|C\|, \qquad \alpha \in (0, 1), \qquad \lambda \in \mathbb{R} \setminus \{0\}. \tag{4.1}$$

(ii) *For every $x, y \in S$,*

$$|V^*(\lambda, \alpha, x) - V^*(\lambda, \alpha, y)| \le 2B\|C\|, \qquad \lambda \in \mathbb{R} \setminus \{0\}, \qquad \alpha \in (0, 1),$$

where B is the constant in (2.11).

Proof. (i) Let $\alpha \in (0, 1)$, $\lambda \in \mathbb{R} \setminus \{0\}$, and $x \in S$ be arbitrary but fixed. It will be shown that

$$\begin{aligned} \lambda V^*(\lambda, \alpha, x) &\le \lambda\alpha V^*(\lambda\alpha, \alpha, x) + |\lambda|\|C\|, \\ \lambda V^*(\lambda, \alpha, x) &\ge \lambda\alpha V^*(\lambda\alpha, \alpha, x) - |\lambda|\|C\|, \end{aligned} \tag{4.2}$$

inequalities that, via Definition 3.1, lead to (4.1). First note that (2.12) yields

$$\lambda V(\lambda, \alpha, \pi, x) = \log\biggl(\mathbb{E}_x^\pi\biggl[\exp\biggl(\lambda \sum_{k=0}^{\infty} \alpha^k C(X_k, A_k)\biggr)\biggr]\biggr), \qquad \pi \in \mathcal{P}.$$

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

Now, observe that

$$\lambda \sum_{k=0}^{\infty} \alpha^k C(X_k, A_k) = \lambda\alpha \sum_{k=0}^{\infty} \alpha^k C(X_k, A_k) + \lambda(1-\alpha) \sum_{k=0}^{\infty} \alpha^k C(X_k, A_k);$$

since $|\sum_{k=0}^{\infty} \alpha^k C(X_k, A_k)| \leq \|C\|/(1-\alpha)$, it follows that

$$\lambda \sum_{k=0}^{\infty} \alpha^k C(X_k, A_k) \leq \lambda\alpha \sum_{k=0}^{\infty} \alpha^k C(X_k, A_k) + |\lambda|\|C\|,$$
$$\lambda \sum_{k=0}^{\infty} \alpha^k C(X_k, A_k) \geq \lambda\alpha \sum_{k=0}^{\infty} \alpha^k C(X_k, A_k) - |\lambda|\|C\|.$$

These two last displays together imply that, for every $\pi \in \mathcal{P}$,

$$\begin{aligned} \lambda V(\lambda, \alpha, \pi, x) &\leq \lambda\alpha V(\lambda\alpha, \alpha, \pi, x) + |\lambda|\|C\|, \\ \lambda V(\lambda, \alpha, \pi, x) &\geq \lambda\alpha V(\lambda\alpha, \alpha, \pi, x) - |\lambda|\|C\|. \end{aligned} \tag{4.3}$$

Next, suppose that $\lambda < 0$. In this context, taking the supremum over $\pi \in \mathcal{P}$ on both sides of the above inequalities, via (2.13) it follows that (4.2) holds. On the other hand, if $\lambda > 0$, taking the infimum with respect to $\pi \in \mathcal{P}$ on both sides of the relations in (4.3), (2.13) yields that (4.2) remains valid in this case.

(ii) Let $\lambda \in \mathbb{R} \setminus \{0\}$ and $\alpha \in (0, 1)$ be arbitrary but fixed. To begin with, let the Markov policy $\pi^{\lambda,\alpha}$ be as in Lemma 2.2(ii), and observe that (2.12) and (2.15) together yield that

$$\exp(\lambda V^*(\lambda, \alpha, x)) = \mathbb{E}_x^{\pi^{\lambda,\alpha}}\left[\exp\left(\lambda \sum_{k=0}^{\infty} \alpha^k C(X_k, A_k)\right)\right], \qquad x \in S. \tag{4.4}$$

Next, let $n \in \mathbb{N} \setminus \{0\}$ and $x, y \in S$ be arbitrary but fixed, and note that

$$\begin{aligned} &\mathbb{E}_x^{\pi^{\lambda,\alpha}}\left[\mathbf{1}(T_y = n) \exp\left(\lambda \sum_{k=0}^{\infty} \alpha^k C(X_k, A_k)\right) \,\middle|\, (X_k, A_k),\ k < n, X_n\right] \\ &= \mathbb{E}_x^{\pi^{\lambda,\alpha}}\left[\mathbf{1}(T_y = n) \exp\left(\lambda \sum_{k=0}^{n-1} \alpha^k C(X_k, A_k)\right)\right. \\ &\qquad \left. \times \exp\left(\lambda\alpha^n \sum_{k=0}^{\infty} \alpha^k C(X_{k+n}, A_{k+n})\right) \,\middle|\, (X_k, A_k),\ k < n, X_n\right] \\ &= \mathbf{1}(T_y = n) \exp\left(\lambda \sum_{k=0}^{n-1} \alpha^k C(X_k, A_k)\right) \\ &\qquad \times \mathbb{E}_x^{\pi^{\lambda,\alpha}}\left[\exp\left(\lambda\alpha^n \sum_{k=0}^{\infty} \alpha^k C(X_{k+n}, A_{k+n})\right) \,\middle|\, (X_k, A_k),\ k < n, X_n\right] \end{aligned}$$

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

$$= \mathbf{1}(T_y = n)\exp\biggl(\lambda\sum_{k=0}^{n-1}\alpha^k C(X_k, A_k)\biggr)\mathbb{E}_{X_n}^{\pi^{\lambda\alpha^n,\alpha}}\biggl[\exp\biggl(\lambda\alpha^n\sum_{k=0}^{\infty}\alpha^k C(X_k, A_k)\biggr)\biggr]$$
$$= \mathbf{1}(T_y = n)\exp\biggl(\lambda\sum_{k=0}^{n-1}\alpha^k C(X_k, A_k)\biggr)\exp(\lambda\alpha^n V^*(\lambda\alpha^n, \alpha, X_n)),$$

where the second equality is due to (2.9), the third one follows from the Markov property and the definition of the policies $\pi^{\lambda,\alpha}$ in (2.16), and (4.4) was used in the last step. Since $X_n = y$ on the event $[T_y = n]$ by (2.9), and $\mathbb{P}_x^{\pi^{\lambda,\alpha}}[T_y < \infty] = 1$ by (2.11), it follows that

$$\mathbb{E}_x^{\pi^{\lambda,\alpha}}\biggl[\exp\biggl(\lambda\sum_{k=0}^{\infty}\alpha^k C(X_k, A_k)\biggr)\biggr]$$
$$= \mathbb{E}_x^{\pi^{\lambda,\alpha}}\biggl[\exp\biggl(\lambda\sum_{k=0}^{T_y-1}\alpha^k C(X_k, A_k) + \lambda\alpha^{T_y}V^*(\lambda\alpha^{T_y}, \alpha, y)\biggr)\biggr],$$

an equality that via (4.4) leads to

$$\exp(\lambda V^*(\lambda, \alpha, x)) = \mathbb{E}_x^{\pi^{\lambda,\alpha}}\biggl[\exp\biggl(\lambda\sum_{k=0}^{T_y-1}\alpha^k C(X_k, A_k) + \lambda\alpha^{T_y}V^*(\lambda\alpha^{T_y}, \alpha, y)\biggr)\biggr],$$

and then

$$\exp(\lambda[V^*(\lambda, \alpha, x) - V^*(\lambda, \alpha, y)])$$
$$= \mathbb{E}_x^{\pi^{\lambda,\alpha}}\biggl[\exp\biggl(\lambda\sum_{k=0}^{T_y-1}\alpha^k C(X_k, A_k) - \lambda[V^*(\lambda, \alpha, y) - \alpha^{T_y}V^*(\lambda\alpha^{T_y}, \alpha, y)]\biggr)\biggr].$$

Observe now that

$$V^*(\lambda, \alpha, y) - \alpha^{T_y}V^*(\lambda\alpha^{T_y}, \alpha, y) = \sum_{k=0}^{T_y-1}[\alpha^k V^*(\lambda\alpha^k, \alpha, y) - \alpha^{k+1}V^*(\lambda\alpha^k, \alpha, y)]$$
$$= \sum_{k=0}^{T_y-1}\alpha^k g_\alpha(\lambda\alpha^k, x),$$

by Definition 3.1. Combining the last two displays, it follows that

$$\exp(\lambda[V^*(\lambda, \alpha, x) - V^*(\lambda, \alpha, y)]) = \mathbb{E}_x^{\pi^{\lambda,\alpha}}\biggl[\exp\biggl(\lambda\sum_{k=0}^{T_y-1}\alpha^k[C(X_k, A_k) - g_\alpha(\lambda\alpha^k, y)]\biggr)\biggr].$$

Observing that $\lambda\sum_{k=0}^{T_y-1}\alpha^k[C(X_k, A_k) - g_\alpha(\lambda\alpha^k, y)] \ge -2|\lambda|\|C\|T_y$ by Lemma 4.1(i), the above equality implies that

$$e^{\lambda[V^*(\lambda,\alpha,x)-V^*(\lambda,\alpha,y)]} \ge \mathbb{E}_x^{\pi^{\lambda,\alpha}}[e^{-2|\lambda|\|C\|T_y}] \ge e^{-2|\lambda|\|C\|\mathbb{E}_x^{\pi^{\lambda,\alpha}}[T_y]} \ge e^{-2|\lambda|\|C\|B},$$

where Jensen's inequality was used in the second step, and the third inequality is due to (2.11).

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

Therefore,

$$\lambda[V^*(\lambda,\alpha,y) - V^*(\lambda,\alpha,x)] \le 2|\lambda|\|C\|B.$$

Interchanging the roles of x and y, it follows that $|\lambda|\,|V^*(\lambda,\alpha,y) - V^*(\lambda,\alpha,x)| \le 2|\lambda|\|C\|B$, and the conclusion follows, since $x, y \in S$ and $\lambda \neq 0$ are arbitrary. □

Given $\alpha \in (0,1)$, the next objective of this section is to relate the mappings $(\lambda, x) \mapsto g_\alpha(\lambda, x)$ and $(\lambda, x) \mapsto g_\alpha(\lambda\alpha, x)$ via a stochastic matrix. The result in this direction is stated as Lemma 4.2 below and involves the following notation. Recall that, for each $\lambda \neq 0$ and $\alpha \in (0,1)$, the stationary policy $f^{\lambda,\alpha}$ satisfies (2.17) and (2.18).

Definition 4.1. For each $\lambda \in \mathbb{R} \setminus \{0\}$ and $\alpha \in (0,1)$, the stochastic matrices $Q^{\lambda,\alpha}$ and $\tilde{Q}^{\lambda,\alpha}$ on S are defined as follows. For every $x, y \in S$,

$$Q^{\lambda,\alpha}_{x,y} := \frac{p_{x,y}(f^{\lambda,\alpha}(x))\mathrm{e}^{\lambda\alpha^2 V^*(\lambda\alpha^2,\alpha,y)}}{\sum_{w\in S} p_{x,w}(f^{\lambda,\alpha}(x))\mathrm{e}^{\lambda\alpha^2 V^*(\lambda\alpha^2,\alpha,w)}} \tag{4.5}$$

and

$$\tilde{Q}^{\lambda,\alpha}_{x,y} := \frac{p_{x,y}(f^{\lambda\alpha,\alpha}(x))\mathrm{e}^{\lambda\alpha^2 V^*(\lambda\alpha^2,\alpha,y)}}{\sum_{w\in S} p_{x,w}(f^{\lambda\alpha,\alpha}(x))\mathrm{e}^{\lambda\alpha^2 V^*(\lambda\alpha^2,\alpha,w)}}.$$

Lemma 4.2. *Under Assumptions 2.1 and 2.2, the following assertions hold for every* $\alpha \in (0,1)$.

(i) *If* $\lambda < 0$ *then*

$$\mathrm{e}^{\lambda g_\alpha(\lambda,x)} \le \mathrm{e}^{\lambda(1-\alpha)C(x,f^{\lambda,\alpha}(x))} \sum_{y\in S} Q^{\lambda,\alpha}_{x,y}\mathrm{e}^{\lambda\alpha g_\alpha(\lambda\alpha,y)}, \qquad x \in S.$$

(ii) *For each* $\lambda > 0$,

$$\mathrm{e}^{\lambda g_\alpha(\lambda,x)} \ge \mathrm{e}^{\lambda(1-\alpha)C(x,f^{\lambda,\alpha}(x))} \sum_{y\in S} Q^{\lambda,\alpha}_{x,y}\mathrm{e}^{\lambda\alpha g_\alpha(\lambda\alpha,y)}, \qquad x \in S, \tag{4.6}$$

and

$$\mathrm{e}^{\lambda g_\alpha(\lambda,x)} \le \mathrm{e}^{\lambda(1-\alpha)C(x,f^{\lambda\alpha,\alpha}(x))} \sum_{y\in S} \tilde{Q}^{\lambda,\alpha}_{x,y}\mathrm{e}^{\lambda\alpha g_\alpha(\lambda\alpha,y)}, \qquad x \in S. \tag{4.7}$$

(iii) *For each* $\lambda \in \mathbb{R}\setminus\{0\}$ *and* $x, y \in S$,

$$Q^{\lambda,\alpha}_{x,y} \ge p_{x,y}(f^{\lambda,\alpha}(x))\mathrm{e}^{-4|\lambda|B\|C\|} \quad \textit{and} \quad \tilde{Q}^{\lambda,\alpha}_{x,y} \ge p_{x,y}(f^{\lambda\alpha,\alpha}(x))\mathrm{e}^{-4|\lambda|B\|C\|}, \tag{4.8}$$

where the finite constant B *is as in (2.11).*

Proof. Let $\alpha \in (0,1)$ and $\lambda \in \mathbb{R}\setminus\{0\}$ be arbitrary but fixed. To begin with, note that (2.1) and the optimality equation (2.14) together yield that

$$\operatorname{sign}(\lambda)\mathrm{e}^{\lambda V^*(\lambda,\alpha,x)} \le \operatorname{sign}(\lambda)\mathrm{e}^{\lambda C(x,a)} \sum_{w\in S} p_{x,w}(a)\mathrm{e}^{\lambda\alpha V^*(\lambda\alpha,\alpha,w)}, \qquad x \in S,\ a \in A(x). \tag{4.9}$$

(i) Suppose that $\lambda < 0$. In this case, observing that $\operatorname{sign}(\lambda) = -1$, the above relation with $\lambda\alpha$ instead of λ leads to

$$\mathrm{e}^{\lambda\alpha V^*(\lambda\alpha,\alpha,x)} \ge \mathrm{e}^{\lambda\alpha C(x,f^{\lambda,\alpha}(x))} \sum_{w\in S} p_{x,w}(f^{\lambda,\alpha}(x))\mathrm{e}^{\lambda\alpha^2 V^*(\lambda\alpha^2,\alpha,w)}, \qquad x \in S.$$

Combining this equality with (2.18), it follows that, for every $x \in S$,

$$\begin{aligned} &\mathrm{e}^{\lambda[V^*(\lambda,\alpha,x)-\alpha V^*(\lambda\alpha,\alpha,x)]} \\ &\quad\le \frac{\mathrm{e}^{\lambda C(x,f^{\lambda,\alpha}(x))}\sum_{y\in S} p_{x,y}(f^{\lambda,\alpha}(x))\mathrm{e}^{\lambda\alpha V^*(\lambda\alpha,\alpha,y)}}{\mathrm{e}^{\lambda\alpha C(x,f^{\lambda,\alpha}(x))}\sum_{w\in S} p_{x,w}(f^{\lambda,\alpha}(x))\mathrm{e}^{\lambda\alpha^2 V^*(\lambda\alpha^2,\alpha,w)}} \\ &\quad= \mathrm{e}^{\lambda(1-\alpha)C(x,f^{\lambda\alpha,\alpha}(x))} \\ &\qquad\times \sum_{y\in S} \frac{p_{x,y}(f^{\lambda,\alpha}(x))\mathrm{e}^{\lambda\alpha^2 V^*(\lambda\alpha^2,\alpha,y)}}{\sum_{w\in S} p_{x,w}(f^{\lambda,\alpha}(x))\mathrm{e}^{\lambda\alpha^2 V^*(\lambda\alpha^2,\alpha,w)}}\mathrm{e}^{\lambda\alpha[V^*(\lambda\alpha,\alpha,y)-\alpha V^*(\lambda\alpha^2,\alpha,y)]} \end{aligned}$$

and the conclusion follows via Definitions 3.1 and 4.1.

(ii) Assume that $\lambda > 0$. In this context, $\operatorname{sign}(\lambda) = 1$, and (4.9) with $\lambda\alpha$ instead of λ now leads to

$$\mathrm{e}^{\lambda\alpha V^*(\lambda\alpha,\alpha,x)} \le \mathrm{e}^{\lambda\alpha C(x,f^{\lambda,\alpha}(x))}\sum_{w\in S} p_{x,w}(f^{\lambda,\alpha}(x))\mathrm{e}^{\lambda\alpha^2 V^*(\lambda\alpha^2,\alpha,w)}, \qquad x \in S.$$

Paralelling the argument used to establish part (i), this inequality and (2.18) together lead to (4.6). To establish (4.7), recall that λ is positive and note that (2.18) with $\lambda\alpha$ instead of λ yields

$$\mathrm{e}^{\lambda\alpha V^*(\lambda\alpha,\alpha,x)} = \mathrm{e}^{\lambda\alpha C(x,f^{\lambda\alpha,\alpha}(x))}\sum_{y\in S} p_{x,y}(f^{\lambda\alpha,\alpha}(x))\mathrm{e}^{\lambda\alpha^2 V^*(\lambda\alpha^2,\alpha,y)}, \qquad x \in S,$$

whereas (4.9) implies that

$$\mathrm{e}^{\lambda V^*(\lambda,\alpha,x)} \le \mathrm{e}^{\lambda C(x,f^{\lambda\alpha,\alpha}(x))}\sum_{y\in S} p_{x,y}(f^{\lambda\alpha,\alpha}(x))\mathrm{e}^{\lambda\alpha V^*(\lambda\alpha,\alpha,y)}, \qquad x \in S.$$

Using these last two displays, (4.7) follows along the same lines used in part (i).

(iii) Let $w^* \in S$ be fixed, and note that (4.5) yields

$$Q^{\lambda,\alpha}_{x,y} = \frac{p_{x,y}(f^{\lambda,\alpha}(x))\mathrm{e}^{\lambda\alpha^2[V^*(\lambda\alpha^2,\alpha,y)-V^*(\lambda\alpha^2,\alpha,w^*)]}}{\sum_{w\in S} p_{x,w}(f^{\lambda,\alpha}(x))\mathrm{e}^{\lambda\alpha^2[V^*(\lambda\alpha^2,\alpha,w)-V^*(\lambda\alpha^2,\alpha,w^*)]}}, \qquad x, y \in S.$$

Using the fact that $|V^*(\lambda\alpha^2,\alpha,\cdot) - V^*(\lambda\alpha^2,\alpha,w^*)| \le 2|B\|C\|$, by Lemma 4.1, it follows that, for every $x, y \in S$,

$$\mathrm{e}^{\lambda\alpha^2[V^*(\lambda\alpha^2,\alpha,y)-V^*(\lambda\alpha^2,\alpha,w^*)]} \ge \mathrm{e}^{-2|\lambda|B\|C\|}$$

and

$$\sum_{w\in S} p_{x,w}(f^{\lambda,\alpha}(x))\mathrm{e}^{\lambda\alpha^2[V^*(\lambda\alpha^2,\alpha,w)-V^*(\lambda\alpha^2,\alpha,w^*)]} \le \mathrm{e}^{2|\lambda|B\|C\|}.$$

These last three displays yield that $Q^{\lambda,\alpha}_{x,y} \ge p_{x,y}(f^{\lambda,\alpha}(x))\mathrm{e}^{-4|\lambda|B\|C|}$ for all $x, y \in S$, whereas the second inequality in (4.8) can be established along similar lines. □

The following continuity property will be useful.

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

Lemma 4.3. *For each $x \in S$, the mapping*

$$(\lambda, h(\cdot)) \mapsto \inf_{a \in A(x)} \left[\sum_{y \in S} p_{x,y}(a) U_\lambda(C(x,a) + h(y)) \right] =: T[\lambda, h](x)$$

is continuous in $(\lambda, h(\cdot)) \in (\mathbb{R} \setminus \{0\}) \times \mathcal{B}(S)$.

Proof. Let $x, y \in S$, $h, \tilde{h} \in \mathcal{B}(S)$, and $\lambda, \tilde{\lambda} \in \mathbb{R} \setminus \{0\}$ be arbitrary, where λ and $\tilde{\lambda}$ have the same sign. Set

$$\varepsilon(\lambda, \tilde{\lambda}, h, \tilde{h}) := |\lambda - \tilde{\lambda}|(\|C\| + \|h\|) + |\tilde{\lambda}| \, \|h - \tilde{h}\|$$

and observe that $\mathrm{e}^{\lambda(C(x,a)+h(y))} \le \mathrm{e}^{\tilde{\lambda}(C(x,a)+\tilde{h}(y))} \mathrm{e}^{\varepsilon(\lambda, \tilde{\lambda}, h, \tilde{h})}$ for every $a \in A(x)$. This inequality and (2.1) immediately yield that

$$T[\lambda, h](x) \begin{cases} \le T[\tilde{\lambda}, \tilde{h}](x) \mathrm{e}^{\varepsilon(\lambda, \tilde{\lambda}, h, \tilde{h})}, & \lambda, \tilde{\lambda} > 0, \\ \ge T[\tilde{\lambda}, \tilde{h}](x) \mathrm{e}^{\varepsilon(\lambda, \tilde{\lambda}, h, \tilde{h})}, & \lambda, \tilde{\lambda} < 0. \end{cases}$$

Interchanging the roles of (λ, h) and $(\tilde{\lambda}, \tilde{h})$, it follows that

$$T[\tilde{\lambda}, \tilde{h}](x) \begin{cases} \le T[\lambda, h](x) \mathrm{e}^{\varepsilon(\tilde{\lambda}, \lambda, \tilde{h}, h)}, & \lambda, \tilde{\lambda} > 0, \\ \ge T[\lambda, h](x) \mathrm{e}^{\varepsilon(\tilde{\lambda}, \lambda, \tilde{h}, h)}, & \lambda, \tilde{\lambda} < 0. \end{cases}$$

Since $\varepsilon(\lambda, \tilde{\lambda}, h, \tilde{h}) + \varepsilon(\tilde{\lambda}, \lambda, \tilde{h}, h) \to 0$ as $\tilde{\lambda} \to \lambda$ in $\mathbb{R} \setminus \{0\}$ and $\tilde{h} \to h$ in $\mathcal{B}(S)$, the desired conclusion follows from the last two displays. $\square$

The following continuity result for the optimal risk-sensitive average cost $g^*(\cdot)$ was established in Proposition 2.1 and Theorem 3.1 of [10].

Lemma 4.4. *Under Assumptions 2.1 and 2.2, the mapping $\lambda \mapsto g^*(\lambda)$ is continuous on $\mathbb{R}$.*

5. The risk-seeking case

In this section Theorem 3.1 will be proved. The argument has been divided into three parts, stated as Lemmas 5.1–5.3 below. The starting point is the following result, where a fundamental convexity property is established.

Lemma 5.1. *Let $\alpha \in (0, 1)$ and $x \in S$ be arbitrary.*

(i) *For each $\pi \in \mathcal{P}$, the mapping $\lambda \mapsto \lambda V(\lambda, \alpha, \pi, x)$ is convex on $\mathbb{R}$.*

(ii) *The function $\lambda \mapsto \lambda V^*(\lambda, \alpha, x)$ is convex on $(-\infty, 0)$.*

(iii) *The function $g_\alpha(\cdot, x)$ is increasing and continuous on $(-\infty, 0)$.*

Proof. (i) Given $\pi \in \mathcal{P}$, note that (2.12) yields

$$\mathrm{e}^{\lambda V(\lambda, \alpha, \pi, x)} = \mathbb{E}_x^\pi[\mathrm{e}^{\lambda W_\alpha}], \qquad \lambda \in \mathbb{R}, \tag{5.1}$$

where $W_\alpha := \sum_{k=0}^{\infty} \alpha^k C(X_k, A_k)$. Now, let $\lambda, \lambda_0, \lambda_1 \in \mathbb{R}$ and $\rho \in (0, 1)$ be such that

$$\lambda = \rho \lambda_0 + (1 - \rho) \lambda_1,$$

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

and observe that Hölder's inequality implies that

$$\mathbb{E}_x^\pi[\mathrm{e}^{\lambda W}] = \mathbb{E}_x^\pi[\mathrm{e}^{\rho\lambda_0 W}\mathrm{e}^{(1-\rho)\lambda_1 W}] \le \mathbb{E}_x^\pi[\mathrm{e}^{\lambda_0 W}]^\rho \mathbb{E}[\mathrm{e}^{\lambda_1 W}]^{(1-\rho)}.$$

Using (5.1), this relation leads to $\mathrm{e}^{\lambda V(\lambda,\alpha,\pi,x)} \le \mathrm{e}^{\rho\lambda_0 V(\lambda_0,\alpha,\pi,x)}\mathrm{e}^{(1-\rho)\lambda_1 V(\lambda_1,\alpha,\pi,x)}$, that is,

$$\lambda V(\lambda,\alpha,\pi,x) \le \rho\lambda_0 V(\lambda_0,\alpha,\pi,x) + (1-\rho)\lambda_1 V(\lambda_1,\alpha,\pi,x),$$

establishing the convexity of the mapping $\lambda \mapsto \lambda V[\lambda,\alpha,\pi,x)$.

(ii) Since the supremum of a family of convex functions is also convex, part (i) implies that the mapping $\lambda \mapsto \sup_{\pi\in\mathcal{P}}[\lambda V(\lambda,\alpha,\pi,x)]$ is convex on $\mathbb{R}$. The conclusion follows by combining this fact with the relation

$$\sup_{\pi\in\mathcal{P}}[\lambda V(\lambda,\alpha,\pi,x)] = \lambda \inf_{\pi\in\mathcal{P}} V(\lambda,\alpha,\pi,x) = \lambda V^*(\lambda,\alpha,x), \qquad \lambda < 0;$$

see (2.13) for the second equality.

(iii) Using the fact that a convex function on an open interval is continuous, part (ii) implies that the function $\lambda \mapsto V^*(\lambda,\alpha,x) = \lambda^{-1}[\lambda V^*(\lambda,\alpha,x)]$ is continuous on $(-\infty,0)$, and then so is $g_\alpha(\cdot,x)$ by Definition 3.1. Now, let $\lambda,\lambda_1 \in (-\infty,0)$ be such that

$$\lambda < \lambda_1.$$

In this case

$$\lambda\alpha > \lambda, \qquad \lambda_1\alpha > \lambda_1, \qquad \lambda\alpha < \lambda_1\alpha$$

so that the extreme points of the interval $[\lambda,\lambda\alpha]$ are less than the corresponding end points of the interval $[\lambda_1,\lambda_1\alpha]$, and then the convexity of the mapping $\mu \mapsto \mu V^*(\mu,\alpha,x)$ on $(-\infty,0)$ established in part (ii) implies that

$$\frac{\lambda V(\lambda,\alpha,x) - \lambda\alpha V(\lambda\alpha,\alpha,x)}{\lambda - \lambda\alpha} \le \frac{\lambda_1 V(\lambda_1,\alpha,x) - \lambda_1\alpha V(\lambda_1\alpha,\alpha,x)}{\lambda_1 - \lambda_1\alpha}.$$

Using the inclusion $\alpha \in (0,1)$, this inequality is equivalent to $g_\alpha(\lambda,x) \le g_\alpha(\lambda_1,x)$, by Definition 3.1, so that $g_\alpha(\cdot,x)$ is increasing on $(-\infty,0)$. □

Next, the limit points of the family $\{g_\alpha(\lambda,\cdot)\}$ as α goes to 1 will be studied using Lemmas 4.2 and 5.1(iii).

Lemma 5.2. *Let $\lambda < 0$ be arbitrary but fixed. Suppose that the sequence $\{\alpha_n\} \subset (0,1)$ satisfies*

$$\lim_{n\to\infty}\alpha_n = 1 \quad \textit{and} \quad \lim_{n\to\infty} g_{\alpha_n}(\lambda,x) =: \tilde{g}(\lambda,x) \quad \textit{exist for each } x \in S. \tag{5.2}$$

In this case,

(i) $\lim_{n\to\infty} g_{\alpha_n}(\lambda\alpha_n,x) = \tilde{g}(\lambda,x)$ *for all $x \in S$, and*

(ii) $\tilde{g}(\lambda,\cdot)$ *is constant, say $\tilde{g}(\lambda,\cdot) = \tilde{g}^*(\lambda)$.*

Proof. Recall that that the inequality $|g_\alpha(\lambda,x)| \le \|C\|$ is always valid, by Lemma 4.1, and let $\{\alpha_n\} \subset (0,1)$ be such that (5.2) holds.

(i) It is sufficient to show that any limit point of $\{g_{\alpha_n}(\lambda\alpha_n, \cdot)\}$ coincides with $\tilde{g}(\lambda, \cdot)$. To achieve this goal, let $\tilde{g}^{(1)}(\lambda, \cdot): S \to \mathbb{R}$ be an arbitrary limit point of $\{g_{\alpha_n}(\lambda\alpha_n, \cdot)\}$ and select a subsequence $\{\beta_k\}$ of $\{\alpha_n\}$ such that

$$\lim_{k\to\infty} g_{\beta_k}(\lambda\beta_k, x) = \tilde{g}^{(1)}(\lambda, x), \qquad x \in S. \tag{5.3}$$

Since the space of stationary policies is a compact metric and the matrices $Q^{\lambda,\alpha}$ in Definition 4.1 are stochastic, taking an additional subsequence (if necessary), without loss of generality it can be supposed that the following limits also exist for every $x, y \in S$:

$$\lim_{k\to\infty} f^{\lambda,\beta_k}(x) =: f^{\lambda}(x) \quad \text{and} \quad \lim_{k\to\infty} Q^{\lambda,\beta_k}_{x,y} =: Q^{\lambda}_{x,y}, \tag{5.4}$$

where, since S is finite, the limit matrix Q^{λ} is stochastic. Moreover, via Lemma 4.2(iii) and Assumption 2.1, the above display yields that $Q^{\lambda}_{x,y} \ge \mathrm{e}^{-4|\lambda|B\|C\|} p_{x,y}(f^{\lambda}(x))$ for every $x, y \in S$, and then Q^{λ} is a communicating matrix, by Assumption 2.2. Consequently, Q^{λ} has an invariant distribution $\rho(\cdot)$ which is positive at each state, that is,

$$\rho(y) > 0 \quad \text{and} \quad \rho(y) = \sum_{x\in S} \rho(x) Q^{\lambda}_{x,y}, \qquad y \in S. \tag{5.5}$$

Observe now that, for every $k \in \mathbb{N}$ and $x \in S$,

$$\mathrm{e}^{\lambda g_{\beta_k}(\lambda,x)} \le \mathrm{e}^{\lambda(1-\beta_k)C(x,f^{\lambda,\beta_k}(x))} \sum_{y\in S} Q^{\lambda,\beta_k}_{x,y} \mathrm{e}^{\lambda\beta_k g_{\beta_k}(\lambda\beta_k,y)},$$

by Lemma 4.2(i), whereas, using the fact that $\lambda < 0$, Lemma 5.1(iii) yields

$$\lambda g_{\beta_k}(\lambda, \cdot) \ge \lambda g_{\beta_k}(\lambda\beta_k, \cdot).$$

Since $\{\beta_k\}$ is a subsequence of $\{\alpha_n\}$, taking the limit on both sides of the inequalities in the last two displays, via (5.2)–(5.4), it follows that

$$\mathrm{e}^{\lambda\tilde{g}(\lambda,x)} \le \sum_{y\in S} Q^{\lambda}_{x,y} \mathrm{e}^{\lambda\tilde{g}^{(1)}(\lambda,y)}, \qquad x \in S,$$

and

$$\lambda\tilde{g}(\lambda, \cdot) \ge \lambda\tilde{g}^{(1)}(\lambda, \cdot). \tag{5.6}$$

Combining these two inequalities with (5.5), it follows that

$$\begin{aligned}
0 &\le \sum_{x\in S} \rho(x) \bigg[\sum_{y\in S} Q^{\lambda}_{x,y} \mathrm{e}^{\lambda\tilde{g}^{(1)}(\lambda,y)} - \mathrm{e}^{\lambda\tilde{g}(\lambda,x)}\bigg] \\
&= \sum_{y\in S}\sum_{x\in S} \rho(x) Q^{\lambda}_{x,y} \mathrm{e}^{\lambda\tilde{g}^{(1)}(\lambda,y)} - \sum_{x\in S} \rho(x) \mathrm{e}^{\lambda\tilde{g}(\lambda,x)} \\
&= \sum_{y\in S} \rho(y) \mathrm{e}^{\lambda\tilde{g}^{(1)}(\lambda,y)} - \sum_{x\in S} \rho(y) \mathrm{e}^{\lambda\tilde{g}(\lambda,x)} \\
&= \sum_{x\in S} \rho(x) [\mathrm{e}^{\lambda\tilde{g}^{(1)}(\lambda,x)} - \mathrm{e}^{\lambda\tilde{g}(\lambda,x)}] \\
&\le 0,
\end{aligned}$$

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

where the last inequality is due to (5.6). Recalling that $\rho(\cdot) > 0$ and λ is nonnull, these last three displays together lead to

$$\tilde{g}^{(1)}(\lambda, x) = \tilde{g}(\lambda, x), \qquad e^{\lambda \tilde{g}(x)} = \sum_{y \in S} Q^{\lambda}_{x,y} e^{\lambda \tilde{g}(y)}, \qquad x \in S. \tag{5.7}$$

As already mentioned, from the first equality we see that

$$\lim_{n\to\infty} g_{\alpha_n}(\lambda\alpha_n, x) = \tilde{g}(\lambda, x) = \lim_{n\to\infty} g_{\alpha_n}(\lambda, x), \qquad x \in S.$$

(ii) Since Q^{λ} has a unique invariant distribution $\rho(\cdot)$, via the ergodic theorem the second equality in (5.7) implies that $e^{\lambda \tilde{g}(\lambda,x)} = \sum_{y\in S} \rho(y) e^{\lambda \tilde{g}(\lambda,y)}$ for every $x \in S$, and then $\tilde{g}(\lambda, \cdot) = \tilde{g}^*(\lambda)$, where $\tilde{g}^*(\lambda) := \lambda^{-1} \log(\sum_{y\in S} \rho(y) e^{\lambda \tilde{g}(\lambda,y)})$. □

The following result is the final step before the proof of Theorem 3.1.

Lemma 5.3. *Given $\lambda \neq 0$, suppose that $\tilde{g}(\lambda) \in \mathbb{R}$ is such that*

$$\lim_{n\to\infty} g_{\alpha_n}(\lambda, x) = \tilde{g}(\lambda), \qquad x \in S, \tag{5.8}$$

where $\{\alpha_n\} \subset (0, 1)$ and $\alpha_n \nearrow 1$. In this case,

$$\tilde{g}(\lambda) = g^*(\lambda) \quad \textit{and} \quad \lim_{n\to\infty} h_{\alpha_n}(\lambda, x) = h^*(\lambda, x), \qquad x \in S,$$

where $(g^*(\lambda), h^*(\lambda, \cdot))$ is the solution of the optimality equation (2.10).

Proof. Let $\{\alpha_n\} \subset (0, 1)$ be a sequence such that $\alpha_n \nearrow 1$ and (5.8) holds. Next, let $\tilde{h}(\lambda, \cdot)$ be any limit point of $\{h_{\alpha_n}(\lambda, \cdot)\}$, and select a subsequence $\{\beta_k\}_{k\in\mathbb{N}}$ of $\{\alpha_n\}$ such that

$$\lim_{k\to\infty} h_{\beta_k}(\lambda, x) =: \tilde{h}(\lambda, x), \qquad x \in S. \tag{5.9}$$

Now, observe that (3.1) implies that $\tilde{h}(\lambda, z) = 0$, whereas, for every $x \in S$,

$$\lim_{k\to\infty} [h_{\beta_k}(\lambda, x) - \beta_k h_{\beta_k}(\lambda\beta_k, x)] = \lim_{k\to\infty} [g_{\beta_k}(\lambda, x) - g_{\beta_k}(\lambda, z)] = \tilde{g}(\lambda) - \tilde{g}(\lambda) = 0,$$

where the second equality is due to (5.8). Since S is finite, the two previous displays together yield that

$$\|\beta_k h_{\beta_k}(\lambda\beta_k, x) - \tilde{h}(\lambda, x)\| \to 0 \quad \text{as } k \to \infty. \tag{5.10}$$

Next, replace α by β_k on both sides of (3.2) and take the limit as k goes to ∞ on both sides of the resulting equality to obtain, via (5.8)–(5.10) and Lemma 4.3, that the pair $(\tilde{g}(\lambda), \tilde{h}(\lambda, \cdot))$ satisfies the optimality equation (2.10). Recalling that $\tilde{h}(\lambda, z) = 0$, it follows that $\tilde{g}(\lambda) = g^*(\lambda)$ and $\tilde{h}(\lambda, \cdot)) = h^*(\lambda, \cdot)$, by Lemma 2.1(iii), and the conclusion follows, since $\tilde{h}(\lambda, \cdot)$ is an arbitrary limit point of $\{h_{\alpha_n}(\lambda, \cdot)\}$. □

Proof of Theorem 3.1. Suppose that Assumptions 2.1 and 2.2 hold, and recall that, by Lemma 4.1, $\|g_\alpha(\lambda, \cdot)\| \leq \|C\|$ and $\|h_\alpha(\lambda, \cdot)\| \leq 2B\|C\|$.

(i) Let $(\tilde{g}(\lambda, \cdot), \tilde{h}(\lambda, \cdot))$ be an arbitrary limit point of the family $\{(g_\alpha(\lambda, \cdot), h_\alpha(\lambda, \cdot))\}_{\alpha\in(0,1)}$ as α increases to 1. To establish (3.3) and (3.4), it is sufficient to show that

$$\tilde{g}(\lambda, \cdot) = g^*(\lambda) \quad \text{and} \quad \tilde{h}(\lambda, \cdot) = h^*(\lambda, \cdot). \tag{5.11}$$

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

To achieve this goal, let $\{\alpha_n\} \subset (0, 1)$ be a sequence such that $\alpha_n \nearrow 1$ and $\lim_{n\to\infty} g_{\alpha_n}(\lambda, x) = \tilde{g}(\lambda, x)$ for every $x \in S$. In this context, Lemma 5.2 yields that $\tilde{g}(\lambda, \cdot)$ is constant, say $\tilde{g}(\lambda)$, and (5.11) follows via Lemma 5.3.

(ii) Let $I \subset (-\infty, 0)$ be a compact interval. Given an arbitrary but fixed state $x \in S$, set

$$\Delta(\alpha) := \sup_{\lambda\in I} \|g_\alpha(\lambda, x) - g^*(\lambda)\|, \qquad \alpha \in (0, 1),$$

and note that, by Lemmas 4.4 and 5.1(iii), for each $\alpha \in (0, 1)$, there exits λ_α such that

$$\lambda_\alpha \in I \quad \text{and} \quad \Delta(\alpha) = |g_\alpha(\lambda_\alpha, x) - g^*(\lambda_\alpha)|, \qquad \alpha \in (0, 1).$$

It must be shown that

$$\Delta(\alpha) \to 0 \quad \text{as } \alpha \nearrow 1,$$

a convergence that is equivalent to the following statement:

$$\text{if } \{\alpha_n\} \subset (0, 1) \text{ and } \alpha_n \nearrow 1 \text{ then } |g_{\alpha_n}(\lambda_{\alpha_n}, x) - g^*(\lambda_{\alpha_n})| \to 0. \tag{5.12}$$

To establish this claim, let $\{\alpha_n\} \subset (0, 1)$ be an arbitrary sequence increasing to 1, and note that, since the sequence $\{\lambda_{\alpha_n}\}$ is contained in the compact interval I, taking a subsequence (if necessary), without loss of generality it can be assumed that

$$\lim_{n\to\infty} \lambda_{\alpha_n} =: \lambda^* \in I \subset (-\infty, 0). \tag{5.13}$$

Next, let $\varepsilon > 0$ be arbitrary. Using the fact that $g_\alpha(\cdot, x)$ converges pointwise to $g^*(\cdot)$ on $(-\infty, 0)$, by part (i), Lemma 5.1(iii) implies that $g^*(\cdot)$ is an increasing function on $(-\infty, 0)$. Recalling that $g^*(\cdot)$ is continuous, by Lemma 4.4, select numbers a, b such that

$$a < \lambda^* < b < 0 \quad \text{and} \quad [g^*(a), g^*(b)] \subset (g^*(\lambda^*) - \varepsilon, g^*(\lambda^*) + \varepsilon). \tag{5.14}$$

To continue, using (5.13), select a positive integer $N_1(\varepsilon)$ such that

$$n > N_1(\varepsilon) \quad \Longrightarrow \quad \lambda_{\alpha_n} \in (a, b)$$

and, recalling that both $g_{\alpha_n}(\cdot, x)$ and $g^*(\cdot)$ are increasing on $(-\infty, 0)$, observe that

$$n > N_1(\varepsilon) \quad \Longrightarrow \quad g_{\alpha_n}(\lambda_{\alpha_n}, x) \in [g_{\alpha_n}(a, x), g_{\alpha_n}(b, x)] \quad \text{and} \quad g^*(\lambda_{\alpha_n}) \in [g^*(a), g^*(b)].$$

On the other hand, part (i) implies that there exists a positive integer $N_2(\varepsilon)$ such that

$$n > N_2(\varepsilon) \quad \Longrightarrow \quad |g_{\alpha_n}(a, x) - g^*(a)| < \varepsilon \quad \text{and} \quad |g_{\alpha_n}(b, x) - g^*(b)| < \varepsilon.$$

The last two displays together lead to

$$n > \max\{N_1(\varepsilon), N_2(\varepsilon)\} \quad \Longrightarrow \quad g^*(\lambda_{\alpha_n}),$$
$$g_{\alpha_n}(\lambda_{\alpha_n}, x) \in (g^*(a) - \varepsilon, g^*(b) + \varepsilon) \subset (g^*(\lambda^*) - 2\varepsilon, g^*(\lambda^*) + 2\varepsilon),$$

where (5.14) was used to set the inclusion. Therefore,

$$n > \max\{N_1(\varepsilon), N_2(\varepsilon)\} \quad \Longrightarrow \quad |g^*(\lambda_{\alpha_n}) - g_{\alpha_n}(\lambda_{\alpha_n}, x)| \le 4\varepsilon,$$

and then $|g^*(\lambda_{\alpha_n}) - g_{\alpha_n}(\lambda_{\alpha_n}, x)| \to 0$, since $\varepsilon > 0$ is arbitrary. This establishes (5.12) and, as already noted, completes the proof of the theorem. □

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

6. The approximation results under risk aversion

In this section Theorem 3.2 will be proved. The argument relies heavily on Theorem 6.1 below, whose statement involves two sequences $\{\lambda_n\}$ and $\{\alpha_n\}$ satisfying the following conditions:

$$\lambda_n \in (0, \infty), \qquad \alpha_n \in (0, 1), \qquad n \in \mathbb{N}, \tag{6.1}$$

and

$$\lim_{n\to\infty} \lambda_n =: \lambda^* \in (0, \infty), \quad \text{whereas} \quad \lim_{n\to\infty} \alpha_n = 1. \tag{6.2}$$

Theorem 6.1. *Suppose that Assumptions 2.1, 2.2, and 3.1 hold, and let the sequences* $\{\lambda_n\}$ *and* $\{\alpha_n\}$ *be as in (6.1) and (6.2). In this case,*

$$\lim_{n\to\infty} g_{\alpha_n}(\lambda_n, x) = g^*(\lambda^*)$$

and

$$\lim_{n\to\infty} h_{\alpha_n}(\lambda_n, x) = h^*(\lambda^*, x) = \lim_{n\to\infty} \alpha_n h_{\alpha_n}(\lambda_n \alpha_n, x), \qquad x \in S.$$

Before presenting the rather technical proof of this theorem, it will be useful to derive Theorem 3.2.

Proof of Theorem 3.2. Suppose that Assumptions 2.1, 2.2, and 3.1 hold, and let $\lambda > 0$ be arbitrary.

(i) To establish (3.3) and (3.4), it is sufficient to show that if $\{\alpha_n\}$ is a sequence of positive numbers increasing to 1, then

$$\lim_{n\to\infty} g_{\alpha_n}(\lambda, x) = g^*(\lambda) \quad \text{and} \quad \lim_{n\to\infty} h_{\alpha_n}(\lambda, x) = h^*(\lambda, x), \qquad x \in S,$$

convergence follows from Theorem 6.1 applied to the $\lambda_n = \lambda$ case for every n.

(ii) Given $x \in S$ and a compact interval I contained in $(0, \infty)$, define

$$\tilde{\Delta}_n(x, I) :- \sup\left\{ |g_\alpha(\lambda, x) - g^*(\lambda)| \;\middle|\; \lambda \in I,\ 1 - \frac{1}{n+1} \le \alpha < 1 \right\}, \qquad n \in \mathbb{N},$$

and note that

$$\tilde{\Delta}_n(x, I) \ge \tilde{\Delta}_{n+1}(x, I) \ge 0, \qquad n \in \mathbb{N}. \tag{6.3}$$

It must be verified that

$$\tilde{\Delta}_n(x, I) \to 0 \quad \text{as } n \to \infty. \tag{6.4}$$

To achieve this goal, for every $n \in \mathbb{N}$, select

$$\lambda_n \in I \quad \text{and} \quad \alpha_n \in \left[1 - \frac{1}{n+1}, 1\right)$$

such that

$$-\frac{1}{n} \le \tilde{\Delta}_n(x, I) - \frac{1}{n} \le |g_{\alpha_n}(\lambda_n, x) - g^*(\lambda_n)|. \tag{6.5}$$

Recalling that the compact interval I is contained in $(0, \infty)$, there exists a subsequence $\{\lambda_{n_k}\}$ such that

$$\lim_{k\to\infty} \lambda_{n_k} = \lambda^* \in I \subset (0, \infty). \tag{6.6}$$

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

From this point, since $\alpha_{n_k} \to 1$, Theorem 6.1 yields

$$\lim_{k\to\infty} g_{\alpha_{n_k}}(\lambda_{n_k}, x) = g^*(\lambda^*).$$

Now, using the fact that $g^*(\cdot)$ is continuous, by Lemma 4.4, note that (6.6) implies that

$$\lim_{k\to\infty} g^*(\lambda_{n_k}) = g^*(\lambda^*).$$

These last two displays and (6.5) together yield that $\lim_{k\to\infty} \tilde{\Delta}_{n_k}(x, I) = 0$, and (6.4) follows via the monotonicity relation (6.3). $\square$

The remainder of this section is dedicated to presenting the somewhat elaborated proof of Theorem 6.1. To ease the presentation, the essential technical tools have been formulated in Lemmas 6.1–6.3 below, and the corresponding proofs are presented in Appendix A. To begin with, some notation is introduced. Throughout the remainder of this section $\{(\lambda_n, \alpha_n)\}$ is a fixed sequence satisfying (6.1) and (6.2), and Assumptions 2.1, 2.2, and 3.1 are enforced. Now, with regard to Lemma 4.1 and Definition 3.1 and, recalling that S is a finite set and $\mathbb{F}$ is a compact metric space, observe that, via Cantor's diagonal method, without loss of generality it can be assumed that, for every $k \in \mathbb{N}$ and $x, y \in S$, the following limits exist:

$$g_k(x) := \lim_{n\to\infty} g_{\alpha_n}(\lambda_n\alpha_n^k, x),$$

$$h_k(x) := \lim_{n\to\infty} h_{\alpha_n}(\lambda_n\alpha_n^k, x), \qquad f^{(k)}(x) = \lim_{n\to\infty} f^{\lambda_n\alpha_n^k,\alpha_n}(x), \tag{6.7}$$

and

$$Q^{(k)}_{x,y} := \lim_{n\to\infty} Q^{\lambda_n\alpha_n^k,\alpha_n}_{x,y}, \qquad \tilde{Q}^{(k)}_{x,y} := \lim_{n\to\infty} \tilde{Q}^{\lambda_n\alpha_n^k,\alpha_n}_{x,y}. \tag{6.8}$$

Note that $Q^{(k)}$ and $\tilde{Q}^{(k)}$ are stochastic matrices, and that

$$\|g_k(\cdot)\| \le \|C\|, \qquad \|h_k(\cdot)\| \le 2\|C\|B, \qquad k \in S, \tag{6.9}$$

by Lemma 4.1. With the above notation, the conclusions of Theorem 6.1 are equivalent to

$$g_0(\cdot) = g^*(\lambda^*) \quad \text{and} \quad h_0(\cdot) = h^*(\lambda^*, \cdot) = h_1(\cdot). \tag{6.10}$$

The starting point to achieve this goal is the following simple result.

Lemma 6.1. (i) *For every $k \in \mathbb{N}$, the following two relations hold:*

$$\mathrm{e}^{\lambda^* g_k(z)+\lambda^* h_k(x)} = \inf_{a\in A(x)}\Bigg[\mathrm{e}^{\lambda^* C(x,a)} \sum_{y\in S} p_{x,y}(a)\mathrm{e}^{\lambda^* h_{k+1}(y)}\Bigg], \qquad x \in S, \tag{6.11}$$

$$h_k(\cdot) - h_{k+1}(\cdot) = g_k(\cdot) - g_k(z) \quad \textit{and} \quad h_k(z) = 0. \tag{6.12}$$

(ii) *If $g_k(\cdot)$ is constant then $g_k(\cdot) = g^*(\lambda^*)$ and $h_k(\cdot) = h^*(\lambda^*, \cdot) = h_{k+1}(\cdot)$.*

Next, let $x, y \in S$ and $k \in \mathbb{N}$ be arbitrary. Replacing the parameters λ and α in (4.8) by $\lambda_n\alpha_n^k$ and α_n, respectively, and taking the limit in the resulting inequalities, via Assumption 2.1, (6.7), and (6.8), it follows that

$$Q^{(k)}_{x,y} \ge p_{x,y}(f^{(k)}(x))\mathrm{e}^{-4\lambda^* B\|C\|} \quad \text{and} \quad \tilde{Q}^{(k)}_{x,y} \ge p_{x,y}(f^{(k+1)}(x))\mathrm{e}^{-4\lambda^* B\|C\|}. \tag{6.13}$$

Similarly, starting from (4.6) and (4.7), Assumption 2.1, (6.7), and (6.8) together lead to

$$\mathrm{e}^{\lambda^* g_k(x)} \ge \sum_{w\in S} Q^{(k)}_{x,w}\mathrm{e}^{\lambda^* g_{k+1}(w)} \quad \text{and} \quad \mathrm{e}^{\lambda^* g_k(x)} \le \sum_{w\in S} \tilde{Q}^{(k)}_{x,w}\mathrm{e}^{\lambda^* g_{k+1}(w)}. \tag{6.14}$$

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

Now, set

$$M_k := \max_{x\in S} g_k(x), \qquad m_k := \min_{x\in S} g_k(x), \qquad k \in \mathbb{N},$$

and note that (6.14) yields $e^{\lambda^* g_k(x)} \geq \sum_{w\in S} Q^{(k)}_{x,w} e^{m_{k+1}} = e^{m_{k+1}}$, as well as $e^{\lambda^* g_k(x)} \leq \sum_{w\in S} \tilde{Q}^{(k)}_{x,w} e^{\lambda^* M_{k+1}} = e^{\lambda^* M_{k+1}}$ for every $x \in S$ and $k \in \mathbb{N}$, so that, since λ^* is positive,

$$m_{k+1} \leq m_k \leq M_k \leq M_{k+1}, \qquad k \in \mathbb{N},$$

and then the limits

$$M^* := \lim_{k\to\infty} M_k, \qquad m^* := \lim_{k\to\infty} m_k \tag{6.15}$$

exist and belong to $[-\|C\|, \|C\|]$, by (6.9). Observe that the last three displays together yield that

$$m^* \leq g_k(\cdot) \leq M^*, \qquad k \in \mathbb{N}. \tag{6.16}$$

On the other hand, using the fact that the state space is finite, it follows that

(i) for every $n \in \mathbb{N}$, the function $g_n(\cdot)$ has a maximizer $x_n \in S$ so that $g_n(x_n) = M_n$, and

(ii) there exists a state x^* such that $x_n = x^*$ for infinitely many nonnegative integers n.

Therefore, the first convergence in (6.15) yields that

$$\text{there exists } \{n_k\} \subset \mathbb{N} \text{ such that } n_k \nearrow \infty \text{ and } \lim_{k\to\infty} g_{n_k}(x^*) = M^*. \tag{6.17}$$

Similarly, it can be shown that there exists $x_* \in S$ such that

$$\lim_{r\to\infty} g_{n_r}(x_*) = m^* \quad \text{for some sequence } \{n_r\} \subset \mathbb{N} \text{ increasing to } \infty.$$

The following two lemmas, which concern the sequences $\{g_k(\cdot)\}$ and $\{f^{(k)}\}$ in (6.7) and the numbers M^* and m^* in (6.15), are the backbone of the argument that will be used below to establish Theorem 6.1. The corresponding proofs rely on Assumption 3.1.

Lemma 6.2. *Suppose that Assumptions 2.1, 2.2, and 3.1 hold.*

(i) *Let $x \in S$ and the sequence $\{n_k\} \subset \mathbb{N}$ be such that*

$$\lim_{k\to\infty} n_k = \infty, \qquad \lim_{k\to\infty} g_{n_k}(x) = M^*, \qquad \lim_{k\to\infty} f^{(1+n_k)} =: f \in \mathbb{F}. \tag{6.18}$$

In this case,

$$p_{x,y}(f(x)) > 0 \quad \Longrightarrow \quad \lim_{k\to\infty} g_{1+n_k}(y) = M^*.$$

(ii) *There exists a sequence $\{n_k\} \subset \mathbb{N}$ increasing to ∞ such that*

$$\lim_{k\to\infty} g_{n_k}(w) = M^*, \qquad w \in S.$$

Lemma 6.3. *Suppose that Assumptions 2.1, 2.2, and 3.1 hold.*

(i) *If $x \in S$ and $\{n_k\} \subset \mathbb{N}$ satisfy*

$$\lim_{k\to\infty} n_k = \infty, \qquad \lim_{k\to\infty} g_{n_k}(x) = m^*, \qquad \lim_{k\to\infty} f^{(n_k)} =: f \in \mathbb{F},$$

then

$$p_{x,y}(f(x)) > 0 \quad \Longrightarrow \quad \lim_{k\to\infty} g_{1+n_k}(y) = m^*.$$

(ii) *There exists a sequence* $\{n_k\} \subset \mathbb{N}$ *such that* $\lim_{k\to\infty} n_k = \infty$ *and*

$$\lim_{k\to\infty} g_{n_k}(y) = m^*, \qquad y \in S.$$

The following consequence of Lemmas 4.3 and 6.1 will be used to provide a fairly direct proof of Theorem 6.1.

Lemma 6.4. *Suppose that* $L \in \mathbb{R}$ *is such that the convergence*

$$\lim_{k\to\infty} g_{n_k}(x) = L, \qquad x \in S,$$

holds for some sequence $\{n_k\} \subset \mathbb{N}$ *increasing to* ∞*. In this case,* $L = g^*(\lambda^*)$*.*

Proof. Recall that $\|h_{n_k}\| \le 2B\|C\|$ for every k. Taking a subsequence (if necessary), without loss of generality assume that $\lim_{k\to\infty} h_{n_k} =: h \in \mathcal{B}(S)$ exists, and note that, via (6.12), $\lim_{k\to\infty} h_{n_{k+1}}(x) = \lim_{k\to\infty} h_{n_k}(x) - \lim_{k\to\infty}[g_{n_k}(x) - g_{n_k}(z)] = h(x) - [L - L] = h(x)$ for every $x \in S$. Now, replace k by n_k in (6.11) and take the limit as k goes to ∞ on both sides of the resulting equality to obtain, via Lemma 4.3, that the pair $(L, h(\cdot))$ satisfies the optimality equation (2.10) corresponding to λ^*, so that $L = g^*(\lambda^*)$, by Lemma 2.1. □

Proof of Theorem 6.1. As already mentioned, it is sufficient to verify (6.10). To achieve this objective, using Lemma 6.2(ii) select a sequence $\{n_k\} \subset \mathbb{N}$ going to ∞ such that $\lim_{k\to\infty} g_{n_k} = M^*$, and observe that Lemma 6.4 yields $M^* = g^*(\lambda^*)$. Similarly, Lemmas 6.3(ii) and 6.4 together lead to $m^* = g^*(\lambda^*)$, and then $g_k(\cdot) = g^*(\lambda^*)$ for every k, by (6.16); in particular, $g_0(\cdot) = g^*(\lambda^*)$ and (6.10) follows from Lemma 6.1(ii). □

7. Proof of the integral formula

In this section a proof of Theorem 3.3 will be provided. The argument relies on the following lemma, which is a consequence of the uniform convergence results in Theorems 3.1(ii) and 3.2(ii).

Lemma 7.1. *Let* $x \in S$ *and* $\lambda \in \mathbb{R} \setminus \{0\}$ *be arbitrary, and suppose that either of the following conditions are valid:*

(i) $\lambda < 0$ *and Assumptions 2.1 and 2.2 hold;*

(ii) $\lambda > 0$ *and Assumptions 2.1, 2.2, and 3.1 hold.*

In this framework,

$$(1-\alpha)\sum_{k=0}^{\infty} \lambda\alpha^k g_\alpha(\lambda\alpha^k, x) \to \int_0^\lambda g^*(s)\,\mathrm{d}s \quad \text{as } \alpha \nearrow 1. \tag{7.1}$$

Proof. Let $x \in S$, $\lambda \neq 0$, $\varepsilon \in (0, |\lambda|)$, and $\alpha \in (0, 1)$ be arbitrary but fixed. Now, set

$$I_\varepsilon(\lambda) := \begin{cases} [\varepsilon, \lambda], & \lambda > 0, \\ [\lambda, -\varepsilon], & \lambda < 0, \end{cases}$$

whereas, for each nonnegative integer k,

$$I_k(\alpha, \lambda) := \begin{cases} [\lambda\alpha^{k+1}, \lambda\alpha^k] & \text{if } \lambda > 0, \\ [\lambda\alpha^k, \lambda\alpha^{k+1}] & \text{if } \lambda < 0. \end{cases}$$

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

Observe that, under either of the conditions (i) or (ii), Theorems 3.1(ii) and 3.2(ii) imply that

$$\sup_{t\in I_\varepsilon(\lambda)} |g_\alpha(t,x) - g^*(t)| \to 0 \quad \text{as } \alpha \nearrow 1. \tag{7.2}$$

Next, let $k^* \equiv k^*(\varepsilon, \lambda, \alpha)$ be the largest integer k satisfying the $|\lambda|\alpha^k \ge \varepsilon$ so that

$$\lambda\alpha^k \in I_\varepsilon(\lambda), \qquad 1 \le k \le k^*, \tag{7.3}$$

and

$$|\lambda|\alpha^{k^*} \ge \varepsilon > |\lambda|\alpha^{k^*+1}. \tag{7.4}$$

Using the fact that $|g(\lambda\alpha^k, x) - g^*(\lambda\alpha^k)| \le 2\|C\|$ for every $k \in \mathbb{N}$, by (2.8) and (4.1), it follows that

$$\begin{aligned}
&\left|\sum_{k=0}^{\infty}(1-\alpha)\lambda\alpha^k g_\alpha(\lambda\alpha^k, x) - \sum_{k=0}^{\infty}(1-\alpha)\lambda\alpha^k g^*(\lambda\alpha^k)\right| \\
&\quad\le (1-\alpha)\sum_{k=0}^{k^*}|\lambda|\alpha^k|g_\alpha(\lambda\alpha^k, x) - g^*(\lambda\alpha^k)| \\
&\qquad + (1-\alpha)\sum_{k=k^*+1}^{\infty}|\lambda|\alpha^k|g_\alpha(\lambda\alpha^k, x) - g^*(\lambda\alpha^k)| \\
&\quad\le (1-\alpha)\sum_{k=0}^{k^*}|\lambda|\alpha^k|g_\alpha(\lambda\alpha^k, x) - g^*(\lambda\alpha^k)| + (1-\alpha)\sum_{k=k^*+1}^{\infty} 2\|C\||\lambda|\alpha^k \\
&\quad\le (1-\alpha)\sum_{k=0}^{k^*}|\lambda|\alpha^k \sup_{t\in I_\varepsilon(\lambda)} |g_\alpha(t,x) - g^*(t)| + 2\|C\||\lambda|\alpha^{k^*+1} \\
&\quad\le |\lambda| \sup_{t\in I_\varepsilon(\lambda)} |g_\alpha(t,x) - g^*(t)| + 2\|C\|\varepsilon,
\end{aligned}$$

where the third inequality is due to (7.3), and (7.4) was used in the last step. Since $\varepsilon \in (0, |\lambda|)$ is arbitrary, this last display and (7.2) together imply that

$$\lim_{\alpha\nearrow 1}\left[\sum_{k=0}^{\infty}(1-\alpha)\lambda\alpha^k g_\alpha(\lambda\alpha^k, x) - \sum_{k=0}^{\infty}(1-\alpha)\lambda\alpha^k g^*(\lambda\alpha^k)\right] = 0. \tag{7.5}$$

On the other hand, recalling that $g^*(\cdot)$ is continuous, by Lemma 4.4, select $\delta > 0$ such that

$$|g^*(s) - g^*(t)| < \varepsilon \quad \text{if } |s|, |t| \le |\lambda| \text{ and } |s-t| \le \delta.$$

Now, suppose that $\alpha > 1 - \delta/|\lambda|$. Observing that the length of $I_k(\alpha, \lambda)$ is $|\lambda|\alpha^k(1-\alpha) \le |\lambda|(1-\alpha)$ and $\lambda\alpha^k \in I_k(\alpha, \lambda)$, the above display yields $|g^*(\lambda\alpha^k) - g^*(s)| < \varepsilon$ for every $s \in I_k(\alpha, \lambda)$ so that

$$\begin{aligned}
\left|(1-\alpha)\lambda\alpha^k g^*(\lambda\alpha^k) - \int_{\lambda\alpha^{k+1}}^{\lambda\alpha^k} g^*(s)\,\mathrm{d}s\right| &= \left|\int_{\lambda\alpha^{k+1}}^{\lambda\alpha^k}[g^*(\lambda\alpha^k) - g^*(s)]\,\mathrm{d}s\right| \\
&\le \int_{I_k(\alpha,\lambda)} |g^*(\lambda\alpha^k) - g^*(s)|\,\mathrm{d}s \\
&\le \varepsilon|\lambda|\alpha^k(1-\alpha).
\end{aligned}$$

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

Combining this relation with the equality $\int_0^\lambda g^*(s)\,\mathrm{d}s = \sum_{k=0}^\infty \int_{\lambda\alpha^{k+1}}^{\lambda\alpha^k} g^*(s)\,\mathrm{d}s$, it follows that

$$\begin{aligned}\left|\sum_{k=0}^\infty (1-\alpha)\lambda\alpha^k g^*(\lambda\alpha^k) - \int_0^\lambda g^*(s)\,\mathrm{d}s\right| &= \left|\sum_{k=0}^\infty \left[(1-\alpha)\lambda\alpha^k g^*(\lambda\alpha^k) - \int_{\lambda\alpha^{k+1}}^{\lambda\alpha^k} g^*(s)\,\mathrm{d}s\right]\right| \\ &\le \sum_{k=0}^\infty \left|(1-\alpha)\lambda\alpha^k g^*(\lambda\alpha^k) - \int_{\lambda\alpha^{k+1}}^{\lambda\alpha^k} g^*(s)\,\mathrm{d}s\right| \\ &\le \sum_{k=0}^\infty \varepsilon(1-\alpha)|\lambda|\alpha^k \\ &= \varepsilon|\lambda|.\end{aligned}$$

Thus, since $\varepsilon \in (0, |\lambda|)$ is arbitrary, it follows that

$$\lim_{\alpha\nearrow 1} \sum_{k=0}^\infty (1-\alpha)\lambda\alpha^k g^*(\lambda\alpha^k) = \int_0^\lambda g^*(s)\,\mathrm{d}s,$$

and (7.1) follows combining this convergence and (7.5). □

Proof of Theorem 3.3. Suppose that either of the conditions in Lemma 7.1 hold. Let $\lambda \neq 0$, $x \in S$, and $\alpha \in (0, 1)$ be arbitrary, and observe that Definition 3.1 yields

$$\begin{aligned}\sum_{k=0}^n \lambda\alpha^k g_\alpha(\lambda\alpha^k, x) &= \sum_{k=0}^n \lambda\alpha^k [V_\alpha^*(\lambda\alpha^k, x) - \alpha V_\alpha^*(\lambda\alpha^{k+1}, x)] \\ &= \sum_{k=0}^n \lambda\alpha^k V_\alpha^*(\lambda\alpha^k, x) - \sum_{k=0}^n \lambda\alpha^{k+1} V_\alpha^*(\lambda\alpha^{k+1}, x) \\ &= \lambda V_\alpha^*(\lambda, x) - \lambda\alpha^{n+1} V_\alpha^*(\lambda\alpha^{n+1}, x), \qquad n \in \mathbb{N}.\end{aligned}$$

Recalling that $|V_\alpha^*(\cdot, x)| \le \|C\|/(1-\alpha)$ and taking the limit as n goes to ∞, the above relation implies that

$$\lambda V_\alpha^*(\lambda, x) = \sum_{k=0}^\infty \lambda\alpha^k g_\alpha(\lambda\alpha^k, x).$$

This equality yields $(1-\alpha)\lambda V^*(\lambda, x) \to \int_0^\lambda g^*(s)\,\mathrm{d}s$ as $\alpha \nearrow 1$, by Lemma 7.1, and (3.5) follows since $\lambda \neq 0$. □

8. Concluding remarks

In this paper Markov decision chains with risk-sensitive average cost criterion were studied. The basic framework was determined by the standard continuity-compactness conditions in Assumption 2.1, and the communication property in Assumption 2.2, which is necessary to ensure that the optimal average cost function is constant. Within that context, assuming that the controller is risk seeking, it was proved in Theorem 3.1 that appropriate normalizations of the optimal α-discounted value functions converge to the solution of the risk-sensitive average cost optimality equation. Such a result was obtained by taking advantage of the fact that the mapping $\lambda \mapsto \lambda V^*(\lambda, \alpha, x)$ is convex on the negative axis for every discount factor $\alpha \in (0, 1)$

https://doi.org/10.1017/apr.2018.10 Published online by Cambridge University Press

and $x \in S$. In the risk-averse case $\lambda > 0$, $\lambda V^*(\lambda, \alpha, x) = \inf_\pi[\lambda V(\lambda, \alpha, \pi, x)]$, and then, since the infimum of a family of convex functions is not necessarily convex, it follows that, *a priori*, the convexity of the mapping $\lambda \mapsto \lambda V^*(\lambda, \alpha, x)$ cannot be ensured. Thus, to extend the approximation results in Theorem 3.1 to the risk averse case, the additional condition in Assumption 3.1 was required, as well as to follow an alternative route of analysis.

Appendix A. Proofs of Lemmas 6.1–6.3

Proof of Lemma 6.1. (i) Let $k \in \mathbb{N}$ and $x \in S$ be arbitrary. With regard to (2.1) and, keeping in mind that $\lambda > 0$, noting that (3.2) with α_n and $\lambda_n\alpha_n^k$ instead of α and λ, respectively, yields

$$\begin{aligned}&\exp(\lambda_n\alpha_n^k[g_{\alpha_n}(\lambda_n\alpha^k, z) + h_{\alpha_n}(\lambda_n\alpha_n^k, x)])\\ &\quad= \inf_{a\in A(x)}\Bigg[\exp(\lambda_n\alpha_n^k C(x,a))\sum_{y\in S} p_{x,y}(a)\exp(\lambda_n\alpha_n^{k+1}h_{\alpha_n}(\lambda_n\alpha_n^{k+1}, y))\Bigg], \qquad x \in S.\end{aligned}$$

Observe now that (6.2) and (6.7) together imply that $\|\alpha_n^{k+1}h_{\alpha_n}(\lambda_n\alpha_n^{k+1}, \cdot) - h_{k+1}(\cdot)\| \to 0$ as $n \to \infty$, since S is finite. Taking the limit as n goes to ∞ on both sides of the above displayed equality, via (6.2), (6.7), and Lemma 4.3, then (6.11) follows, whereas (6.12) is a consequence of (3.1) and (6.7).

(ii) Suppose that $g_k(\cdot)$ is constant. In this case, $h_k = h_{k+1}$ and $h_k(z) = 0$, by (6.12), whereas (2.1) and (6.11) show that the pair $(g_k(z), h_k(\cdot))$ satisfies the optimality equation (2.10) corresponding to the risk sensitivity coefficient λ^*, and the conclusion follows from Lemma 2.1(iii). □

Proof of Lemma 6.2. (i) Since the matrix $\tilde{Q}^{(n_k)}$ is stochastic on the finite state space S, after taking a subsequence (if necessary), without loss of generality it can be assumed that

$$\lim_{k\to\infty} \tilde{Q}_{w,v}^{(n_k)} =: \tilde{Q}_{w,v}, \qquad w, v \in S,$$

so that, using Assumption 2.1, (6.13), and (6.18) together imply that

$$\tilde{Q}_{w,v} \ge \mathrm{e}^{-4\lambda^* B\|C\|} p_{w,v}(f(w)), \qquad w, v \in S. \tag{A.1}$$

Next, observe that the second inequality in (6.14) yields

$$\mathrm{e}^{\lambda^* g_{n_k}(x)} \le \sum_{w\in S} \tilde{Q}_{x,w}^{(n_k)} \mathrm{e}^{\lambda^* g_{1+n_k}(w)},$$

and, via (6.16), it follows that, for every $y \in S$,

$$\begin{aligned}\mathrm{e}^{\lambda^* g_{n_k}(x)} &\le \tilde{Q}_{x,y}^{(n_k)} \mathrm{e}^{\lambda^* g_{1+n_k}(y)} + \sum_{w\in S\setminus\{y\}} \tilde{Q}_{x,w}^{(n_k)} \mathrm{e}^{\lambda^* g_{1+n_k}(w)}\\ &\le \tilde{Q}_{x,y}^{(n_k)} \mathrm{e}^{\lambda^* g_{1+n_k}(y)} + \sum_{w\in S\setminus\{y\}} \tilde{Q}_{x,w}^{(n_k)} \mathrm{e}^{\lambda^* M^*}\\ &= \tilde{Q}_{x,y}^{(n_k)} \mathrm{e}^{\lambda^* g_{1+n_k}(y)} + (1 - \tilde{Q}_{x,y}^{(n_k)})\mathrm{e}^{\lambda^* M^*}.\end{aligned}$$

Recalling that λ^* is positive, after taking the inferior limit as k goes to ∞ in the above display, via (6.18), it follows that

$$\mathrm{e}^{\lambda^* M^*} \le \tilde{Q}_{x,y} \mathrm{e}^{\lambda^* \liminf_{k\to\infty} g_{1+n_k}(y)} + (1 - \tilde{Q}_{x,y})\mathrm{e}^{\lambda^* M^*},$$

and this immediately yields

$$\tilde{Q}_{x,y} > 0 \quad\Longrightarrow\quad \liminf_{k\to\infty} g_{1+n_k}(y) \geq M^* \quad\Longrightarrow\quad \lim_{k\to\infty} g_{1+n_k}(y) = M^*,$$

where the second implication is due to (6.16). From this point, the conclusion follows observing that $\tilde{Q}_{x,y} > 0$ when $p_{x,y}(f(x)) > 0$, by (A.1).

(ii) The proof of this part relies heavily on Assumption 3.1. Define the class $\mathcal{G}$ of subsets of S as follows: $G \subset S$ belongs to $\mathcal{G}$ if and only if G is nonempty and there exists a sequence $\{n_k\} \subset \mathbb{N}$ going to ∞ such that

$$\lim_{k\to\infty} g_{n_k}(x) = M^*, \qquad x \in G.$$

From this definition, it follows that the desired conclusion is equivalent to the inclusion $S \in \mathcal{G}$. To achieve this goal, observe that $\mathcal{G}$ is nonempty since the singleton $\{x^*\}$ belongs to $\mathcal{G}$, by (6.17). Now, let $G^* \in \mathcal{G}$ be a maximal element of $\mathcal{G}$ with respect to the inclusion relation so that

$$G \in \mathcal{G} \quad \text{and} \quad G^* \subset G \quad\Longrightarrow\quad G^* = G. \tag{A.2}$$

Next, let $\{n_k\} \subset \mathbb{N}$ be a sequence converging to ∞ such that

$$\lim_{k\to\infty} g_{n_k}(x) = M^*, \qquad x \in G^*,$$

and, without loss of generality, assume that $\lim_{k\to\infty} f^{(1+n_k)} =: f$ exists. In this case setting

$$\hat{G} := \{y \in S \mid p_{x,y}(f(x)) > 0 \text{ for some } x \in G^*\}, \tag{A.3}$$

part (i) implies that $\lim_{k\to\infty} g_{1+n_k}(y) = M^*$ for every $y \in \hat{G}$ so that $\hat{G} \in \mathcal{G}$, by the definition of the class $\mathcal{G}$. To conclude, observe that, by Assumption 3.1, the inclusion $G^* \subset \hat{G}$ follows from (A.3). Using (A.2), these two last displays together imply that $G^* = \hat{G}$ so that G^* is closed with respect to the transition matrix $[p_{w,v}(f(w))]_{w,v\in S}$. Since this last matrix is communicating, by Assumption 2.2, it follows that $S = G^* \in \mathcal{G}$, completing the proof. □

Lemma 6.3 can be established by paralleling the argument in the above proof of Lemma 6.2.

Acknowledgements

The authors are deeply grateful to the reviewers for their careful reading of the original manuscript and for their helpful suggestions to improve the paper. This work was partially supported by the PSF Organization (grant number 0116-18) and by CONACYT (grant number 254166).

References

[1] Arapostathis, A. *et al.* (1993). Discrete-time controlled Markov processes with average cost criterion: a survey. *SIAM J. Control Optimization* **31,** 282–344.

[2] Balaji, S. and Meyn, S. P. (2000). Multiplicative ergodicity and large deviations for an irreducible Markov chain. *Stoch. Process. Appl.* **90,** 123–144.

[3] Bäuerle, N. and Rieder, U. (2011). *Markov Decision Processes with Applications to Finance*. Springer, Heidelberg.

[4] Bäuerle, N. and Rieder, U. (2014). More risk-sensitive Markov decision processes. *Math. Operat. Res.* **39,** 105–120.

[5] Bäuerle, N. and Rieder, U. (2017). Zero-sum risk-sensitive stochastic games. *Stoch. Process. Appl.* **127,** 622–642.
[6] Borkar, V. S. and Meyn, S. P. (2002). Risk-sensitive optimal control for Markov decision processes with monotone cost. *Math. Operat. Res.* **27,** 192–209.
[7] Cavazos-Cadena, R. (2009). Solutions of the average cost optimality equation for finite Markov decision chains: risk-sensitive and risk-neutral criteria. *Math. Meth. Operat. Res.* **70,** 541–566.
[8] Cavazos-Cadena, R. and Cruz-Suárez, D. (2017). Discounted approximations to the risk-sensitive average cost in finite Markov chains. *J. Math. Anal. Appl.* **450,** 1345–1362.
[9] Cavazos-Cadena, R. and Fernández-Gaucherand, E. (2000). The vanishing discount approach in Markov chains with risk-sensitive criteria. *IEEE Trans. Automatic Control* **45,** 1800–1816.
[10] Chávez-Rodríguez, S., Cavazos-Cadena, R. and Cruz-Suárez, H. (2015). Continuity of the optimal average cost in Markov decision chains with small risk-sensitivity. *Math. Meth. Operat. Res.* **81,** 269–298.
[11] Denardo, E. V. and Rothblum, U. G. (2006). A turnpike theorem for a risk-sensitive Markov decision process with stopping. *SIAM J. Control Optimization* **45,** 414–431.
[12] Di Masi, G. B. and Stettner, L. (1999). Risk-sensitive control of discrete-time Markov processes with infinite horizon. *SIAM J. Control Optimization* **38,** 61–78.
[13] Di Masi, G. B. and Stettner, L. (2000). Infinite horizon risk sensitive control of discrete time Markov processes with small risk. *Systems Control Lett.* **40,** 15–20.
[14] Di Masi, G. B. and Stettner, L. (2007). Infinite horizon risk sensitive control of discrete time Markov processes under minorization property. *SIAM J. Control Optimization* **46,** 231–252.
[15] Hernández-Hernández, D. and Marcus, S. I. (1996). Risk sensitive control of Markov processes in countable state space. *Systems Control Lett.* **29,** 147–155.
[16] Hernández-Hernández, D. and Marcus, S. I. (1999). Existence of risk-sensitive optimal stationary polices for controlled Markov processes. *Appl. Math. Optimization* **40,** 273–285.
[17] Hernández-Lerma, O. (1989). *Adaptive Markov Control Processes*. Springer, New York.
[18] Howard, R. A. and Matheson, J. E. (1972). Risk-sensitive Markov decision processes. *Manag. Sci.* **18,** 356–369.
[19] Jaśkiewicz, A. (2007). Average optimality for risk sensitive control with general state space. *Ann. Appl. Prob.* **17,** 654–675.
[20] Kontoyiannis, I. and Meyn, S. P. (2003). Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Prob.* **13,** 304–362.
[21] Pitera, M. and Stettner, L. (2016). Long run risk sensitive portfolio with general factors. *Math. Meth. Operat. Res.* **83,** 265–293.
[22] Puterman, M. L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley, New York.
[23] Shen, Y., Stannat, W. and Obermayer, K. (2013). Risk-sensitive Markov control processes. *SIAM J. Control Optimization* **51,** 3652–3672.
[24] Sladký, K. (2008). Growth rates and average optimality in risk-sensitive Markov decision chains. *Kybernetika* **44,** 205–226.
[25] Stettner, L. (1999). Risk sensitive portfolio optimization. *Math. Meth. Operat. Res.* **50,** 463–474.
[26] Thomas, L. C. (1981). Connectedness conditions for denumerable state Markov decision processes. In *Recent Developments in Markov Decision Processes*, Academic Press, New York, pp. 181–204.
[27] Tijms, H. C. (2003). *A First Course in Stochastic Models*. John Wiley, Chichester.