SIAM J. CONTROL OPTIM.
Vol. 38, No. 4, pp. 1204–1222
© 2000 Society for Industrial and Applied Mathematics

CONTINUITY OF OPTIMAL VALUES AND SOLUTIONS FOR CONTROL OF MARKOV CHAINS WITH CONSTRAINTS*

MABEL M. TIDBALL†, ARIEL LOMBARDI‡, ODILE POURTALLIER§, AND EITAN ALTMAN§

Abstract. We consider in this paper constrained Markov decision processes. This type of control model has many applications in telecommunications and other fields [E. Altman and A. Shwartz, *IEEE Trans. Automat. Control*, 34 (1989), pp. 1089–1102, E. A. Feinberg and M. I. Reiman, *Probab. Engrg. Inform. Sci.*, 8 (1994), pp. 463–489, A. Hordijk and F. Spieksma, *Adv. in Appl. Probab.*, 21 (1989), pp. 409–431, A. Lazar, *IEEE Trans. Automat. Control*, 28 (1983), pp. 1001–1007, P. Nain and K. W. Ross, *IEEE Trans. Automat. Control*, 31 (1986), pp. 883–888, K. W. Ross and B. Chen, *IEEE Trans. Automat. Control*, 33 (1988), pp. 261–267]. We address the issue of the convergence of the value and optimal policies of the problem with discounted costs, to the ones for the problem with expected average cost. We consider the general multichain ergodic structure. We present two stability results in this paper. We establish the continuity of optimal values and solutions of as well as some type of robustness of some suboptimal solutions in the discount factor. Our proof relies on same general theory on continuity of values and solutions in convex optimization that relies on well-known notions of Γ-convergence.

Key words. optimization, sensitivity analysis, constrained Markov decision processes

AMS subject classifications. 93E20, 93B35, 90C40

PII. S0363012997280294

1. Introduction.

We consider a sequence $\mathbf{MP}_n$, $n = 1, 2, \ldots$ of constrained Markov decision processes (CMDPs), and a "limit" one, denoted by $\mathbf{MP}_\infty$, or simply by **MP**. These are defined on some vector spaces, possibly infinite dimensional ones. **MP** is assumed to be feasible (it has at least one solution). However, for any given n, $\mathbf{MP}_n$ need not be feasible, and even if it is, it need not possess an optimal solution (i.e., it may only have ϵ-optimal solutions). We are interested in the following questions.

(i) Do the values of $\mathbf{MP}_n$ converge to the value of **MP**?
(ii) Do optimal (or almost optimal) policies converge in some sense?
(iii) Given an (almost) optimal policy for $\mathbf{MP}_n$, will it be an almost optimal policy for **MP** if n is sufficiently large?
(iv) Conversely, given an optimal policy for **MP**, will it be an almost optimal policy for $\mathbf{MP}_n$ for all n sufficiently large?

By reducing our control problem to equivalent mathematical programs, we show that the epigraph theory and the Γ-convergence theory provide sufficient conditions for having convergence in the sense of (i) and (ii) above. It turns out that the answers for (iii) and for (iv) are in general negative, unlike the unconstrained case. The reason is that an optimal policy for $\mathbf{MP}_n$ may be unfeasible for **MP**, and vice versa. We shall, however, establish sufficient conditions for the following slightly weaker versions of (iii) and (iv).

*Received by the editors December 2, 1997; accepted for publication (in revised form) August 23, 1999; published electronically April 18, 2000.
http://www.siam.org/journals/sicon/38-4/28029.html
†Institut National de Recherche en Agronomie, Economie et Sociologie Rurales, 2 Place Viala 34060, Montpellier Cedex 1, France (tidball@ensam.inra.fr).
‡Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires. Ciudad Universtaria, 1428 Capital Federal, Argentina (aldoc7@mate.dm.uba.ar).
§INRIA, Centre Sophia-Antipolis, 2004 Route des Lucioles, B.P. 93, 06902 Sophia-Antipolis Cedex, France (pourtali@sophia.inria.fr, altman@sophia.inria.fr).

(iii′) Given an optimal policy for $\mathbf{MP}_n$, can we perturb it "slightly" so that it becomes almost optimal for $\mathbf{MP}$ if n is sufficiently large?

(iv′) Given an optimal policy for $\mathbf{MP}$, can we perturb it "slightly" so that it becomes almost optimal for $\mathbf{MP}_n$ for all n sufficiently large?

The answers for (iii′) and (iv′) follow from continuity properties of the solutions and values of mathematical programming. We use in particular notions of Γ-convergence (see [16] and [37]). We extend and adapt existing theorems and tools for the sensitivity of solutions of convex optimization to obtain the appropriate framework for analyzing CMDPs.

We focus here on the convergence in the discount factor α_n as it converges to some limit within the unit interval. We present sufficient conditions for the convergence of the values and of optimal policies, as well as some robustness properties of suboptimal policies. Related results were already obtained in [1] but had a strong restriction on the ergodic structure of the controlled model. It was required to have a single ergodic class under any stationary policy. This condition enabled us to restrict our problem to stationary policies, for which some general theorem on approximation [2] could be used. In the present paper we make no assumption on the ergodic structure, thus allowing a multichain situation. For such ergodic structure, it is known that stationary policies need not be optimal (nor even ϵ-optimal) for the expected average cost criterion, and one has to use either Markov policies (see [25, 27]) or mixed-stationary policies (this term was raised by Feinberg [21] in a similar context; it refers to policies that are highly nonstationary). We use the latter approach to establish, with the help of the results from the first part of the paper, the convergence of the values and policies.

We briefly mention some related work on the continuity and sensitivity analysis of mathematical programs and of control problems. Many papers and books studied similar problems in the case of finite dimensional state, e.g., [17, 24, 36]. Several special issues of scientific journals were devoted to these questions, as well as other related sensitivity, stability, and parametric analysis: *Mathematical Programming* 21, 1984, *Annals of Operations Research* 27, 1990. Convergence results for constrained dynamic control problems were obtained in [1, 2, 4, 6, 8]. Conditions were obtained there for the convergence in the transition probabilities, in the horizon, and in the immediate cost. These results were applied to adaptive control problems [6] and to problems of finite state approximations of CMDPs. Similar questions to those addressed in this paper were studied in [23] and in [43], and some of the results there are close to those in the first part of our paper. The main difference lies in the types of assumptions made. In [23] the central assumption is stated in terms of constraints set convergence making the use of a metric. In [43] less regularity on the constraints sets are required (convergence of the constraint sets in the Hausdorff topology). Nevertheless only points (i) and (ii) are studied there. Some related questions but in the context of min-max problems and Stakelberg equilibrium are studied in [31, 32, 34]. Some other references related to the current paper are [13, 16, 28, 33, 42, 44].

2. CMDPs: The convergence in the discount factor. We consider CMDPs, known also as controlled Markov chains, with a general ergodic structure (multichain). We consider the discounted cost and the average cost. We shall obtain new results on the convergence of the values and optimal policies of the discounted cost, as the discount factor tends to one, to the value and to optimal policies corresponding to the expected average cost. A similar result for the special unichain case was obtained

in [1].

Consider a Markov decision process (MDP) with a finite state space $\mathbf{X} = \{0, 1, \ldots, N\}$ and a finite *action space* $\mathbf{A}$. Without loss of generality, we assume that in any state x all actions in $\mathbf{A}$ are available. The probability to go from state x to state y given that action a is used is given by the transition probability P_{xay}. A policy u in the *policy space* $\mathbf{U}_h$ is described as $u = (u_1, u_2, \ldots)$, where u_t, applied at time epoch t, is a probability measure over $\mathbf{A}$ conditioned on the whole history of actions and state prior to t, as well as the state at time t. Given an initial distribution β on $\mathbf{X}$, each policy u induces a probability measure denoted by P^u_β on the space of sample paths of states and actions (which serves as the canonical sample space Ω). The corresponding expectation operator is denoted by E^u_β. On this probability space are defined the state and action processes, X_t, A_t, $t = 1, 2, \ldots$.

A *Markov policy* $u \in \mathbf{U}_M$ is characterized by the dependence of u_{t+1} on the current state and the time only. A *stationary policy* $g \in \mathbf{U}_S$ is characterized by a single conditional probability measure $p^g_{\cdot|x}$ over $\mathbf{A}$, so that $p^g_{\mathbf{A}|x} = 1$; under g, X_t becomes a Markov chain with stationary transition probabilities, given by $P^g_{xy} = \sum_{a\in\mathbf{A}} p^g_{a|x} P_{xay}$. The class of *stationary deterministic policies* $\mathbf{U}_D$ is a subclass of $\mathbf{U}_S$, and every $g \in \mathbf{U}_D$ is characterized by a mapping g:$\mathbf{X} \to \mathbf{A}$, so that $p^g_{\cdot|x} = \delta_{g(x)}(\cdot)$ is concentrated at the point $g(x)$ for each x. Let L be the number of stationary deterministic policies among $\mathbf{U}_D$, and enumerate the policies in $\mathbf{U}_D$ such that $\mathbf{U}_D = \{u^1, \ldots, u^L\}$.

It will often be useful to extend the definition of a policy $u = (u_1, u_2, \ldots)$ so as to allow u_t to depend not only on the history, but also on some additional randomizing mechanism. In particular, for any finite class of policies $G \subset \mathbf{U}_h$, we define $\overline{M}(G)$ to be the class of mixed policies generated by G. We call these mixed-G policies. A mixed-G policy $\hat{q}$ is identified with a distribution q over G; the controller first uses q to choose some policy $u \in G$ and then proceeds with that policy from time 1 onwards. Define $\mathcal{U} := \overline{M}(U_D)$. Finally, we denote $\mathbf{U} = \mathbf{U}_h \cup \mathcal{U}$.

DEFINITION 2.1. *For any initial distribution q over the set $\mathbf{U}_D$, we shall identify the policy $m(q) \in \mathcal{U}$ to be the one that chooses initially the policy u^j with probability q_j.*

Remark 2.2. Although we consider here a more general definition of policies than $\mathbf{U}_h$, these policies are in fact equivalent to those in $\mathbf{U}_h$ (details are given in Remark 3.1). A randomization over actual classes of policies has already been considered in [20, 29]. An alternative for the randomization over policies is the randomization over the strategic measures they generate. If we identify a policy with the strategic measure it generates, then it follows that the large class $\overline{M}(\mathbf{U}_h)$ is equivalent to the class of Markov policies $\mathbf{U}_M$, as was shown in [19].

For any given distribution β for the initial state (at time 1) and a policy u, define a probability measure P^u_β on which the stochastic processes X_t and A_t of the states and actions are defined. When β is concentrated on some state x (i.e., $\beta = \delta_x$), we shall use the notation P^u_x instead of P^u_β.

Let $c : \mathbf{X} \times \mathbf{A} \to \mathbb{R}$ and $d : \mathbf{X} \times \mathbf{A} \to \mathbb{R}^K$ be immediate cost functions, $d = (d^1, d^2, \ldots, d^k)$.

Fix some discount factor $\alpha \in [0, 1)$, and define the normalized discounted costs corresponding to an initial distribution β and a policy u by

$$C_\alpha(\beta, u) = \sum_{t=1}^{\infty} (1-\alpha) E^u_\beta \alpha^{t-1} c(X_t, A_t),$$

$$D^k_\alpha(\beta,u)=\sum_{t=1}^{\infty}(1-\alpha)E^u_\beta\alpha^{t-1}d^k(X_t,A_t),\qquad k=1,\dots,K.$$

Define the average costs associated with a policy u and with an initial distribution β on $\mathbf{X}$:

$$C_{ea}(\beta,u)=\overline{\lim_{t\to\infty}}\,\frac{1}{t}E^u_\beta\left[\sum_{s=1}^{t}c(X_s,A_s)\right],$$
$$D^k_{ea}(\beta,u)=\overline{\lim_{t\to\infty}}\,\frac{1}{t}E^u_\beta\left[\sum_{s=1}^{t}d^k(X_s,A_s)\right],\qquad k=1,\dots,K.$$

Given a vector $V\in\mathbb{R}^K$, we consider the subset $\Pi_V\subset\mathbf{U}$ of policies satisfying the constraints

$$D(\beta,u)\le V. \tag{2.1}$$

A policy $u\in\Pi_V$ is called feasible. We introduce the following constrained problem (COP): find a policy u^* that achieves

$$C(\beta):=\inf_{u\in\Pi_V}C(\beta,u). \tag{2.2}$$

In (2.1) and (2.2), the costs stand for either the discounted or the average cost. The COP is said to be feasible if Π_V is nonempty.

We are now ready to state the first main result.

THEOREM 2.3 (convergence of the value). *Assume that the following Slater condition holds: there exists some policy $v\in\mathbf{U}$ such that*

$$D^k(\beta,v)\le V^k-\eta\qquad\forall k=1,\dots,K \tag{2.3}$$

for some $\eta>0$. Then, the value converges in the discount factor:

$$\lim_{\alpha\to1}C_\alpha(\beta)=C_{ea}(\beta).$$

The proof of this theorem, as well as the convergence of optimal policies, are delayed to the next section.

3. Convergence of the values and the policies. In order to be able to define the convergence of optimal policies, we shall show that one may restrict the search of optimal policies to the simple subclasses of policies $\mathcal{U}$, without loss of optimality. Moreover, we should relate the solutions of the COP to solutions of mathematical programming, in order to be able to apply the tools that we developed. Note that the control problem is already of the form of a mathematical program, but the cost is not convex in the policies. We shall show that when restricting to $\mathcal{U}$, the costs are convex functions.

There are several ways to solve (2.2). For the discounted cost, the solution was given by Kallenberg in [27] using a linear program (LP) approach. For the expected average cost, there are several possible LP approaches: the one by Hordijk and Kallenberg [25, 27], the one by Feinberg [21], and a related one by Altman and Shwartz [9]; for a definition slightly different than in (2.2), an efficient LP method for computing ϵ-optimal solutions was obtained by Ross and Varadarajan; see [40, 41]. Lagrangian

techniques have also been used to solve CMDPs with a single constraint; see Beutler and Ross [11, 12] and Sennott [45, 46]. The relation between the Lagrange and the LP approaches was pointed out in [10].

Remark 3.1. All the references above considered the solution of the constrained MDP among the policies $\mathbf{U}_h$. However, a standard argument due to Derman and Strauch [18] shows (in a constructive way) that for any policy $u \in \mathbf{U}$, there exists an equivalent policy $\chi \in \mathbf{U}_M$ under which the marginal probabilities of the states and actions are the same as those under u, and in particular, both the discounted and the expected average costs are the same. Thus below, whenever we obtain an optimal policy among $\mathcal{U}$, one may consider the policy χ instead without loss of optimality. Note, however, that one cannot in general restrict further to the stationary policies $\mathbf{U}_S$. Indeed, it is shown in [25, 27] that they are not sufficient for the expected average cost with a general multichain structure.

We would like to prove the convergence results by showing that there is a correspondence between values and optimal solutions of the control problem, and values and optimal solutions of related LPs, and then use the general results from the previous section. While this is possible for the unichain case, it turns out that for the multichain case, the LP introduced by Kallenberg [27] for the discounted cost is completely different than any of the LPs for the expected average cost (e.g., the number of decision variables is different). Therefore, as a first step, we shall introduce a new LP method for computing the value and optimal policies for the discounted cost problem, which is an adaptation of the one for the expected average cost in Feinberg [21] and Altman and Shwartz [9]. This will allow us to have the same type of LP for both the discounted and the average cost.

Denote the simplex $S(L) := \{\gamma \in R^L : \gamma_i \geq 0, i = 1, \ldots, L, \ \sum_{i=1}^L \gamma_i = 1\}$. Introduce the following $\mathbf{LP}_\alpha$: Find $\gamma^* \in S(L)$ that achieves

$$\mathcal{C}_\alpha^* := \min_{\gamma \in S(L)} \sum_{i=1}^L \gamma_i C_\alpha(\beta, u^i), \qquad \textit{such that } (s.t.) \tag{3.1}$$

$$\mathcal{D}_\alpha^k(\gamma) := \sum_{i=1}^L \gamma_i D_\alpha^k(\beta, u^i) \leq V^k, \ k = 1, \ldots, K. \tag{3.2}$$

Define $\mathcal{C}_\alpha(\gamma) := \sum_{i=1}^L \gamma_i C_\alpha(\beta, u^i)$. We say that the LP is feasible if the subset of $S(L)$ satisfying the constraints (3.2) is nonempty.

THEOREM 3.2 (relation between LP and the CMDP, the discounted cost).

(i) *For any $\gamma \in S(L)$, the policy $m(\gamma) \in \mathcal{U}$ (see Definition 2.1) satisfies*

$$C_\alpha(\beta, m(\gamma)) = \mathcal{C}_\alpha(\gamma), \qquad D_\alpha^k(\beta, m(\gamma)) = \mathcal{D}_\alpha^k(\gamma), \qquad k = 1, \ldots, K.$$

(ii) *For any vector of costs*

$$\{C_\alpha(\beta, u), D_\alpha^k(\beta, u), k = 1, \ldots, K\},$$

achievable by some policy $u \in \mathbf{U}$, there exists some $v \in \mathcal{U}$ achieving the same vector of costs.

(iii) *COP_α is feasible if and only if LP_α is and the optimal values are the same: $\mathcal{C}_\alpha^* = C_\alpha(\beta)$. Moreover, if γ^* is optimal for LP_α, then $m(\gamma^*)$ is optimal for COP_α.*

Proof. Denote

$$f_\alpha(\beta,u;y,a) := (1-\alpha)\sum_{t=1}^{\infty}\alpha^{t-1}P_\beta^u(X_t=y,A_t=a), \qquad x\in\mathbf{X}, a\in\mathbf{A},$$

and let $f_\alpha(\beta,u)$ be the vector whose (y,a)th elements are given by $f_\alpha(\beta,x;y,a)$. Then (see, e.g., [14])

$$C_\alpha(\beta,u) = c\cdot f_\alpha(\beta,u) = \sum_{y\in\mathbf{X}}\sum_{a\in\mathbf{A}} c(y,a) f_\alpha(\beta,u;y,a), \tag{3.3}$$

with a similar representation for the costs $D_\alpha^k(\beta,u)$. For any class of policies $\overline{U}$, denote $\mathbf{L}_\alpha(\beta,\overline{U}) = \cup_{u\in\overline{U}} f_\alpha(\beta,u)$. It is known that the set $\mathbf{L}_\alpha(\beta,\mathbf{U}_h)$ is convex, compact, and its extreme points are contained in $\{f_\alpha(\beta,u), u\in\mathbf{U}_D\}$; see [14, 27]. For any probability γ over $\mathbf{U}_D$, we clearly have

$$f_\alpha(\beta,m(\gamma)) = \sum_{i=1}^{L}\gamma_i f_\alpha(\beta,u^i). \tag{3.4}$$

Consequently, $\mathbf{L}_\alpha(\beta,\mathbf{U})$ is convex, compact, and its extreme points are $\{f_\alpha(\beta,u), u\in\mathbf{U}_D\}$.

Combining this with (3.3), we conclude that the set of achievable costs

$$\cup_{u\in\mathbf{U}}\{C_\alpha(\beta,u), D_\alpha^k(\beta,u), k=1,\dots,K\} \tag{3.5}$$

is also convex, compact, and its extreme points are

$$\{C_\alpha(\beta,u), D_\alpha^k(\beta,u), k=1,\dots,K, u\in\mathbf{U}_D\}. \tag{3.6}$$

By combining (3.3) with (3.4), we get, for any probability γ over $\mathbf{U}_D$,

$$C_\alpha(\beta,m(\gamma)) = \sum_{i=1}^{L}\gamma_i C_\alpha(\beta,u^i), \qquad D_\alpha^k(\beta,m(\gamma)) = \sum_{i=1}^{L}\gamma_i D_\alpha(\beta,u^i),\ k=1,\dots,K.$$

Hence, the set of performance measures achievable by $u\in\mathcal{U}$ is also convex, compact, with the extreme points in the set (3.6), and thus, equal to the set (3.5) achievable by all policies. This establishes (i) and (ii) and implies (iii). □

The LP method corresponding to (3.1) for the expected average cost, due to Feinberg [21], is the same: $\mathbf{LP_{ea}}$: Find $\gamma^*\in S(K)$ that achieves

$$\mathcal{C}_{ea}^* := \min_{\gamma\in S(L)}\sum_{i=1}^{L}\gamma_i C_{ea}(\beta,u^i), \qquad s.t. \tag{3.7}$$

$$\mathcal{D}_{ea}^k(\gamma) := \sum_{i=1}^{L}\gamma_i D_{ea}^k(\beta,u^i) \le V^k,\ k=1,\dots,K. \tag{3.8}$$

Define $\mathcal{C}_{ea}(\gamma) := \sum_{i=1}^{L}\gamma_i C_{ea}(\beta,u^i)$.

THEOREM 3.3 (relation between LP and the CMDP, the expected average cost).

(i) *For any $\gamma\in S(L)$, the policy $m(\gamma)\in\mathcal{U}$ (see Definition 2.1) satisfies*

$$C_{ea}(\beta,m(\gamma)) = \mathcal{C}_{ea}(\gamma), \qquad D_{ea}^k(\beta,m(\gamma)) = \mathcal{D}_{ea}^k(\gamma), \qquad k=1,\dots,K.$$

(ii) *For any vector of costs*

$$\{C_{ea}(\beta,u), D^k_{ea}(\beta,u), k=1,\dots,K\}$$

achievable by some policy $u\in\mathbf{U}$*, there exists a dominating* $v\in\mathcal{U}$*, i.e., such that*

$$C_{ea}(\beta,v)\le C_{ea}(\beta,u),\quad D^k_{ea}(\beta,v)\le D^k_{ea}(\beta,u), k=1,\dots,K.$$

(iii) COP_{ea} *is feasible if and only if* LP_{ea} *is and the optimal values are the same:* $\mathcal{C}^*_{ea}=C_{ea}(\beta)$*. Moreover, if* γ^* *is optimal for* LP_{ea}*, then* $m(\gamma^*)$ *is optimal for* COP_{ea}*.*

Proof. Denote

$$f^t_{ea}(\beta,u;y,a):=t^{-1}\sum_{s=1}^{t}P^u_\beta(X_s=y,A_s=a),\qquad x\in\mathbf{X},a\in\mathbf{A},$$

and let $f^t_{ea}(\beta,u)$ be the vector whose (y,a)th elements are given by $f^t_{ea}(\beta,x;y,a)$. For any $u\in U_D$, since the state process is a Markov chain, it is known that

$$f_{ea}(\beta,u)=\lim_{t\to\infty}f^t_{ea}(\beta,u)\ \text{ exists,}\tag{3.9}$$

and it is straightforward to show that

$$C_{ea}(\beta,u)=c\cdot f_{ea}(\beta,u)=\sum_{y\in\mathbf{X}}\sum_{a\in\mathbf{A}}c(y,a)f_{ea}(\beta,u;y,a),\tag{3.10}$$

with a similar representation for the costs $D^k_{ea}(\beta,u)$. It is then clear that (3.9) and (3.10) hold in fact for any $u\in\mathcal{U}$, which establishes (i).

For a fixed initial distribution β, for any policy $v\in\mathbf{U}_h$, and for any accumulation point f of the sequence $f^t_{ea}(\beta,u)$, there exists some $u\in\mathcal{U}$ such that $f_{ea}(\beta,u)=f$. This is a direct consequence of Theorem 2 in [25] and is a special case of the result in [21] (who studies the semi-Markov case). Combining this with the fact that (3.9) and (3.10) hold for any $u\in\mathcal{U}$ establishes (ii), by using Corollary 2.5 in [7]. Finally, (iii) is a consequence of (i) and (ii). □

For a given $u\in\mathcal{U}$ we shall understand below $\pi_\delta(u)=m(\pi_\delta(\gamma))$, where γ is such that $u=m(\gamma)$ and π_δ is defined in (B.17). We are now ready to state the second main result for MDPs.

THEOREM 3.4. *Assume that the Slater condition* (2.3) *holds. Consider a sequence* α_n *converging to* 1*, and let* COP_n *be the constrained optimal control problem corresponding to the discount factor* α_n*. Let* δ *be such that* $\eta>\delta>0$*. Then we have the following.*

(i) *Let* $u_n\in\mathcal{U}$ *be* ϵ_n*-optimal for* COP_n*,* $\limsup_n\epsilon_n\le\epsilon$*. Then there exists* $N(\epsilon,\delta)$ *such that* $\forall n\ge N(\epsilon,\delta)$*,* $\pi_\delta(u_n)$ *is* $O(\epsilon+\delta)$*-optimal for* COP_{ea}*.*

(ii) *Let* $u\in\mathcal{U}$ *be optimal for* COP_{ea}*. Then there exists* $N(\epsilon,\delta)$ *such that* $\forall n\ge N(\epsilon,\delta)$*,* $\pi_\delta(u)$ *is* $O(\epsilon+\delta)$*-optimal for* COP_n*.*

(iii) *Let* $u_n\in\mathcal{U}$ *be optimal for* COP_n *and let* $\gamma_n\in S(L)$ *be such that* $u_n=m(\gamma_n)$*. Assume that* γ_n *converges to some* γ*. Then* $m(\gamma)$ *is optimal for* COP_{ea}*.*

Proof of Theorems 2.3 *and* 3.4. We apply below Theorems B.1, B.4, and B.15, which present sensitivity results for general convex optimization problems, to obtain the convergence of the optimal values and the convergence and robustness of policies

LP_α to the optimal value of LP_{ea}, and consequently, by Theorems 3.2(ii), (iii), and 3.3(ii), (iii), the convergence for the original constrained optimal control problems. It remains to show that the required conditions in Theorems B.1, B.4, and B.15 hold.

It is well known that for any $u \in U_D$,

$$\lim_{\alpha\to 1} C_\alpha(\beta,u) = C_{ea}(\beta,u), \qquad \lim_{\alpha\to 1} D^k_\alpha(\beta,u) = D^k_{ea}(\beta,u),\ k=1,\dots,K. \tag{3.11}$$

Choose an arbitrary $u \in \mathcal{U}$, and let γ be such that $u = m(\gamma)$. Then, due to Theorems 3.2(i) and 3.3(i),

$$\begin{aligned} &|C_\alpha(\beta,u) - C_{ea}(\beta,u)| \\ &\le \left| \sum_{j=1}^{L} \gamma_j \left[C_\alpha(\beta,u^j) - C_{ea}(\beta,u^j) \right] \right| \\ &\le \max_{1\le j\le L} |C_\alpha(\beta,u^j) - C_{ea}(\beta,u^j)|, \end{aligned}$$

which does not depend on u.

With the same applied to the costs D^k, this implies that (3.11) holds for any $u \in \mathcal{U}$, and that the convergence is uniform over $\mathcal{U}$. Equivalently, for any $\gamma \in S(L)$,

$$\lim_{\alpha\to 1} \mathcal{C}_\alpha(\gamma) = \mathcal{C}_{ea}(\gamma), \qquad \lim_{\alpha\to 1} \mathcal{D}^k_\alpha(\gamma) = \mathcal{D}^k_{ea}(\gamma),\ k=1,\dots,K \tag{3.12}$$

uniformly in γ. This establishes conditions (A.2) and (A.3).

The set $S(L)$ is clearly convex. As the costs are linear in γ, conditions (A.4) and (A.5) hold. Since γ is bounded in a simplex, this implies condition (A.6).

It follows from condition (2.3) and from Theorem 3.3 that there exists some $\eta > 0$ and some $\gamma \in S(L)$ such that the policy $m(\gamma)$ satisfies

$$\mathcal{D}^k_{ea}(\gamma) = D^k_{ea}(\beta, m(\gamma)) \le V^k - \eta, \qquad k = 1,\dots,K.$$

This establishes condition (A.7). Finally, condition (A.8) trivially holds, as $V = V_n$ do not depend on n. □

4. Concluding remarks. We have presented sufficient conditions for the continuity of the optimal values of constrained optimization problems and established several results on convergence of optimal solutions. Using these general tools, we obtained a new result for the convergence of discounted MDPs to the expected average cost one, under a general multichain ergodic structure. This was done by showing that one could restrict without loss of optimality to some subclass of policies, and then that an equivalent LP can be used to compute the values and the optimal policies. Since our results in the first part of the paper hold for convex programs and not just for LPs, this suggests that they could be used for establishing convergence properties in control problems with more complex cost functions.

One can further use the continuity of the optimal values and solutions of constrained opimtization to obtain other important features in constrained control of Markov chains. One can obtain convergence of the values and optimal policies of finite horizon problems to infinite ones, and one can establish the convergence of problems with finite state spaces to those with infinite state spaces (see [1, 3]). Another interesting question is on the structure of optimal policies in CMDPs. Ross [38] has shown that CMDPs with finite state and action spaces have optimal stationary policies that require at most $K+1$ randomizations, where K is the number of

constraints. Borkar [15] has extended this to a countable state space. Our approach allows us to obtain the same result as Borkar using the fact that limits of policies that are optimal for truncated (finite state) problems (which can be chosen with no more than $K+1$ randomization, according to [38]) are optimal for the original countable state problem (see details in [1, 3]).

Appendix A. Solutions of convex optimization. Let X be a convex subset of a topological vector space $V_n = (V_n^1, \dots, V_n^K) \in \mathbb{R}^K$, $n = 1, 2, \dots, \infty$, $(V_\infty = V)$ and

$$C_n : X \to \mathbb{R}, \quad n = 1, \dots, \infty \quad (C_\infty = C),$$

$$D_n : X \to \mathbb{R}^K, \quad n = 1, \dots, \infty \quad (D_\infty = D),$$

$$\Delta_n = \{x \in X : D_n(x) \le V_n\}, \quad n = 1, 2, \dots, \infty \quad (\Delta_\infty = \Delta).$$

We define the values of the constrained problems:

$$R_n = \inf_{u \in \Delta_n} C_n(u); \qquad R = \inf_{u \in \Delta} C(u).$$

We want to answer the following questions.

(i) Does $R_n \to R$ when $n \to \infty$?

(ii) Convergence of policies: Let $\pi : X \to \Delta$, and fix some $\epsilon \ge 0$. Let ϵ_n be a sequence of positive real numbers such that $\overline{\lim}_{n\to\infty} \epsilon_n \le \epsilon$. Assume that u_n^* is an ϵ_n-optimal policy for the nth optimal cost function R_n. Is $\pi(u_n^*)$ "almost" optimal for the limit optimal cost function R, for n large enough?

(iii′) Robustness of the optimal policy: If u^* is ϵ-optimal for the limit optimal cost function, can we derive of it an "almost" optimal policy for the nth approximating optimal cost function for all n large enough?

(iv′) Let $\overline{u} \in X$ be some limit point of u_n^*, defined above. Is $\overline{u}$ ϵ-optimal for the limit optimal cost function?

We can give an answer to questions (i) and (iv′) using the notion of Γ-convergence [16]. In fact, let X be a topological space, $\mathcal{N}(u)$ the set of all open neighborhoods of $u \in X$, and F_n a sequence of functions from X to $\mathbb{R} \cup \infty$.

We define

$$\Gamma - \underline{\lim}_{n\to\infty} F_n(u) = \sup_{V \in \mathcal{N}(u)} \underline{\lim}_{n\to\infty} \inf_{y \in V} F_n(y),$$

$$\Gamma - \overline{\lim}_{n\to\infty} F_n(u) = \sup_{V \in \mathcal{N}(u)} \overline{\lim}_{n\to\infty} \inf_{y \in V} F_n(y).$$

We say that F_n Γ-converge to F if and only if

$$\Gamma - \overline{\lim}_{n\to\infty} F_n(u) \le F(u) \le \Gamma - \underline{\lim}_{n\to\infty} F_n(u). \tag{A.1}$$

It is known that if there exists F that verifies (A.1) we obtain properties about minima and minimizers of function F (see [16]). As these properties are related with questions (i) and (iv′) of our problem, we are going to rewrite them in the context of Γ-convergence and ask for some assumptions on the data problem in order to obtain (A.1) for an appropriate definition of F_n and F.

We assume there exists $M_1 > 0$ such that

$$\lim_{n\to\infty} D_n = D, \quad \text{uniformly in} \quad \bigcup_{k\geq M_1} \Delta_k \cup \Delta, \tag{A.2}$$

$$\lim_{n\to\infty} C_n = C, \quad \text{uniformly in} \quad \bigcup_{k\geq M_1} \Delta_k \cup \Delta, \tag{A.3}$$

$$D : X \to \mathbb{R}^K \text{ is a lower semicontinuous and convex function,} \tag{A.4}$$

$$C : X \to \mathbb{R} \text{ is a lower semicontinuous and convex function,} \tag{A.5}$$

$$\exists M > 0 \quad \text{such that} \quad -M \leq C(x) \quad \forall x \in X, \tag{A.6}$$

$$\exists v \in U, \exists \eta > 0 \quad \text{such that} \quad D^k(v) \leq V^k - \eta \ \forall k = 1, \ldots, K, \tag{A.7}$$

$$V_n \to V, \quad n \to \infty. \tag{A.8}$$

We shall denote by (**H**) the set of hypotheses (A.2)–(A.8).

For any vector $\mathcal{V} \in \mathbb{R}^K$ and any constant $v \in \mathbb{R}$, we shall understand below $\mathcal{V} + v$ to mean the vector in $\mathbb{R}^K$ obtained by adding the constant v to each of the components of $\mathcal{V}$. We shall say that $x \in X$ is ϵ-optimal for R_n if $x \in \Delta_n$ and $C_n(x) \leq R_n + \epsilon$.

Appendix B. Key theorems for approximations. In this section we shall prove the approximation theorems.

Let us define

$$F_n(u) = \begin{cases} C_n(u) & \text{if} \quad u \in \Delta_n, \\ \infty & \text{if} \quad u \notin \Delta_n \end{cases} \tag{B.1}$$

and

$$F(u) = \begin{cases} C(u) & \text{if} \quad u \in \Delta, \\ \infty & \text{if} \quad u \notin \Delta. \end{cases} \tag{B.2}$$

We first establish the Γ-converge of F_n.

THEOREM B.1. *If* (**H**) *holds, then* F_n Γ-*converge to* F *as* $n \to \infty$.

Proof. Let x be such that $D(x) < V$ (then $F(x) = C(x)$). Let ϵ be a positive real number. Because of the lower semicontinuity (l.s.c.) of C there exists $U \in \mathcal{N}(x)$ such that

$$C(x) - \epsilon < C(y) \qquad \forall y \in U. \tag{B.3}$$

Now, given $\delta > 0$, we can take $N_1 \geq M_1$ such that if $n \geq N_1$, then, from (A.3),

$$|C_n(y) - C(y)| < \delta \qquad \text{when} \quad y \in \bigcup_{k\geq M_1} \Delta_k \cup \Delta, \tag{B.4}$$

and finally, by (A.2) and (A.8), it is possible to choose $N \geq N_1$ such that $x \in \Delta_n$ for all $n \geq N$. Then we have

$$\inf_{y \in U} F_n(y) = \inf_{y \in U \cap \Delta_n} C_n(y) \qquad \text{when} \quad n \geq N. \tag{B.5}$$

Taking into account the relations

$$U \cap \Delta_n \subset \bigcup_{k \geq M_1} \Delta_k \cup \Delta,$$

$$U \cap \Delta_n \subset U,$$

we obtain, from (B.3) and (B.4) for all $n \geq N$,

$$C(x) - \epsilon - \delta \leq \inf_{y \in U \cap \Delta_n} C(y) - \delta \leq \inf_{y \in U \cap \Delta_n} C_n(y),$$

and from the arbitrariness of δ it follows that

$$C(x) - \epsilon \leq \liminf_{n \to \infty} \inf_{y \in U} F_n. \tag{B.6}$$

On the other hand, if $V \in \mathcal{N}(x)$ and $n \geq N$, then

$$\inf_{y \in V} F_n(y) \leq C_n(x),$$

and taking into account that $C_n(x)$ converge to $C(x)$, we obtain

$$\limsup_{n \to \infty} \inf_{y \in V} F_n(x) \leq C(x) \qquad \forall V \in \mathcal{N}(x). \tag{B.7}$$

Finally, from (B.6) and (B.7) and because ϵ is also arbitrarily chosen, we deduce

$$\sup_{V \in \mathcal{N}(x)} \liminf_{n \to \infty} \inf_{y \in V} F_n(y) = C(x),$$

$$\sup_{V \in \mathcal{N}(x)} \limsup_{n \to \infty} \inf_{y \in V} F_n(y) = C(x);$$

that is,

$$(\Gamma - \lim F_n)(x) = F(x).$$

Now, we consider the second case: let x be such that $D(x) > V$ (then $F(x) = \infty$). Because of the l.s.c. of D, given $\lambda > 0$, there exists $U \in \mathcal{N}(x)$ such that if ϵ is sufficiently small, then

$$V + \epsilon < D(x) - \lambda < D(y) \qquad \forall y \in U.$$

Let $N \geq M_1$ be such that (by hypotheses (A.8) and (A.2))

$$|V - V_n| < \frac{\epsilon}{2} \qquad \text{and} \qquad |D(y) - D_n(y)| < \frac{\epsilon}{2} \qquad \forall y \in \bigcup_{k \geq M_1} \Delta_k \cup \Delta$$

whenever $n \geq N$. Therefore, if $n \geq N$ and $y \in U \cap (\bigcup_{k \geq M_1} \Delta_k \cup \Delta)$, then $D_n(y) > V_n$ and then $F_n(y) = \infty$. On the other hand, if $y \in U \setminus (\bigcup_{k \geq M_1} \Delta_k \cup \Delta)$, we also have $F_n(y) = \infty$. Thus, for all $n \geq N$

$$F_n(y) = \infty \qquad \forall y \in U,$$

and from that it follows that

$$\liminf_{n \to \infty} \inf_{y \in U} F_n(y) = \limsup_{n \to \infty} \inf_{y \in U} F_n(y) = \infty;$$

that is,

$$(\Gamma - \lim F_n)(x) = F(x).$$

Now, we consider the last case. Let x be such that $D(x) = V$. We will use the following property: If $U \in \mathcal{N}(x)$ and $\delta > 0$, there exists $y \in U$ such that

$$D(y) < V \qquad \text{and} \qquad C(y) < C(x) + \delta. \tag{B.8}$$

In fact, because of (A.7) there exists $v \in \Delta$ such that $D(v) < V$. From the convexity of C and D it follows that the segment $[v, x]$ is contained in Δ, and we have

$$D((1-t)v + tx) \leq (1-t)D(v) + tD(x) < V,$$

$$C((1-t)v + tx) \leq (1-t)C(v) + tC(x)$$

for all $t \in [0, 1]$. Then we can choose t sufficiently close to 1 and $y = (1-t)v + tx$ such that $y \in U$ and $C(y) < C(x) + \delta$.

Now, let $V \in \mathcal{N}(x)$, and let $\delta > 0$. We can choose $y \in V$ such that (B.8) holds. Then, because of (A.2) and (A.8), there exists $k \in \mathbb{N}$ such that $y \in \Delta_n$ for all $n \geq k$. Then

$$\liminf_n \inf_{x \in V} F_n(x) \leq \liminf_n C_n(y) = C(y) \leq C(x) + \delta,$$

and from the arbitrariness of δ it follows that

$$\liminf_n \inf_{x \in V} F_n(x) \leq C(x). \tag{B.9}$$

Let ϵ be a positive real number. Let $U \in \mathcal{N}(x)$ be such that (because of the l.s.c. of C)

$$C(x) - \frac{\epsilon}{2} < C(y) \qquad \forall y \in U, \tag{B.10}$$

and let $x_0 \in U$ be such that (by the property proved above)

$$D(x_0) < V.$$

Then x_0 satisfies $C(x) - \epsilon < C(x_0)$. There exists (by assumptions (A.2) and (A.8)) $N_1 \geq M_1$ such that $D_n(x_0) < V_n$ for all $n \geq N_1$. Finally, taking into account (B.10) and (A.3), we can take $N \geq N_1$ such that if $n \geq N$, then

$$C(x) - \epsilon < C_n(y) \qquad \forall y \in U \cap \Big(\bigcup_{k \geq M_1} \Delta_k \cup \Delta \Big). \tag{B.11}$$

Then we have

$$C(x) - \epsilon \leq \liminf_{n\to\infty} \inf_{y\in U} F_n(y),$$

where the last inequality follows from (B.11) and the relation

$$x_0 \in U \cap \Delta_n \subset U \cap \Big(\bigcup_{k\geq M_1} \Delta_k \cup \Delta \Big) \qquad \forall n \geq N.$$

Now, because of the arbitrariness of ϵ, we have

$$\sup_{V\in\mathcal{N}(x)} \liminf_{n} \inf_{x\in V} F_n(x) \geq C(x). \tag{B.12}$$

From (B.9) it holds that for all $V \in \mathcal{N}(x)$ and (B.12) we have

$$(\Gamma\text{-}\liminf_n F_n)(x) = F(x). \qquad \square$$

The Γ-convergence of F_n to F is proved. With some additional hypotheses we can answer (i) (Theorem B.3) and (iv) (Theorem B.5). Theorem B.4 gives an answer to (iv) for $\epsilon = 0$.

DEFINITION B.2. *We say that the sequence (G_n) is equicoercive if for all $t \in \mathbb{R}$ there exists a closed countably compact subset K_t of X such that $\{G_n \leq t\} \subseteq K_t$ for every $n \in \mathbb{N}$.*

THEOREM B.3 (see Dal Maso [16]). *Suppose that (G_n) is equicoercive in X and that Γ-converge to a function G in X. Then*

$$\min_{x\in X} G(x) = \lim_{n\to\infty} \inf_{x\in X} G_n(x).$$

THEOREM B.4 (see Dal Maso [16]). *Assume that (F_h) Γ-converge to a function F in X. For every $h \in \mathbb{N}$, let x_h be a minimizer of F_h in X (or, more generally, an ϵ_h-minimizer, where (ϵ_h) is a sequence of real numbers converging to 0). If x is a cluster point of (x_h), then x is a minimizer of F in X, and*

$$F(x) = \limsup_{h\to\infty} F_h(x_h).$$

If (x_h) converges to x in X, then x is a minimizer of F in X, and

$$F(x) = \lim_{h\to\infty} F_h(x_h).$$

THEOREM B.5 (see Rockafellar–Wets [37]). *Let $X = \mathbb{R}^N$. Suppose that (f_n) Γ-converge to f with $-\infty < \inf f < \infty$. Then we have the following.*

1. *$\inf f_n \to \inf f$ if and only if there exist for every $\epsilon > 0$ a compact set $B \subset \mathbb{R}^N$ and $k \in \mathbb{N}$ such that*

$$\inf_B f_n \leq \inf f_n + \epsilon \qquad \forall n \geq k.$$

2. *For all $\epsilon \geq 0$,*

$$\limsup_n (\epsilon\text{-}\operatorname{argmin} f_n) \subset \epsilon\text{-}\operatorname{argmin} f,$$

and consequently,

$$\limsup_n (\epsilon_n\text{-}\operatorname{argmin} f_n) \subset \operatorname{argmin} f \quad \textit{whenever} \quad \epsilon_n \searrow 0.$$

3. *Under the assumption that* $\inf f_n \to \inf f$*, there exists a sequence* $\epsilon_n \searrow 0$ *such that* $\epsilon_n-\text{argmin} f_n \to \text{argmin} f$*. Conversely, if such a sequence exists, and if* $\text{argmin} f \neq \emptyset$*, then* $\inf f_n \to \inf f$*.*

COROLLARY B.6. *Suppose* (f_n) *is an equicoercive sequence of real functions in* $\mathbb{R}^n$ *that* Γ*-converges to* f*. Let* $\epsilon > 0$ *and* $\delta > 0$*, and let* $x_n \in \epsilon-\text{argmin} f_n$*. Then there exists* $N \in \mathbb{N}$ *such that*

$$B(x_n, \delta) \cap \epsilon-\text{argmin} f \neq \emptyset$$

whenever $n \geq N$*.*

Proof. If it is not true, then there exists a subsequence (x_{n_k}) such that

$$B(x_{n_k}, \delta) \cap \epsilon-\text{argmin} f = \emptyset. \tag{B.13}$$

Now, if $t = \epsilon + \sup(\inf_{x \in \mathbb{R}_n} f_n, n \in \mathbb{N})$ (it is finite), then there exists a compact set K_t such that $\{f_n \leq t\} \subseteq K_t$ for every $n \in \mathbb{N}$, and then the sequence (x_{n_k}) has a limit point x. By statement 2 of Theorem B.5, $x \in \epsilon-\text{argmin} f$. But, if k is sufficiently large, then $x \in B(x_{n_k}, \delta)$, and this contradicts (B.13). □

We cannot answer (iii), as we cannot directly construct the function of the type of π from the Γ-convergence. But we can answer the following question:

(iii″) If x_0 is ϵ-optimal for R, is it possible to find an arbitrarily small neighborhood of x_0 and an ϵ-optimal policy $x(n)$ for R_n for n sufficiently large?

THEOREM B.7. *Suppose that the sequence* (F_n) Γ*-converges to a function* F *and*

$$\min_{x \in X} F(x) = \lim_n \inf_{x \in X} F_n(x). \tag{B.14}$$

If $x_0 \in \epsilon-\text{argmin} F$*,* $U \in \mathcal{N}(x_0)$*, and* $\bar{\epsilon} > \epsilon$*, then there exists* $N \in \mathbb{N}$ *such that*

$$U \cap \bar{\epsilon}-\text{argmin} F_n \neq \emptyset \qquad \textit{when} \quad n \geq N.$$

Proof. By definition,

$$F(x_0) = \sup_{V \in \mathcal{N}(x_0)} \liminf_n \inf_{y \in V} F_n(y) = \sup_{V \in \mathcal{N}(x_0)} \limsup_n \inf_{y \in V} F_n(y).$$

Then, if $U \in \mathcal{N}(x_0)$, we have

$$\limsup_n \inf_{y \in U} F_n(y) \leq F(x_0).$$

Thus, given $\lambda > 0$, there exists N_1 such that

$$\inf_{y \in U} F_n(y) \leq F(x_0) + \lambda \qquad \text{when} \quad n \geq N_1.$$

For each $n \geq N_1$ we can choose $y_n \in U$ such that

$$F_n(y_n) \leq F(x_0) + 2\lambda \qquad \text{when} \quad n \geq N_1. \tag{B.15}$$

On the other hand, it follows from (B.14) that there exists $N \geq N_1$ such that

$$F(x_0) - \epsilon - \lambda \leq \inf_{x \in X} F_n(x) \qquad \text{when} \quad n \geq N. \tag{B.16}$$

Finally, from (B.15) and (B.16), we obtain

$$0 \leq F_n(y_n) - \inf_{x \in X} F_n(x) \leq \epsilon + 3\lambda$$

for all $n \geq N$, and then, by choosing λ sufficiently small, we have

$$y_n \in U \cap \bar{\epsilon}\text{−argmin}F_n \qquad \forall n \geq N. \qquad \square$$

To answer questions (ii) and (iii), we present the following definitions and results.

DEFINITION B.8. *For each δ such that $0 < \delta < \eta$ (where η is defined in* (A.7)*) and for each $u \in X$ we set*

$$\epsilon_\delta(u) = \min\{\lambda \ : \ \lambda \in [0,1], \text{and } \lambda D(v) + (1-\lambda)D(u) \leq V - \delta\},$$

and

$$\pi_\delta(u) = \epsilon_\delta(u)v + (1 - \epsilon_\delta(u))u. \tag{B.17}$$

Remark B.9. $\pi_\delta(u) \in \Delta$, because of the convexity of X we have $\pi_\delta(u) \in X$, and by the definition of $\epsilon_\delta(u)$ and the fact that D is a convex function (A.4), we have

$$D(\pi_\delta(u)) \leq V - \delta. \tag{B.18}$$

Remark B.10. Let δ be such that $0 < \delta < \eta$. If (A.2), (A.4), (A.7), and (A.8) hold, there exists $\overline{N}(\delta)$ such that $\pi_\delta(u) \in \Delta_n$ if $n \geq \overline{N}(\delta)$ $\forall u \in X$. In fact, by (A.2) there exists $N(\delta)$ such that if $n \geq N(\delta)$, $D_n(\pi_\delta(u)) \leq D(\pi_\delta(u)) + \delta/2 \ \ \forall u \in X$, and by (B.18)

$$D_n(\pi_\delta(u)) \leq D(\pi_\delta(u)) + \frac{\delta}{2} \leq V - \delta + \frac{\delta}{2} \leq V - \frac{\delta}{2}.$$

Then, by (A.8) we have that there exists $\hat{N}$ such that $V - \delta/2 \leq V_n$ if $n \geq \hat{N}$, so, we have $D_n(\pi_\delta(u)) \leq V_n$ for all $n \geq \overline{N}(\delta) = \max(N(\delta), \hat{N})$.

LEMMA B.11. *Let δ be such that $0 < \delta < \eta$. If* (A.2)*,* (A.7)*, and* (A.8) *hold, then*

$$\limsup_{n\to\infty} \left\{ \sup_{u\in\Delta_n} \epsilon_\delta(u) \right\} \leq \frac{\delta}{\eta}.$$

Proof. By definition of Δ_n we have that $\forall u \in \Delta_n$, $D_n(u) \leq V_n$. By (A.2) and (A.8) we have

$$\forall \hat{\epsilon} > 0 \quad \exists N(\hat{\epsilon}) \text{ such that } \quad \forall n \geq N(\hat{\epsilon}), \forall u \in \Delta_n, \quad D(u) \leq V + 2\hat{\epsilon}. \tag{B.19}$$

By (A.7) and (B.19) we have $\forall \hat{\epsilon} > 0$, $u \in \Delta_n$, $\lambda \in [0,1]$,

$$\lambda D(v) + (1-\lambda)D(u) \leq \lambda(V - \eta) + (1-\lambda)(V + 2\hat{\epsilon}). \tag{B.20}$$

But

$$\lambda(V-\eta) + (1-\lambda)(V+2\hat{\epsilon}) \leq V - \delta \iff \lambda \geq \frac{2\hat{\epsilon} + \delta}{2\hat{\epsilon} + \eta} \quad \forall \hat{\epsilon} > 0. \tag{B.21}$$

By (B.20) and (B.21) we obtain

$$\left[\frac{2\hat{\epsilon} + \delta}{2\hat{\epsilon} + \eta}, 1\right] \subseteq \{\lambda \in [0,1] : \lambda D(v) + (1-\lambda)D(u) \leq V - \delta\} \quad \forall \hat{\epsilon} > 0 \ \ \forall u \in \Delta_n,$$

and then, by definition of $\epsilon_\delta(u)$, we have

$$\epsilon_\delta(u) \le \frac{2\hat{\epsilon}+\delta}{2\hat{\epsilon}+\eta} \quad \forall u \in \Delta_n, \quad n \ge N(\hat{\epsilon}) \quad \forall \hat{\epsilon} > 0.$$

From this last inequality we deduce

$$\sup_{u\in\Delta_n} \epsilon_\delta(u) \le \frac{2\hat{\epsilon}+\delta}{2\hat{\epsilon}+\eta} \quad n \ge N(\hat{\epsilon}) \quad \forall \hat{\epsilon} > 0,$$

and this last inequality implies

$$\limsup_{n\to\infty} \left[\sup_{u\in\Delta_n} \epsilon_\delta(u) \right] \le \frac{\delta}{\eta}. \qquad \square$$

Remark B.12. Let δ be such that $\eta > \delta > 0$. In the same way of Lemma B.11, and if (A.7) holds, we can prove that

$$\sup_{u\in\Delta} \epsilon_\delta(u) \le \frac{\delta}{\eta}.$$

For the proof it is only necessary to remark that (B.20) and (B.21) become

$$\lambda D(v) + (1-\lambda)D(u) \le \lambda(V-\eta) + (1-\lambda)V,$$

$$\lambda(V-\eta) + (1-\lambda)V \le V - \delta \Longleftrightarrow \lambda \ge \frac{\delta}{\eta}.$$

LEMMA B.13. *Let δ be such that $\eta > \delta > 0$. If* (A.2), (A.5), (A.6), (A.7), *and* (A.8) *hold, then*

$$\limsup_{n\to\infty} \left[\sup_{u\in\Delta_n} \{C(\pi_\delta(u)) - C(u)\} \right] \le (C(v)+M)\frac{\delta}{\eta}.$$

Proof. $\forall u \in \Delta_n$ by (A.5) and (B.17) we obtain

$$C(\pi_\delta(u)) - C(u) \le \epsilon_\delta(u)C(v) + (1-\epsilon_\delta(u))C(u) - C(u) = \epsilon_\delta(u)[C(v) - C(u)].$$

Then by this last equation, (A.6), and Lemma B.11 we have

$$\limsup_{n\to\infty} \sup_{u\in\Delta_n} \epsilon_\delta(u)[C(v)-C(u)] \le (C(v)+M) \limsup_{n\to\infty} \sup_{u\in\Delta_n} \epsilon_\delta(u) \le (C(v)+M)\frac{\delta}{\eta}. \qquad \square$$

Remark B.14. Let δ be such that $0 < \delta < \eta$. By Remark B.12 and if (A.5) and (A.7) hold, we can obtain

$$\sup_{u\in\Delta} \{C(\pi_\delta(u)) - C(u)\} \le (C(v)+M)\frac{\delta}{\eta}.$$

THEOREM B.15. *Suppose that* (**H**) *holds and let F_n and F defined by* (B.1) *and* (B.2), *respectively. We also suppose*

$$\lim_n \inf_{x\in X} F_n(x) = \inf_{x\in X} F(x).$$

Then,

1. *If* $u \in \epsilon-\text{argmin}F$ *and* $\lambda > 0$*, there exist* $N \in \mathbb{N}$ *and* $\delta > 0$ *such that if* $n \geq N$*, then*

$$\pi_\delta(u) \in (\epsilon + \lambda)-\text{argmin}F_n. \tag{B.22}$$

2. *Let* $u_n \in \epsilon-\text{argmin}F_n$ *and* $\lambda > 0$*. There exist* $N \in \mathbb{N}$ *and* $\delta_0 > 0$ *such that if* $n \geq N$ *and* $\delta < \delta_0$*, then*

$$\pi_\delta(u_n) \in (\epsilon + \lambda)-\text{argmin}F. \tag{B.23}$$

Proof.

1. We have

$$C_n(\pi_\delta(u)) - C(u) = [C_n(\pi_\delta(u)) - C(\pi\delta(u))] + [C(\pi_\delta(u)) - C(u)]\,.$$

From the convergence of (C_n) to C, Remark B.14, and taking into account that for each $\delta > 0$ if n is sufficiently large then $\pi_\delta(u) \in \Delta$, it follows that if $\lambda > 0$ is given, we can choose $N_1 \in \mathbb{N}$ and $\delta > 0$ sufficiently small, such that if $n > N$, then each term in the right side of the last equation is less than $\frac{\lambda}{3}$, and the equation becomes

$$C_n(\pi_\delta(u)) - C(u) < \frac{2}{3}\lambda.$$

Then, keeping in mind the ϵ-optimality of u, we have

$$C_n(\pi_\delta(u)) < C(u) + \frac{2}{3}\lambda < \inf_{x\in X} F(x) + \epsilon + \frac{2}{3}\lambda < \inf_{x\in X} F_n(x) + \epsilon + \lambda$$

if $n \geq N$ for some $N \geq N_1$, and that is (B.22).

2. We have

$$C(\pi_\delta(u_n)) - C_n(u_n) = [C(\pi_\delta(u_n)) - C(u_n)] + [C(u_n) - C_n(u_n)]\,.$$

From Lemma B.13 and the convergence of (C_n) to C, if $\lambda > 0$ is given, we can choose $N_1 \in \mathbb{N}$ and $\delta > 0$ sufficiently small, such that if $n > N_1$, then each term in the right side of the last equation is less than $\frac{\lambda}{3}$, and the equation becomes

$$C(\pi_\delta(u_n)) - C_n(u_n) < \frac{2}{3}\lambda.$$

Then, keeping in mind the ϵ-optimality of u_n, we have

$$C(\pi_\delta(u_n)) < C_n(u_n) + \frac{2}{3}\lambda < \inf_{x\in X} F_n(x) + \epsilon + \frac{2}{3}\lambda < \inf_{x\in X} F(x) + \epsilon + \lambda$$

if $n \geq N$ for some $N \geq N_1$, and that is (B.23). ∎

PROPOSITION B.16. *Suppose* **(H)** *holds. We set* $K_{\epsilon,n} = \epsilon-\text{argmin}F_n$*,* $K_\epsilon = \epsilon-\text{argmin}F$*,* $R_n = \inf_{x\in X} C_n(x)$*, and* $R = \inf_{x\in X} C(x)$*. Then*

$$K_{\epsilon,n} \cap K_\epsilon \neq \emptyset.$$

Proof. Let $x \in K_{\frac{\epsilon}{4}} \subset K_\epsilon$; then $D(x) \le V$. We will consider the case $D(x) = V$. The case $D(x) < V$ is similar and simpler. Let $w \in \Delta$ be such that $D(w) < V$. Then the segment $[w, x]$ is contained in Δ. Besides,

$$C((1-t)w + tx) \le (1-t)C(w) + tC(x) < C(x) + \frac{\epsilon}{4}$$

if $t \in [T, 1]$, for some T with $0 < T < 1$. We take a such t, and we put $x_0 = (1-t)w + tx$. Then $x_0 \in K_{\frac{\epsilon}{2}} \subset K_\epsilon$, and $D(x_0) < V$, from which it easily follows that $x_0 \in \Delta_n$ if $n \ge N_1$ for some $N_1 \in \mathbb{N}$. We can take $N \ge N_1$ such that for every $n \ge N$

$$|C_n(x_0) - C(x_0)| < \frac{\epsilon}{4} \qquad \text{and} \qquad |R_n - R| < \frac{\epsilon}{4}.$$

Therefore, if $n \ge N$, we have

$$C_n(x_0) < C(x_0) + \frac{\epsilon}{4} < R + \frac{3}{4}\epsilon < R_n + \epsilon,$$

that is $x_0 \in K_{\epsilon,n}$, and then

$$K_\epsilon \cap K_{\epsilon,n} \neq \emptyset \qquad \forall n \ge N. \qquad \square$$

REFERENCES

[1] E. Altman, *Asymptotic properties of constrained Markov decision processes*, Z. Oper. Res., 37 (1993), pp. 151–170.

[2] E. Altman, *Denumerable constrained Markov decision problems and finite approximations*, Math. Oper. Res., 19 (1994), pp. 169–191.

[3] E. Altman, *Constrained Markov Decision Processes*, in Stochastic Modeling, Chapman and Hall/CRC, Boca Raton, FL, 1999.

[4] E. Altman and V. A. Gaitsgory, *Stability and singular perturbations in constrained Markov decision problems*, IEEE Trans. Automat. Control, 38 (1993), pp. 971–975.

[5] E. Altman and A. Shwartz, *Optimal priority assignment: A time sharing approach*, IEEE Trans. Automat. Control, AC-34 (1989), pp. 1089–1102.

[6] E. Altman and A. Shwartz, *Adaptive control of constrained Markov chains*, IEEE Trans. Automat. Control, 36 (1991), pp. 454–462.

[7] E. Altman and A. Shwartz, *Markov decision problems and state-action frequencies*, SIAM J. Control Optim., 29 (1991), pp. 786–809.

[8] E. Altman and A. Shwartz, *Sensitivity of constrained Markov decision problems*, Ann. Oper. Res., 32 (1991), pp. 1–22.

[9] E. Altman and A. Shwartz, *Time-sharing policies for controlled Markov chains*, Oper. Res., 41 (1993), pp. 1116–1124.

[10] E. Altman and F. Spieksma, *The linear program approach in Markov decision problems revisited*, Z. Oper. Res., 42 (1995), pp. 169–188.

[11] F. J. Beutler and K. W. Ross, *Optimal policies for controlled Markov chains with a constraint*, J. Math. Anal. Appl., 112 (1985), pp. 236–252.

[12] F. J. Beutler and K. W. Ross, *Time-average optimal constrained semi-Markov decision processes*, Adv. Appl. Probab., 18 (1986), pp. 341–359.

[13] J. R. Birge and R. J. Wets, *Designing approximatin schemes for stochastic optimization problems*, Math. Programming Stud., 27 (1986), pp. 54–102.

[14] V. S. Borkar, *A convex analytic approach to Markov decision processes*, Probab. Theory Related Fields, 78 (1988), pp. 583–602.

[15] V. S. Borkar, *Ergodic control of Markov chains with constraints—the general case*, SIAM J. Control Optim., 32 (1994), pp. 176–186.

[16] G. Dal Maso, *An Introduction to* Γ*-Convergence*, Progr. Nonlinear Differential Equations Appl. 8, Birkhäuser, Basel, Boston, 1993.

[17] G. B. Dantzig, J. Folkman, and N. Shapiro, *On the continuity of the minimum set of a continuous function*, J. Math. Anal. Appl., 17 (1967), pp. 519–548.

[18] C. Derman and R. E. Strauch, *A note on memoryless rules for controlling sequential control processes*, Ann. Math. Statist., 37 (1966), pp. 276–278.
[19] E. Dynkin and A. Yushkevich, *Controlled Markov Processes*, Springer-Verlag, Berlin, 1979.
[20] E. A. Feinberg, *Sufficient classes of strategies in discrete dynamic programming* I: *Decomposition of randomized strategies and embedded models*, SIAM Theory Probab. Appl., 31 (1986), pp. 658–668.
[21] E. A. Feinberg, *Constrained semi-Markov decision processes with average rewards*, Z. Oper. Res., 39 (1995), pp. 257–288.
[22] E. A. Feinberg and M. I. Reiman, *Optimality of randomized trunk reservation*, Probab. Engrg. Inform. Sci., 8 (1994), pp. 463–489.
[23] A. V. Fiacco, *Convergence properties of local solutions of convex optimization problems*, J. Optim. Theory Appl., 13 (1974), pp. 1–12.
[24] V. A. Gaitsgory and A. A. Pervozvanskii, *Perturbation theory for mathematical programming problems*, J. Optim. Theory Appl., 49 (1986), pp. 389–410.
[25] A. Hordijk and L. C. M. Kallenberg, *Constrained undiscounted stochastic dynamic programming*, Math. Oper. Res., 9 (1984), pp. 276–289.
[26] A. Hordijk and F. Spieksma, *Constrained admission control to a queuing system*, Adv. in Appl. Probab., 21 (1989), pp. 409–431.
[27] L. C. M. Kallenberg, *Linear Programming and Finite Markovian Control Problems*, Mathematical Centre Tracts 148, Amsterdam, 1983.
[28] P. Kanniappan and S. M. A. Sastry, *Uniform convergence of convex optimization problems*, J. Math. Anal. Appl., 96 (1974), pp. 1–12.
[29] N. V. Krylov, *The construction of an optimal strategy for a finite controlled chain*, Theory Probab. Appl., 10 (1965), pp. 45–54.
[30] A. Lazar, *Optimal flow control of a class of queuing networks in equilibrium*, IEEE Trans. Automat. Control, 28 (1983), pp. 1001–1007.
[31] M. B. Lignota and J. Morgan, *Topological existence and stability of min sup problems*, J. Math. Anal. Appl., 151 (1990), pp. 164–180.
[32] M. B. Lignota and J. Morgan, *Convergences of marginal functions with dependent constraints*, Optimization, 23 (1992), pp. 189–213.
[33] R. Lucchetti and R. J. B. Wets, *Convergence of minima of integral functionals, with applications to optimal control and stochastic optimization*, Statist. Decisions, 11 (1993), pp. 69–84.
[34] J. Morgan and R. Raucci, *Continuity properties of ϵ-solutions for generalized parametric saddle point problems and application to hierarchical games*, J. Math. Anal. Appl., 211 (1997), pp. 30–48.
[35] P. Nain and K. W. Ross, *Optimal priority assignment with hard constraint*, IEEE Trans. Automat. Control, 31 (1986), pp. 883–888.
[36] A. A. Pervozvanskii and V. A. Gaitsgory, *Theory of Suboptimal Decision: Decomposition and Aggregation*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1988.
[37] R. T. Rockafellar and R. J. B. Wets, *Variational Analysis*, Grundlehren Math. Wiss. 317, Springer-Verlag, Berlin, 1998.
[38] K. W. Ross, *Randomized and past-dependent policies for Markov decision processes with multiple constraints*, Oper. Res., 37 (1989), pp. 474–477.
[39] K. W. Ross and B. Chen, *Optimal scheduling of interactive and non interactive traffic in telecommunication systems*, IEEE Trans. Automat. Control, 33 (1988), pp. 261–267.
[40] K. Ross and R. Varadarajan, *Markov decision processes with sample path constraints: The communicating case*, Oper. Res., 37 (1989), pp. 780–790.
[41] K. Ross and R. Varadarajan, *Multichain Markov decision processes with a sample path constraint: A decomposition approach*, Math. Oper. Res., 16 (1991), pp. 195–207.
[42] M. Schal, *A selection theorem for optimization problems*, Arch. Math., 25 (1974), pp. 219–224.
[43] I. E. Schochetman, *Pointwise versions of the maximum theorem with applications to optimization*, Appl. Math. Lett., 3 (1990), pp. 89–92.
[44] I. E. Schochetman and R. L. Smith, *Convergence of selections with applications in optimization*, J. Math. Anal. Appl., 155 (1991), pp. 278–242.
[45] L. I. Sennott, *Constrained discounted Markov decision chains*, Probab. Engrg. Inform. Sci., 5 (1991), pp. 463–475.
[46] L. I. Sennott, *Constrained average cost Markov decision chains*, Probab. Engrg. Inform. Sci., 7 (1993), pp. 69–83.