TOP (2014) 22:1074–1102
DOI 10.1007/s11750-013-0313-9

ORIGINAL PAPER

Nonzero-sum constrained discrete-time Markov games: the case of unbounded costs

Wenzhao Zhang · Yonghui Huang · Xianping Guo

Received: 16 April 2013 / Accepted: 16 December 2013 / Published online: 8 February 2014

Abstract In this paper, we consider discrete-time N-person constrained stochastic games with discounted cost criteria. The state space is denumerable and the action space is a Borel set, while the cost functions are admitted to be unbounded from below and above. Under suitable conditions *weaker* than those in (Alvarez-Mena and Hernández-Lerma, Math Methods Oper Res 63:261–285, 2006) for bounded cost functions, we also show the existence of a Nash equilibrium for the constrained games by introducing two approximations. The first one, which is as in (Alvarez-Mena and Hernández-Lerma, Math Methods Oper Res 63:261–285, 2006), is to construct a sequence of finite games to approximate a (constrained) *auxiliary* game with an initial distribution that is concentrated on a finite set. However, without hypotheses of bounded costs as in (Alvarez-Mena and Hernández-Lerma, Math Methods Oper Res 63:261–285, 2006), we also establish the existence of a Nash equilibrium for the *auxiliary* game with unbounded costs by developing more shaper error bounds of the approximation. The second one, which is *new*, is to construct a sequence of the auxiliary-type games above and prove that the limit of the sequence of Nash equilibria for the auxiliary-type games is a Nash equilibrium for the original constrained games. Our results are illustrated by a controlled queueing system.

Keywords Markov games · Occupation measure · Unbounded costs · Constrained Nash equilibria

W. Zhang · Y. Huang · X. Guo (✉)
School of Mathematics and Computational Science,
Sun Yat-Sen University, Guangzhou 510275, People's Republic of China
e-mail: mcsgxp@mail.sysu.edu.cn

W. Zhang
e-mail: zhangwenzhao1987@163.com

Y. Huang
e-mail: hyongh5@mail.sysu.edu.cn

Mathematics Subject Classification (2000) 90C40 · 91A25

1 Introduction

Discrete-time nonzero-sum Markov games have applications in various areas such as telecommunication networks, economics, and others (see, Filar and Vrieze 1997; Maitra and Sudderth 1996; Nowak 2003, 2007) and have been widely studied, see, for instance, Fink (1964) and Sobel (1971) for finite states; Federgruen (1978), Küenle (1994) and Sennott (1994) for denumerable states; Nowak (2003) for a uncountable state space. A look at these references shows that most of the literatures deal with the *unconstrained* nonzero-sum Markov games. In contrast, to the best of our knowledge, there is just a handful of work for *constrained* nonzero-sum Markov games, which can be classified into two groups:

(1) The first deals with N-player constrained stochastic games where each player independently controls a Markov chain and does not know the information on the actions of other players. Altman et al (2008) prove the existence of Nash equilibria for constrained games in which all the players use average cost criteria.
(2) The second deals with N-player constrained games where all players jointly control a single Markov chain and each player has the information on the actions of other players. To the best of our knowledge, the existence of constrained Nash equilibria for this kind of constrained game with discounted and average cost criteria is first established in Altman and Shwartz (2000) for finite states and actions. More precisely, Altman and Shwartz (2000) consider constrained games in the following different cases: (a) all the players have discounted costs; (b) all the players have average costs; (c) some players have discounted costs, while others have average costs. Using an interesting finite-state approximating technique, Alvarez-Mena and Hernández-Lerma (2006) generalize the main results in Altman and Shwartz (2000) about discounted constrained games [i.e., the case (a)] to the case of a denumerable state space and Borel action sets. However, the main results in Alvarez-Mena and Hernández-Lerma (2006) are based on a special structure of transition (Assumption 3.4 in Alvarez-Mena and Hernández-Lerma 2006) and *bounded* cost functions, and the assumption of bounded costs is required in the arguments for the main results in Alvarez-Mena and Hernández-Lerma (2006).

In this paper, we also study the second group of constrained Markov games in the case (a). We consider the general case when costs are allowed to be *unbounded from above and from below*. Our motivation for studying unbounded costs comes from the fact that unbounded costs are very common in economic models, queueing systems and so on; see Bäuerle and Rieder (2011), Durán (2000), González-Trejo et al (2002), Puterman (1994), Sennott (1999) and Hernández-Lerma and Lasserre (1999). Under suitable conditions, we show the existence of a Nash equilibrium for the constrained games with unbounded costs. Moreover, we also give *weaker* conditions than Alvarez-Mena and Hernández-Lerma (2006) for the existence of Nash equilibria for constrained games with bounded costs, see Remark 8. We next briefly point out the differences between the methods in this paper and those in Altman and Shwartz

(2000) and Alvarez-Mena and Hernández-Lerma (2006) for constrained games in the case (a):

(1′) In Altman and Shwartz (2000) and Alvarez-Mena and Hernández-Lerma (2006), the common idea is to show that the original game can be approximated by some particular games in which Nash equilibria exist. More precisely, suppose that $\mathscr{G}'$ denotes the original game with finite states considered in Altman and Shwartz (2000) or the original game with denumerable states considered in Alvarez-Mena and Hernández-Lerma (2006). The proof in Altman and Shwartz (2000) and Alvarez-Mena and Hernández-Lerma (2006) can be roughly divided into two steps: In the first step, the authors construct a sequence $\{\mathscr{G}_n'\}$ of constrained games associated with the original game $\mathscr{G}'$, and show that each $\mathscr{G}_n'$ has a Nash equilibrium $\boldsymbol{\varphi}_n$. In the second step, the authors prove that $\mathscr{G}'$ can be approximated by the sequence $\{\mathscr{G}_n'\}$, which means that the limit of the sequence $\{\boldsymbol{\varphi}_n\}$ (in the sense of Definition 3(d)) is a Nash equilibrium of $\mathscr{G}'$. In Altman and Shwartz (2000), each $\mathscr{G}_n'$ has the same state space as that of $\mathscr{G}'$. However, in Alvarez-Mena and Hernández-Lerma (2006), the state space of each $\mathscr{G}_n'$ which is the so-called *finite game* is a finite subset of that of $\mathscr{G}'$.

(2′) Compared to Alvarez-Mena and Hernández-Lerma (2006), instead of using the *finite games* to approximate the denumerable state game directly, we introduce two approximations. The first one, which is as in Alvarez-Mena and Hernández-Lerma (2006), is to construct a sequence of finite games to approximate a (constrained) *auxiliary* game with the distribution of initial state that is concentrated on a finite set. To do this, using a different estimation technique from the one in Alvarez-Mena and Hernández-Lerma (2006) for the case of bounded costs, we can obtain more shaper error bounds of the approximation; see Remark 5. Hence, we also show the existence of a Nash equilibrium for the *auxiliary* game with unbounded costs. The second one, which is *new*, is to construct a sequence of the auxiliary-type games above and prove that the original constrained game can be approximated by these auxiliary-type games. It should be mentioned that the two approximating sequences introduced in this paper are both different from the sequence $\{\mathscr{G}_n'\}$ in Altman and Shwartz (2000).

The rest of this paper is organized as follows. In Sect. 2, we introduce the models of constrained discrete-time Markov games. Then, in Sect. 3, we show the existence of constrained Nash equilibria. Finally, we illustrate our main assumptions and results by a controlled queueing system in Sect. 4.

2 The game model

In this section, we introduce the discrete-time N-person nonzero-sum constrained stochastic game model ($\mathscr{G}$) we are concerned with. First, we will use the following notation. If X is a Borel space, we denote by $\mathscr{B}(X)$ its Borel σ-algebra, by D^c the complement of a set $D \subseteq X$ (with respect to X), and by $\mathscr{P}(X)$ the set of probability measures on $\mathscr{B}(X)$ endowed with the topology of weak convergence. Let $I := \{1, \ldots, N\}$, $S := \{1, \ldots, p\}$, $\mathbb{N} := \{1, 2, \ldots\}$ and $\overline{\mathbb{N}} := \mathbb{N} \bigcup \{+\infty\}$, where p is some positive integer. For each $x, y \in \mathbb{N}$, we define $x \vee y := \max\{x, y\}$.

Now, we give the model of discrete-time N-person nonzero-sum constrained stochastic game

$$\mathscr{G} := \Big\{X, \big(A_i, \{A_i(x)|x \in X\}, c_i^0(x, \boldsymbol{a}), \{(c_i^k(x, \boldsymbol{a}), d_i^k)|k \in S\}\big)_{i\in I}, Q(\cdot|x, \boldsymbol{a}), \nu\Big\}, \tag{1}$$

with the following components:

(a) A denumerable set X denotes the state space endowed with the discrete topology. Since the set X can be suitably enumerated in the form $\{0, 1, 2, \ldots\}$, we will assume that $X = \{0, 1, \ldots\}$, the set of nonnegative integers. (See Example 1 in Sect. 4 for a bi-dimensional state space that can be suitably enumerated in the form $\{0, 1, 2, \ldots\}$.)

(b) For each $i \in I$, A_i is the action space for player i which is assumed to be a Borel space endowed with the Borel σ-algebra $\mathscr{B}(A_i)$. The set $A_i(x) \in \mathscr{B}(A_i)$ represents the nonempty set of available actions or decisions for player i at state $x \in X$. Let

$$K_i := \big\{(x, a_i)|x \in X, a_i \in A_i(x)\big\} \tag{2}$$

be the set of all feasible state-action pairs for player i. Denote by

$$K := \big\{(x, \boldsymbol{a})|x \in X, \boldsymbol{a} \in \boldsymbol{A}(x)\big\} \text{ where } \boldsymbol{A}(x) := \times_{i=1}^N A_i(x), \tag{3}$$

the set of all feasible state-action vectors for $\mathscr{G}$ which is a Borel subset of $X \times \boldsymbol{A}$ with $\boldsymbol{A} := \times_{i=1}^N A_i$.

(c) For each $i \in I$, the function $c_i^0(x, \boldsymbol{a})$ on K corresponds to the cost function that is to be minimized by player i, and $c_i^k(x, \boldsymbol{a})$ $(k \in S)$ on K corresponds to the cost functions on which some constraints are imposed. The real number d_i^k denotes the constraints for each player i and $k \in S$.

(d) Q denotes the transition function, that is, $Q(y|x, \boldsymbol{a})$ is the probability of moving from state x to y, if the action vector $\boldsymbol{a} \in \boldsymbol{A}(x)$ is chosen by the players.

(e) ν is the initial state distribution.

To precisely define the optimality criterion, we need to introduce the concept of a policy.

Definition 1 A *randomized history-dependent policy* for player i is a sequence $\pi^i := \big\{\pi_t^i, t = 0, 1, 2, \ldots\big\}$ of stochastic kernels π_t^i on the action space A_i given H_t satisfying

$$\pi_t^i(A_i(x_t)|h_t) = 1 \qquad \forall h_t := (x_0, \boldsymbol{a}_0, x_1, \boldsymbol{a}_1, \ldots, x_t) \in H_t, \qquad t = 0, 1, 2, \ldots,$$

where $H_0 := X$, $H_t := K^t \times X$ and $\boldsymbol{a}_t = (a_{1,t}, \ldots, a_{N,t}) \in \boldsymbol{A}(x_t)$.

The set of all randomized history-dependent policies for player i is denoted by Π_h^i. A multi-strategy is a vector $\boldsymbol{\pi} := (\pi^1, \ldots, \pi^N) \in \boldsymbol{\Pi}_h$, where $\boldsymbol{\Pi}_h := \times_{i=1}^N \Pi_h^i$. For each $i \in I$, let Φ^i be the set of all stochastic kernels on A_i given X.

Definition 2 (i) A randomized history-dependent policy $\pi^i = \{\pi_t^i\} \in \Pi_h^i$ for player i is said to be *randomized Markov* if there is a sequence $\{\varphi_t^i\} \subseteq \Phi^i$ such that $\pi_t^i(\cdot|h_t) = \varphi_t^i(\cdot|x_t)$ for each $h_t \in H_t$.

(ii) A randomized history-dependent policy $\pi^i = \{\pi_t^i\} \in \Pi_h^i$ for player i is said to be *randomized stationary* if there is a stochastic kernel $\varphi^i \in \Phi^i$ such that $\pi_t^i(\cdot|h_t) = \varphi^i(\cdot|x_t)$ for each $h_t \in H_t$. We will write such a stationary policy as φ^i which is obviously randomized Markovian.

The family of all randomized Markov (resp. stationary) policies for player i is denoted by Π_m^i (resp. Π_s^i), and $\boldsymbol{\Pi}_m := \times_{i=1}^N \Pi_m^i$ (resp. $\boldsymbol{\Pi}_s := \times_{i=1}^N \Pi_s^i$) denotes the set of all randomized Markov (resp. stationary) multi-strategies.

Let $\Omega := (X \times \boldsymbol{A})^\infty$, and $\mathscr{F}$ be the corresponding product σ-algebra. Then, for each $\boldsymbol{\pi} \in \boldsymbol{\Pi}_h$ and each initial distribution $\nu \in \mathscr{P}(X)$, the well-known Tulcea's Theorem (Hernández-Lerma and Lasserre, 1996, p.178) gives the existence of a unique probability measure $P_\nu^{\boldsymbol{\pi}}$ on $(\Omega, \mathscr{F})$ such that, for each $B \in \mathscr{B}(X)$ and $h_t \in H_t$,

$$P_\nu^{\boldsymbol{\pi}}(x_{t+1} \in B|h_t, \boldsymbol{a}_t) = Q(B|x_t, \boldsymbol{a}_t), \quad t = 0, 1, 2, \ldots, \tag{4}$$

where x_t and $\boldsymbol{a}_t$ denote the state and the action-vector at the decision epoch t, respectively. The expectation operator with respect to $P_\nu^{\boldsymbol{\pi}}$ is denoted by $E_\nu^{\boldsymbol{\pi}}$. If ν is concentrated at some state x, we will write $P_\nu^{\boldsymbol{\pi}}$ and $E_\nu^{\boldsymbol{\pi}}$ as $P_x^{\boldsymbol{\pi}}$ and $E_x^{\boldsymbol{\pi}}$, respectively.

For each multi-strategy $\boldsymbol{\pi} = (\pi^1, \ldots, \pi^N) \in \boldsymbol{\Pi}_h$ and policy $\pi' \in \Pi_h^i$, we define $\boldsymbol{\pi}^{-i} := (\pi^j; j \in I, j \neq i)$ to be the $N-1$ dimensional multi-strategy; moreover, we denote by $[\boldsymbol{\pi}^{-i}, \pi']$ the multi-strategy where player j uses π^j for each $j \neq i$, while player i uses π'. Similarly, for each $a_i \in A_i$, we denote by $[\boldsymbol{\pi}^{-i}, a_i]$ the N-vector in which the jth component is π^j for each $j \neq i$, while the ith component is a_i. Fix any discounted factor $\alpha \in (0, 1)$. For each $\boldsymbol{\pi} \in \boldsymbol{\Pi}_h$ and each initial state $x \in X$, the *expected discounted cost* is defined for each player i as

$$V_i^k(x, \boldsymbol{\pi}) := (1-\alpha) E_x^{\boldsymbol{\pi}} \left[\sum_{t=0}^\infty \alpha^t c_i^k(x_t, \boldsymbol{a}_t) \right] \text{ for each } \quad 0 \le k \le p. \tag{5}$$

If the initial distribution is ν, then we define

$$V_i^k(\nu, \boldsymbol{\pi}) := (1-\alpha) E_\nu^{\boldsymbol{\pi}} \left[\sum_{t=0}^\infty \alpha^t c_i^k(x_t, \boldsymbol{a}_t) \right] \text{ for each } \quad 0 \le k \le p. \tag{6}$$

Definition 3 (a) For a fixed $\boldsymbol{\pi} = (\pi^1, \ldots \pi^N) \in \boldsymbol{\Pi}_h$ and $i \in I$, $\pi' \in \Pi_h^i$ is called a *feasible policy* for player i against $\boldsymbol{\pi}$ if $V_i^k(\nu, [\boldsymbol{\pi}^{-i}, \pi']) \le d_i^k$ for each $k \in S$. We denote by

$$U_i(\boldsymbol{\pi}) := \left\{ \pi' \in \Pi_h^i \middle| V_i^k(\nu, [\boldsymbol{\pi}^{-i}, \pi']) \le d_i^k, \text{ for each } k \in S \right\}, \tag{7}$$

the set of *feasible policies* of player i against $\boldsymbol{\pi}$.

(b) A multi-strategy $\boldsymbol{\pi} = (\pi^1, \ldots, \pi^N) \in \boldsymbol{\Pi}_h$ is said to be *feasible* for $\mathscr{G}$ if $\pi^i \in U_i(\boldsymbol{\pi})$ for each $i \in I$. We denote by U the set of all *feasible multi-strategies*.

(c) (Constrained Nash equilibrium) $\boldsymbol{\pi}^* \in \boldsymbol{\Pi}_h$ is called a constrained Nash equilibrium of $\mathscr{G}$ if

$$\boldsymbol{\pi}^* \in U, \text{ and} V_i^0(\nu, \boldsymbol{\pi}^*) = \inf_{\pi' \in U_i(\boldsymbol{\pi}^*)} V_i^0(\nu, [\boldsymbol{\pi}^{*-i}, \pi']) \text{ for each } i \in I. \tag{8}$$

(d) For each $i \in I$, a sequence $\{\varphi_n\} \subseteq \Pi_s^i$ is said to converge weakly to $\varphi \in \Pi_s^i$, if $\varphi_n(\cdot|x) \to \varphi(\cdot|x)$ weakly in $\mathscr{P}(A_i(x))$ for each $x \in X$. A sequence $\{\boldsymbol{\varphi}_n\} \subseteq \boldsymbol{\Pi}_s$ with $\boldsymbol{\varphi}_n = (\varphi_n^1, \ldots, \varphi_n^N)$ for each $n \in \mathbb{N}$ is said to converge weakly to $\boldsymbol{\varphi} = (\varphi^1, \ldots, \varphi^N) \in \boldsymbol{\Pi}_s$ if the sequence $\{\varphi_n^i\}$ converges weakly to φ^i for each $i \in I$.

Our main goal in this paper is to give conditions to ensure the existence of constrained Nash equilibria.

3 Existence of constrained Nash equilibria

In this section, we show the existence of constrained Nash equilibria for game $\mathscr{G}$. In order to obtain the existence result, we need the following assumptions:

Assumption 1 Suppose that there exists a nondecreasing *moment* function $\omega \geq 1$ on X and constants $M > 0$ and $\beta > 0$ with $1 \leq \beta^2 < \frac{1}{\alpha}$ such that

(a) For each $i \in I$, $A_i(x)$ is a compact set for each $x \in X$;
(b) The functions $Q(y|x, \boldsymbol{a})$ and $c_i^k(x, \boldsymbol{a})$ are all continuous in $\boldsymbol{a} \in \boldsymbol{A}(x)$, for each $x, y \in X$, $i \in I$ and $0 \leq k \leq p$;
(c) $\sum_{y \in X} Q(y|x, \boldsymbol{a})\omega(y)$ is continuous in $\boldsymbol{a} \in \boldsymbol{A}(x)$ for each $x \in X$;
(d) $|c_i^k(x, \boldsymbol{a})| \leq M\omega(x)$ for each $i \in I$, $(x, \boldsymbol{a}) \in K$ and $0 \leq k \leq p$;
(e) $\sum_{y \in X} Q(y|x, \boldsymbol{a})\omega^2(y) \leq \beta^2\omega^2(x)$ for each $(x, \boldsymbol{a}) \in K$;
(f) $\sum_{x \in X} \omega^2(x)\nu(x) < \infty$.

Remark 1 (a) Assumptions 1(a–c) are standard continuity-compactness conditions which are widely used; see, for instance, Bäuerle and Rieder (2011), González-Trejo et al (2002), Guo and Yang (2008), Puterman (1994), Hernández-Lerma and Lasserre (1999), Nowak (1999) and their extensive references. By Assumption 1(a), for each $i \in I$ and $x \in X$, we know that the space $\mathscr{P}(A_i(x))$ and $\mathscr{P}(\boldsymbol{A}(x))$ with the topology of weak convergence are also compact. Hence, by the Tychonoff's theorem, Π_s^i and $\boldsymbol{\Pi}_s$ are compact.

(b) Assumptions 1(d)–(e) are used to guarantee the finiteness of the expected discounted cost which have been widely used in Bäuerle and Rieder (2011), Hernández-Lerma and Lasserre (1999) and Puterman (1994) for discrete-time Markov decision processes (DTMDPs); Guo and Yang (2008) and González-Trejo et al (2002) for discrete-time Markov games and the references therein. ω in Assumption 1 is a *moment* function means that there exists an increasing sequence of finite sets $Z_n \uparrow X$ such that $\lim \inf_{x \notin Z_n} \omega(x) = +\infty$, where $\inf \emptyset := +\infty$.

(c) Assumption 1(f) is a condition on the "tail" of the initial distribution and typical for constrained problem; see, for instance, Altman (1999), Hernández-Lerma and Lasserre (1999), Guo and Hernández-Lerma (2009) and their extensive reference.

(d) The functions $\omega(x)$ and $c_i^k(x, \boldsymbol{a})$ in Assumption 1 can be indeed unbounded. For example, if let $\omega(x) := x + 1$ for each $x \in X$, then the function cost $c_i^k(x, \boldsymbol{a}) := x + 1$ satisfies Assumption 1(d) which is obvious unbounded. In Example 1 for queueing system in Sect. 4, the holding cost functions c_i^0 $(i = 1, 2)$ which satisfy Assumption 1(d) are indeed unbounded. Hence, our formulation is more general than that in Alvarez-Mena and Hernández-Lerma (2006).

(e) A key role of Assumption 1(a–c,e) is to get a more general property as Assumption 3.4 in Alvarez-Mena and Hernández-Lerma (2006), see Lemma 2 and Remark 4.

Assumption 2 The $\mathscr{G}$ satisfies the slater condition, which means that, for each multi-strategy $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$ and each player i, there exists $\pi \in \Pi_h^i$ such that

$$V_i^k(\nu, [\boldsymbol{\varphi}^{-i}, \pi]) < d_i^k, \quad \text{for each } k \in S.$$

Remark 2 Assumption 2 has been used in Altman and Shwartz (2000) and Alvarez-Mena and Hernández-Lerma (2006) for the existence of constrained Nash equilibria.

3.1 Construction of auxiliary game and finite game

In this subsection, we introduce an auxiliary game and the corresponding finite games before studying the original game $\mathscr{G}$.

First, let $X_n := \{0, 1, \ldots, n\}$ for each $n \geq 1$ and $m_0 \geq 1$ be an arbitrary fixed integer. We introduce an auxiliary game $\widetilde{\mathscr{G}}_\infty(m_0)$ associated with m_0 by:

$$\Big\{X, \big(A_i, \{A_i(x)|x \in X\}, \widetilde{c}_{i,\infty}^0(x, \boldsymbol{a}), \{(\widetilde{c}_{i,\infty}^k(x, \boldsymbol{a}), \widetilde{d}_{i,\infty}^{k,m_0})|k \in S\}\big)_{i \in I}, \widetilde{Q}_\infty(\cdot|x, \boldsymbol{a}), \widetilde{\nu}_\infty^{m_0}\Big\}, \tag{9}$$

where

$$\widetilde{c}_{i,\infty}^k(x, \boldsymbol{a}) := c_i^k(x, \boldsymbol{a}) \text{ and } \widetilde{Q}_\infty(\cdot|x, \boldsymbol{a}) := Q(\cdot|x, \boldsymbol{a}),$$
$$\text{for each } i \in I, 0 \leq k \leq p, (x, \boldsymbol{a}) \in K, \tag{10}$$
$$\widetilde{\nu}_\infty^{m_0}(x) := \begin{cases} \nu(x) & \text{if } x \in X_{m_0} \backslash \{m_0\}, \\ \nu(m_0) + \nu(X_{m_0}^c) & \text{if } x = m_0, \\ 0 & \text{otherwise}, \end{cases} \tag{11}$$

and the exact definition of $\widetilde{d}_{i,\infty}^{k,m_0}$ will be given in (21) below and the other components are the same as in (1). By (21) below, we can see that $\widetilde{d}_{i,\infty}^{k,m_0}$ $(i \in I$ and $k \in S)$ depends on m_0. Hence, we add the superscript m_0 in $\widetilde{\nu}_\infty^{m_0}(x)$ and $\widetilde{d}_{i,\infty}^{k,m_0}$.

Next, we will show that $\widetilde{\mathscr{G}}_\infty(m_0)$ has a stationary Nash equilibrium. To do so, we define a sequence $\{\widetilde{Q}_n\}$ of the transition function for each $n \in \mathbb{N}$ by: for each $(x, \boldsymbol{a}) \in K$,

$$\widetilde{Q}_n(y|x,\boldsymbol{a}) := \begin{cases} Q(y|x,\boldsymbol{a}) & \text{if } y \in X_n\backslash\{n\}, \\ Q(n|x,\boldsymbol{a}) + Q(X_n^c|x,\boldsymbol{a}) & \text{if } y = n, \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Furthermore, for each $n \in \mathbb{N}$, with $\widetilde{Q}_n$ we can define a constrained game $\widetilde{\mathscr{G}}_n(m_0)$ associated with m_0 by:

$$\left\{X, \left(A_i, \{A_i(x)|x \in X\}, \widetilde{c}^0_{i,n}(x,\boldsymbol{a}), \{(\widetilde{c}^k_{i,n}(x,\boldsymbol{a}), \widetilde{d}^{k,m_0}_{i,n})|k \in S\}\right)_{i\in I}, \widetilde{Q}_n(\cdot|x,\boldsymbol{a}), \widetilde{\nu}^{m_0}_n\right\}, \tag{13}$$

where

$$\widetilde{c}^k_{i,n}(x,\boldsymbol{a}) := \begin{cases} c^k_i(x,\boldsymbol{a}) & \text{if} x \in X_n, \boldsymbol{a} \in \boldsymbol{A}(x), \\ 0 & \text{otherwise,} \end{cases} \quad \text{for each } i \in I, k \in S \text{ and } n \in \mathbb{N}, \tag{14}$$

$$\widetilde{\nu}^{m_0}_n(x) := \widetilde{\nu}^{m_0}_\infty(x) \quad \text{for each } x \in X. \tag{15}$$

The explicit definition of $\widetilde{d}^{k,m_0}_{i,n}$ $(n \in \mathbb{N})$ will be given in (35) below. Then, we will show that $\widetilde{\mathscr{G}}_\infty(m_0)$ can be approximated by $\widetilde{\mathscr{G}}_n(m_0)$, that is, if $\{\boldsymbol{\varphi}^*_n\} \subset \boldsymbol{\Pi}_s$ such that $\boldsymbol{\varphi}^*_n$ is a Nash equilibrium of $\widetilde{\mathscr{G}}_n(m_0)$ for each $n \in \mathbb{N}$ and $\boldsymbol{\varphi}^*_n \to \boldsymbol{\varphi}_\infty$ weakly in $\boldsymbol{\Pi}_s$, then $\boldsymbol{\varphi}_\infty$ is a constrained Nash equilibrium of $\widetilde{\mathscr{G}}_\infty(m_0)$.

Remark 3 By (11), (12) and (15), each $\widetilde{\mathscr{G}}_n(m_0)$ satisfies Assumption 3.2 in Alvarez-Mena and Hernández-Lerma (2006) with $\widetilde{\mathscr{X}} = X_{n\vee m_0}$ for each $n \in \mathbb{N}$. Therefore, each $\widetilde{\mathscr{G}}_n(m_0)$ $(n \in \mathbb{N})$ is the so-called *finite* game in Alvarez-Mena and Hernández-Lerma (2006).

For each $n \in \overline{\mathbb{N}}$, $\boldsymbol{\pi} \in \boldsymbol{\Pi}_h$ and initial state $x \in X$ (resp. initial distribution $\widetilde{\nu}^{m_0}_n$), we denote by $\widetilde{P}^{\boldsymbol{\pi}}_{x,n}$ (resp. $\widetilde{P}^{\boldsymbol{\pi}}_{\widetilde{\nu}^{m_0}_n}$) the probability measure for game $\widetilde{\mathscr{G}}_n(m_0)$, by $\widetilde{E}^{\boldsymbol{\pi}}_{x,n}$ (resp. $\widetilde{E}^{\boldsymbol{\pi}}_{\widetilde{\nu}^{m_0}_n}$) the expectation operator with respect to $\widetilde{P}^{\boldsymbol{\pi}}_{x,n}$ (resp. $\widetilde{P}^{\boldsymbol{\pi}}_{\widetilde{\nu}^{m_0}_n}$), and by $\widetilde{V}^k_{i,n}(x,\boldsymbol{\pi})$ and $\widetilde{V}^k_{i,n}(\widetilde{\nu}^{m_0}_n,\boldsymbol{\pi})$ the corresponding cost criteria for each game $\widetilde{\mathscr{G}}_n(m_0)$ as in (5) and (6).

Under Assumption 1, we can show the finiteness of the expected discounted costs.

Lemma 1 *Suppose that Assumptions 1(d–f) hold. Let $n \in \overline{\mathbb{N}}$, $i \in I$ and $0 \le k \le p$, then for each $x \in X$ and $\boldsymbol{\pi} \in \boldsymbol{\Pi}_h$,*

$$|\widetilde{V}^k_{i,n}(x,\boldsymbol{\pi})| \le M(1-\alpha)\frac{\omega(x)}{1-\alpha\beta}, \tag{16}$$

$$|\widetilde{V}^k_{i,n}(\widetilde{\nu}^{m_0}_n,\boldsymbol{\pi})| \le \frac{M(1-\alpha)}{1-\alpha\beta}\sum_{x\in X}\omega(x)\nu(x) \quad \text{and}$$

$$|V^k_i(\nu,\boldsymbol{\pi})| \le \frac{M(1-\alpha)}{1-\alpha\beta}\sum_{x\in X}\omega(x)\nu(x), \tag{17}$$

and

$$\sup_{\boldsymbol{\varphi}\in\boldsymbol{\Pi}_s} |\widetilde{V}^k_{i,\infty}(\widetilde{\nu}^{m_0}_\infty, \boldsymbol{\varphi}) - V^k_i(\nu, \boldsymbol{\varphi})| \le \frac{2(1-\alpha)M}{1-\alpha\beta} \sum_{x\in X^c_{m_0}} \omega(x)\nu(x). \tag{18}$$

Proof Since the ω is nondecreasing, by Jensen's inequality, Assumption 1(e), (10) and (12), for each $n \in \mathbb{N}$, $(x, \boldsymbol{a}) \in K$, it follows that

$$\begin{aligned}\sum_{y\in X} \widetilde{Q}_n(y|x,\boldsymbol{a})\omega(y) &= \sum_{y\in X_n} Q(y|x,\boldsymbol{a})\omega(y) + Q(X^c_n|x,\boldsymbol{a})\omega(n)\\ &\le \sum_{y\in X_n} Q(y|x,\boldsymbol{a})\omega(y) + \sum_{y\in X^c_n} Q(y|x,\boldsymbol{a})\omega(y) \le \beta\omega(x),\end{aligned}$$

which together with Assumption 1(e) implies that

$$\sum_{y\in X} \widetilde{Q}_n(y|x,\boldsymbol{a})\omega(y) \le \beta\omega(x) \text{ for each } x \in X \text{ and } n \in \overline{\mathbb{N}}. \tag{19}$$

Then, under Assumption 1(e), by (19) and a direct calculation, it follows that

$$\widetilde{E}^{\boldsymbol{\pi}}_{x,n}[\omega(x_t)] \le \beta\widetilde{E}^{\boldsymbol{\pi}}_{x,n}[\omega(x_{t-1})] \le \beta^2\widetilde{E}^{\boldsymbol{\pi}}_{x,n}[\omega(x_{t-2})] \le \cdots \le \beta^t\omega(x)$$

for each $x \in X$, $\boldsymbol{\pi} \in \boldsymbol{\Pi}_h$ and $n \in \overline{\mathbb{N}}$, which together with Assumption 1(d) implies (16). Moreover, by (11) and (15), we have that

$$\sum_{x\in X_{m_0}} \omega(x)\widetilde{\nu}^{m_0}_\infty(x) = \sum_{x\in X_{m_0}} \omega(x)\nu(x) + \omega(m_0)\nu(X^c_{m_0}) \le \sum_{x\in X} \omega(x)\nu(x) < \infty, \tag{20}$$

which together with (15)–(16) and (6) implies (17). Then, by (16), we have that

$$\begin{aligned}&\sup_{\boldsymbol{\varphi}\in\boldsymbol{\Pi}_s} |\widetilde{V}^k_{i,\infty}(\widetilde{\nu}^{m_0}_\infty, \boldsymbol{\varphi}) - V^k_i(\nu, \boldsymbol{\varphi})|\\ &\quad= \sup_{\boldsymbol{\varphi}\in\boldsymbol{\Pi}_s} \Big|\sum_{x\in X} \widetilde{V}^k_{i,\infty}(x, \boldsymbol{\varphi})\widetilde{\nu}^{m_0}_\infty(x) - \sum_{x\in X} V^k_i(x, \boldsymbol{\varphi})\nu(x)\Big|\\ &\quad\le \sup_{\boldsymbol{\varphi}\in\boldsymbol{\Pi}_s} \Big(|V^k_i(m_0, \boldsymbol{\varphi})\nu(X^c_{m_0})| + \sum_{x\in X^c_{m_0}} |V^k_i(x, \boldsymbol{\varphi})|\nu(x)\Big)\\ &\quad\le \frac{(1-\alpha)M}{1-\alpha\beta}(\omega(m_0)\nu(X^c_{m_0}) + \sum_{x\in X^c_{m_0}} \omega(x)\nu(x))\\ &\quad\le \frac{2(1-\alpha)M}{1-\alpha\beta} \sum_{x\in X^c_{m_0}} \omega(x)\nu(x),\end{aligned}$$

where the first inequality follows from the fact that $\widetilde{V}_{i,\infty}^{k}(x,\boldsymbol{\varphi}) = V_i^k(x,\boldsymbol{\varphi})$ for each $0 \le k \le p$, $x \in X$ and $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$. □

Now, by (18), we know that the following definition is well defined

$$\widetilde{d}_{i,\infty}^{k,m_0} := d_i^k + \sup_{\boldsymbol{\varphi}\in\boldsymbol{\Pi}_s} |\widetilde{V}_{i,\infty}^{k}(\widetilde{\nu}_{\infty}^{m_0},\boldsymbol{\varphi}) - V_i^k(\nu,\boldsymbol{\varphi})| \text{ for each } i \in I \quad \text{and} \quad k \in S. \tag{21}$$

Given $n \in \overline{\mathbb{N}}$, $i \in I$, multi-strategy $\boldsymbol{\varphi} = (\varphi^1,\ldots,\varphi^N) \in \boldsymbol{\Pi}_s$ and policy $\varphi' \in \Pi_s^i$, we define the following functions: for each $x, y \in X$ and $a_i \in A_i(x)$,

$$\begin{aligned}
&\widetilde{Q}_n(y|x,[\boldsymbol{\varphi}^{-i},\varphi'])\\
&:= \int_{\boldsymbol{A}(x)} \widetilde{Q}_n(y|x,\boldsymbol{a})\varphi^N(da_N|x)\cdots\varphi^{i+1}(da_{i+1}|x)\varphi'(da_i|x)\varphi^{i-1}(da_{i-1}|x)\\
&\quad\cdots\varphi^1(da_1|x), \qquad (22)\\
&\widetilde{Q}_n(y|x,[\boldsymbol{\varphi}^{-i},a_i])\\
&:= \int_{A_1(x)} \varphi^1(da_1|x)\cdots \int_{A_{i-1}(x)} \varphi^{i-1}(da_{i-1}|x) \int_{A_{i+1}(x)} \varphi^{i+1}(da_{i+1}|x)\\
&\quad\cdots \int_{A_N(x)} \widetilde{Q}_n(y|x,a_1,\ldots,a_{i-1},a_i,a_{i+1},\ldots,a_N)\varphi^N(da_N|x). \qquad (23)
\end{aligned}$$

Similarly, we can also define the functions $\widetilde{c}_{i,n}^{k}(x,[\boldsymbol{\varphi}^{-i},\varphi'])$ and $\widetilde{c}_{i,n}^{k}(x,[\boldsymbol{\varphi}^{-i},a_i])$.

Now, we will take much effort to give an exact definition of $\widetilde{d}_{i,n}^{k,m_0}$ for each $n \in \mathbb{N}$. To do so, we define

$$\delta(r,n) := \sup_{x\in X_r,\boldsymbol{a}\in\boldsymbol{A}(x)} \sum_{y\notin X_n} \omega(y)Q(y|x,\boldsymbol{a}), \tag{24}$$

which has been used in Altman (1999) and Cavazos-Cadena (1986) to study the finite-state approximation for DTMDP. For the function δ, we have the following fact.

Lemma 2 *Suppose that Assumptions 1(a–c,e) hold. Then* $\lim_{n\to\infty}\delta(r,n) = 0$ *for each* $r \in \mathbb{N}$.

Proof By Assumptions 1(b-c) and the finiteness of X_n, we see that $\sum_{y\in X_n^c}\omega(y) Q(y|x,\boldsymbol{a})$ is continuous in $\boldsymbol{a} \in \boldsymbol{A}(x)$ for each $x \in X$ and $n \in \mathbb{N}$. Then the rest proof is similar to the arguments at Page 207–208 in Altman (1999). □

Remark 4 A similar transition property as in Lemma 2 has been required in Assumption 3.4 of Alvarez-Mena and Hernández-Lerma (2006). Since $\omega(x) \ge 1$ for each $x \in X$, under Assumptions 1(a–b), it follows from Lemma 2 that

$$\lim_{n\to\infty} \sup_{x\in X_r} \sup_{\boldsymbol{a}\in\boldsymbol{A}(x)} Q(X_n^c|x,\boldsymbol{a}) = 0, \quad \text{for each} \quad r \in \mathbb{N}. \tag{25}$$

Then, it is easy to show that Assumption 3.4 in Alvarez-Mena and Hernández-Lerma (2006) is equivalent to the statement (25). Hence, Assumption 3.4 in Alvarez-Mena and Hernández-Lerma (2006) can follow from Lemma 2.

By Lemma 2, for each fixed arbitrarily small constant $\tau > 0$, the following function $e_n(\tau)$ is well-defined and finite:

$$e_0(\tau) := m_0 \quad \text{(recall that } m_0 \text{ is the fixed integer)} \tag{26}$$

and recursively for each $n = 1, 2, \ldots,$

$$e_n(\tau) := e(\tau, e_{n-1}(\tau)), \text{ where } e(\tau, r) := \min\{s \in \mathbb{N} : \delta(r, s) \le \tau\}. \tag{27}$$

We define $l(\tau) := \max\{e_n(\tau), n = 0, 1, \ldots, l\}$ $(l = 0, 1, \ldots)$ and the ω-norm $||\cdot||_\omega^C$ of a function u with respect to a set $C \subseteq X$ by : $||u||_\omega^C := \sup_{x\in C} \frac{|u(x)|}{\omega(x)}$.

In order to give the exact definition of $\widetilde{d}_{i,n}^{k,m_0}$, we need the following proposition.

Proposition 1 *Suppose that Assumptions 1(a–e) hold. Then*

$$\lim_{n\to\infty} \sup_{\boldsymbol{\varphi}\in\boldsymbol{\Pi}_s} |\widetilde{V}_{i,n}^k(\widetilde{\nu}_n^{m_0}, \boldsymbol{\varphi}) - \widetilde{V}_{i,\infty}^k(\widetilde{\nu}_\infty^{m_0}, \boldsymbol{\varphi})| = 0 \text{ for each } k \in S \quad \text{and} \quad i \in I.$$

Proof For any fixed $i \in I$, $k \in S$ and $\varepsilon > 0$, we take large enough $l > 0$ such that $\frac{2M(1-\alpha)}{1-\alpha\beta}(\alpha\beta)^l < \frac{\varepsilon}{2}$. Let $\tau = \frac{(1-\alpha\beta)^2}{4M\alpha(1-\alpha)}\varepsilon$. Then for each $n > l(\tau)$ and $r < l$, we have $e_{r+1}(\tau) \le l(\tau)$ which together with the definition of $X_{e_{r+1}(\tau)}$ and X_n implies that $X_{e_{r+1}(\tau)} \subseteq X_{l(\tau)} \subseteq X_n$ and $n \notin X_{e_{r+1}(\tau)}$. And thus, let $x \in X$ and $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$, it follows from (12) that

$$\widetilde{Q}_n(y|x, \boldsymbol{\varphi}) = \widetilde{Q}_\infty(y|x, \boldsymbol{\varphi}) = Q(y|x, \boldsymbol{\varphi}) \tag{28}$$

for each $n > l(\tau)$, and $y \in X_{e_{r+1}(\tau)}$ (where $r < l$). Hence, for each $n > l(\tau)$, $l > r$ and $x \in X_{e_r(\tau)}$, it follows from (28) and Assumption 1(e) that

$$\begin{aligned} &\frac{\alpha}{\omega(x)}\Big|\sum_{y\in X_{e_{r+1}(\tau)}} [\widetilde{Q}_n(y|x, \boldsymbol{\varphi})\widetilde{V}_{i,n}^k(y, \boldsymbol{\varphi}) - \widetilde{Q}_\infty(y|x, \boldsymbol{\varphi})\widetilde{V}_{i,\infty}^k(y, \boldsymbol{\varphi})]\Big| \\ &\le \alpha \sum_{y\in X_{e_{r+1}(\tau)}} \frac{|\widetilde{V}_{i,n}^k(y, \boldsymbol{\varphi}) - \widetilde{V}_{i,\infty}^k(y, \boldsymbol{\varphi})|}{\omega(y)} \frac{Q(y|x, \boldsymbol{\varphi})\omega(y)}{\omega(x)} \\ &\le \alpha\beta||\widetilde{V}_{i,n}^k(\cdot, \boldsymbol{\varphi}) - \widetilde{V}_{i,\infty}^k(\cdot, \boldsymbol{\varphi})||_\omega^{X_{e_{r+1}(\tau)}}, \end{aligned} \tag{29}$$

$$\begin{aligned} &\frac{\alpha}{\omega(x)}\Big|\sum_{y\in X_n\backslash X_{e_{r+1}(\tau)}} Q(y|x, \boldsymbol{\varphi})\widetilde{V}_{i,n}^k(y, \boldsymbol{\varphi})\Big| \\ &= \alpha\Big|\sum_{y\in X_n\backslash X_{e_{r+1}(\tau)}} \frac{\widetilde{V}_{i,n}^k(y, \boldsymbol{\varphi})}{\omega(y)} \frac{Q(y|x, \boldsymbol{\varphi})\omega(y)}{\omega(x)}\Big| \end{aligned}$$

$$\leq \alpha \sup_{y\in X_n\setminus X_{e_{r+1}(\tau)}} \frac{|\widetilde{V}^k_{i,n}(y,\boldsymbol{\varphi})|}{\omega(y)} \sum_{y\in X_n\setminus X_{e_{r+1}(\tau)}} \frac{Q(y|x,\boldsymbol{\varphi})\omega(y)}{\omega(x)}, \tag{30}$$

$$\frac{\alpha}{\omega(x)}|\widetilde{V}^k_{i,n}(n,\boldsymbol{\varphi})Q(X^c_n|x,\boldsymbol{\varphi})| = \alpha\Big|\sum_{y\notin X_n} \frac{\widetilde{V}^k_{i,n}(n,\boldsymbol{\varphi})}{\omega(y)}\frac{Q(y|x,\boldsymbol{\varphi})\omega(y)}{\omega(x)}\Big|$$

$$\leq \alpha \frac{|\widetilde{V}^k_{i,n}(n,\boldsymbol{\varphi})|}{\omega(n)} \sum_{y\notin X_n} \frac{Q(y|x,\boldsymbol{\varphi})\omega(y)}{\omega(x)}$$

$$\leq \alpha \sup_{y\in X_n\setminus X_{e_{r+1}(\tau)}} \frac{|\widetilde{V}^k_{i,n}(y,\boldsymbol{\varphi})|}{\omega(y)} \sum_{y\notin X_n} \frac{Q(y|x,\boldsymbol{\varphi})\omega(y)}{\omega(x)}, \tag{31}$$

$$\frac{\alpha}{\omega(x)}\Big|\sum_{y\notin X_{e_{r+1}(\tau)}} \widetilde{Q}_\infty(y|x,\boldsymbol{\varphi})\widetilde{V}^k_{i,\infty}(y,\boldsymbol{\varphi})\Big|$$

$$= \alpha\Big|\sum_{y\notin X_{e_{r+1}(\tau)}} \frac{\widetilde{V}^k_{i,\infty}(y,\boldsymbol{\varphi})}{\omega(y)}\frac{Q(y|x,\boldsymbol{\varphi})\omega(y)}{\omega(x)}\Big|$$

$$\leq \alpha \sup_{y\notin X_{e_{r+1}(\tau)}} \frac{|\widetilde{V}^k_{i,\infty}(y,\boldsymbol{\varphi})|}{\omega(y)} \sum_{y\notin X_{e_{r+1}(\tau)}} \frac{Q(y|x,\boldsymbol{\varphi})\omega(y)}{\omega(x)}, \tag{32}$$

the first and second inequality in (31) follow from the fact that the function ω is nondecreasing and $n \notin X_{e_{r+1}(\tau)}$, respectively. For each $n \in \overline{\mathbb{N}}$ and $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$, by Markov property, $\widetilde{V}^k_{i,n}(x,\boldsymbol{\varphi})$ satisfies the following equation:

$$u(x) = (1-\alpha)\widetilde{c}^k_{i,n}(x,\boldsymbol{\varphi}) + \alpha\sum_{y\in X} u(y)\widetilde{Q}_n(y|x,\boldsymbol{\varphi}) \text{ for each } x \in X. \tag{33}$$

Hence, by (16), (29)–(33), for each $n > l(\tau)$, $l > r$ and $x \in X_{e_r(\tau)}$,

$$\frac{|\widetilde{V}^k_{i,n}(x,\boldsymbol{\varphi}) - \widetilde{V}^k_{i,\infty}(x,\boldsymbol{\varphi})|}{\omega(x)}$$

$$= \frac{\alpha}{\omega(x)}\Big|\sum_{y\in X_{e_{r+1}(\tau)}} [\widetilde{Q}_n(y|x,\boldsymbol{\varphi})\widetilde{V}^k_{i,n}(y,\boldsymbol{\varphi}) - \widetilde{Q}_\infty(y|x,\boldsymbol{\varphi})\widetilde{V}^k_{i,\infty}(y,\boldsymbol{\varphi})]$$

$$+ \sum_{y\in X_n\setminus X_{e_{r+1}(\tau)}} Q(y|x,\boldsymbol{\varphi})\widetilde{V}^k_{i,n}(y,\boldsymbol{\varphi}) + \widetilde{V}^k_{i,n}(n,\boldsymbol{\varphi})Q(X^c_n|x,\boldsymbol{\varphi})$$

$$- \sum_{y\notin X_{e_{r+1}(\tau)}} \widetilde{Q}_\infty(y|x,\boldsymbol{\varphi})\widetilde{V}^k_{i,\infty}(y,\boldsymbol{\varphi})\Big|$$

$$\leq \alpha\beta||\widetilde{V}^k_{i,n}(\cdot,\boldsymbol{\varphi}) - \widetilde{V}^k_{i,\infty}(\cdot,\boldsymbol{\varphi})||^{X_{e_{r+1}(\tau)}}_\omega + \alpha \sup_{y\in X_n\setminus X_{e_{r+1}(\tau)}} \frac{|\widetilde{V}^k_{i,n}(y,\boldsymbol{\varphi})|}{\omega(y)}$$

$$
\begin{aligned}
&\times \left[\sum_{y \in X_n \setminus X_{e_{r+1}(\tau)}} \frac{Q(y|x,\boldsymbol{\varphi})\omega(y)}{\omega(x)} + \sum_{y \notin X_n} \frac{Q(y|x,\boldsymbol{\varphi})\omega(y)}{\omega(x)} \right] \\
&+\alpha \sup_{y \notin X_{e_{r+1}(\tau)}} \frac{|\widetilde{V}^k_{i,\infty}(y,\boldsymbol{\varphi})|}{\omega(y)} \sum_{y \notin X_{e_{r+1}(\tau)}} \frac{Q(y|x,\boldsymbol{\varphi})\omega(y)}{\omega(x)} \\
&\le \alpha\beta ||\widetilde{V}^k_{i,n}(\cdot,\boldsymbol{\varphi}) - \widetilde{V}^k_{i,\infty}(\cdot,\boldsymbol{\varphi})||^{X_{e_{r+1}(\tau)}}_{\omega} + \alpha \sup_{y \notin X_{e_{r+1}(\tau)}} \frac{|\widetilde{V}^k_{i,n}(y,\boldsymbol{\varphi})|}{\omega(y)} \\
&\times \sum_{y \notin X_{e_{r+1}(\tau)}} \frac{Q(y|x,\boldsymbol{\varphi})\omega(y)}{\omega(x)} \\
&+\alpha \sup_{y \notin X_{e_{r+1}(\tau)}} \frac{|\widetilde{V}^k_{i,\infty}(y,\boldsymbol{\varphi})|}{\omega(y)} \sum_{y \notin X_{e_{r+1}(\tau)}} \frac{Q(y|x,\boldsymbol{\varphi})\omega(y)}{\omega(x)} \\
&\le \alpha\beta ||\widetilde{V}^k_{i,n}(\cdot,\boldsymbol{\varphi}) - \widetilde{V}^k_{i,\infty}(\cdot,\boldsymbol{\varphi})||^{X_{e_{r+1}(\tau)}}_{\omega} + 2\alpha \frac{M(1-\alpha)}{1-\alpha\beta}\tau, \qquad (34)
\end{aligned}
$$

the last inequality uses the fact that $\omega \ge 1$.

Hence, for each $l(\tau) < n \in \mathbb{N}$, by (16) and iterating (34), it follows that

$$
\begin{aligned}
&||\widetilde{V}^k_{i,n}(\cdot,\boldsymbol{\varphi}) - \widetilde{V}^k_{i,\infty}(\cdot,\boldsymbol{\varphi})||^{X_{m_0}}_{\omega} \\
&\le (\alpha\beta)^l ||\widetilde{V}^k_{i,n}(\cdot,\boldsymbol{\varphi}) - \widetilde{V}^k_{i,\infty}(\cdot,\boldsymbol{\varphi})||^{X_{e_l(\tau)}}_{\omega} + 2M[1-(\alpha\beta)^l]\frac{\alpha(1-\alpha)}{(1-\alpha\beta)^2}\tau \\
&\le 2(\alpha\beta)^l \frac{M(1-\alpha)}{1-\alpha\beta} + 2M\frac{\alpha(1-\alpha)}{(1-\alpha\beta)^2}\tau < \varepsilon.
\end{aligned}
$$

Since $l(\tau)$ does not depend on $\boldsymbol{\varphi}$, we see that $\lim_{n\to\infty} \sup_{\boldsymbol{\varphi}\in\boldsymbol{\Pi}_s} ||\widetilde{V}^k_{i,n}(\cdot,\boldsymbol{\varphi}) - \widetilde{V}^k_{i,\infty}(\cdot,\boldsymbol{\varphi})||^{X_{m_0}}_{\omega} = 0$, for each $i \in I$ and $k \in S$. Then, by the finiteness of X_{m_0}, (15) and (20), we have that

$$
\begin{aligned}
&\lim_{n\to\infty} \sup_{\boldsymbol{\varphi}\in\boldsymbol{\Pi}_s} |\widetilde{V}^k_{i,n}(\widetilde{\nu}^{m_0}_n,\boldsymbol{\varphi}) - \widetilde{V}^k_{i,\infty}(\widetilde{\nu}^{m_0}_\infty,\boldsymbol{\varphi})| \\
&\le \lim_{n\to\infty} \sup_{\boldsymbol{\varphi}\in\boldsymbol{\Pi}_s} \sum_{x\in X_{m_0}} |\widetilde{V}^k_{i,n}(x,\boldsymbol{\varphi}) - \widetilde{V}^k_{i,\infty}(x,\boldsymbol{\varphi})|\widetilde{\nu}^{m_0}_\infty(x) \\
&\le \lim_{n\to\infty} \sup_{\boldsymbol{\varphi}\in\boldsymbol{\Pi}_s} ||\widetilde{V}^k_{i,n}(\cdot,\boldsymbol{\varphi}) - \widetilde{V}^k_{i,\infty}(\cdot,\boldsymbol{\varphi})||^{X_{m_0}}_{\omega} \sum_{x\in X_{m_0}} \omega(x)\widetilde{\nu}^{m_0}_\infty(x) = 0.
\end{aligned}
$$

□

Remark 5 The proof of Proposition 1 is inspired by the estimation technique applied in Altman (1999) for DTMDP with total cost criterion (see, Theorem 16.3 in Altman 1999). Different estimation technique also has been applied in the proof of Theorem 3.6(c) in Alvarez-Mena and Hernández-Lerma (2006). However, the hypotheses of bounded cost functions and special transition structure (Assumption 3.4 in Alvarez-Mena and Hernández-Lerma 2006) are required for the original game.

Now, by Proposition 1, for each $k \in S$ and $n \in \mathbb{N}$, we define

$$\widetilde{d}_{i,n}^{k,m_0} := \widetilde{d}_{i,\infty}^{k,m_0} + \sup_{\boldsymbol{\varphi}\in\boldsymbol{\Pi}_s} |\widetilde{V}_{i,n}^k(\widetilde{\nu}_n^{m_0}, \boldsymbol{\varphi}) - \widetilde{V}_{i,\infty}^k(\widetilde{\nu}_\infty^{m_0}, \boldsymbol{\varphi})| \to \widetilde{d}_{i,\infty}^{k,m_0} \text{ as } n \to \infty. \quad (35)$$

3.2 Existence of constrained Nash equilibria for auxiliary game

In this subsection, we show the existence of constrained Nash equilibria for auxiliary game introduced in Sect. 3.1. To achieve this purpose, we assume for a moment that $\widetilde{\mathscr{G}}_\infty(m_0)$ satisfies the slater condition with constraints $\{\widetilde{d}_{i,\infty}^{k,m_0} | k \in S, i \in I\}$ for cost criteria $\widetilde{V}_{i,\infty}^k(\widetilde{\nu}_\infty^{m_0}, \boldsymbol{\pi})$. In the proof of Theorem 1, it will be shown that this condition can be deduced from Assumption 2.

To begin with, we study the convergent property of constrained Nash equilibria for finite games, which actually is a N-dimension convergence problem. In order to convert this problem to be N 1-dimension convergence problems, we define a model $\widetilde{\mathscr{M}}_{i,n}^{\boldsymbol{\varphi}}(m_0)$ ($n \in \overline{\mathbb{N}}$, $i \in I$ and $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$) for constrained DTMDPs with respect to $\widetilde{\mathscr{G}}_n(m_0)$ by:

$$\Big\{X, (A_i, \{A_i(x) | x \in X\}), \widetilde{Q}_{i,n}^{\boldsymbol{\varphi}^{-i}}(\cdot|x, a_i), \widetilde{c}_{i,n}^{0,\boldsymbol{\varphi}^{-i}}(x, a_i), \{(\widetilde{c}_{i,n}^{k,\boldsymbol{\varphi}^{-i}}(x, a_i), \widetilde{d}_{i,n}^{k,m_0}) | k \in S\}, \widetilde{\nu}_n^{m_0}\Big\}, \quad (36)$$

where

$$\widetilde{Q}_{i,n}^{\boldsymbol{\varphi}^{-i}}(y|x, a_i) := \widetilde{Q}_n(y|x, [\boldsymbol{\varphi}^{-i}, a_i]) \text{ for each } y \in X \text{ and } (x, a_i) \in K_i, \quad (37)$$

$$\widetilde{c}_{i,n}^{k,\boldsymbol{\varphi}^{-i}}(x, a_i) := \widetilde{c}_{i,n}^k(x, [\boldsymbol{\varphi}^{-i}, a_i]) \text{ for each } 0 \le k \le p \text{ and } (x, a_i) \in K_i, \quad (38)$$

the other components are the same as in (15) and (35). For each fixed $i \in I$, let $\Omega_i := (X \times A_i)^\infty$ and $\mathscr{F}_i$ be the corresponding product σ-algebra. Let $n \in \overline{\mathbb{N}}$, $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$ and $\widetilde{Q}_{i,n}^{\boldsymbol{\varphi}^{-i}}$ be the transition law defined as in (37). For each $\pi \in \Pi_m^i$ and each initial distribution $\widetilde{\nu}_n^{m_0}$, the Tuclea's Theorem ensures the existence of a unique probability measure $\widehat{P}_{\boldsymbol{\varphi}^{-i},\widetilde{\nu}_n^{m_0}}^{\pi}$ on $(\Omega_i, \mathscr{F}_i)$ such that, for each $B \in \mathscr{B}(X)$ and $(x_t, a_{i,t}) \in K_i$,

$$\begin{aligned}\widehat{P}_{\boldsymbol{\varphi}^{-i},\widetilde{\nu}_n^{m_0}}^{\pi}(x_{t+1} \in B | x_t, a_{i,t}) &= \widetilde{Q}_{i,n}^{\boldsymbol{\varphi}^{-i}}(B | x_t, a_{i,t}) \\ &= \widetilde{Q}_n(B | x_t, [\boldsymbol{\varphi}^{-i}, a_{i,t}]),\ t = 0, 1, \ldots,\end{aligned}$$

where $(x_t, a_{i,t})$ denotes the state and the action of player i at the decision epoch t, and the subscript $\boldsymbol{\varphi}^{-i}$ in $\widehat{P}_{\boldsymbol{\varphi}^{-i},\widetilde{\nu}_n^{m_0}}^{\pi}$ is used to emphasize that the transition laws $\widetilde{Q}_{i,n}^{\boldsymbol{\varphi}^{-i}}$ to construct $\widehat{P}_{\boldsymbol{\varphi}^{-i},\widetilde{\nu}_n^{m_0}}^{\pi}$ is dependent on $\boldsymbol{\varphi}^{-i}$. The expectation operator with respect to

$\widehat{P}^{\pi}_{\boldsymbol{\varphi}^{-i},\widetilde{\nu}_n^{m_0}}$ is denoted by $\widehat{E}^{\pi}_{\boldsymbol{\varphi}^{-i},\widetilde{\nu}_n^{m_0}}$. By the construction of $\widehat{P}^{\pi}_{\boldsymbol{\varphi}^{-i},\widetilde{\nu}_n^{m_0}}$ and $\widetilde{P}^{[\boldsymbol{\varphi}^{-i},\pi]}_{\widetilde{\nu}_n^{m_0}}$, it is easy to see that $\widehat{P}^{\pi}_{\boldsymbol{\varphi}^{-i},\widetilde{\nu}_n^{m_0}}$ actually is the marginal distribution of $\widetilde{P}^{[\boldsymbol{\varphi}^{-i},\pi]}_{\widetilde{\nu}_n^{m_0}}$ on Ω_i. For each fixed $n \in \overline{\mathbb{N}}$, $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$ and $\pi' \in \Pi^i_m$, the expected discounted cost for player i is defined by

$$\widetilde{V}^{k,\boldsymbol{\varphi}}_{i,n}(\widetilde{\nu}_n^{m_0}, \pi) := (1-\alpha)\widehat{E}^{\pi}_{\boldsymbol{\varphi}^{-i},\widetilde{\nu}_n^{m_0}}\left[\sum_{t=0}^{\infty}\alpha^t \widetilde{c}^{k,\boldsymbol{\varphi}^{-i}}_{i,n}(x_t, a_{i,t})\right] \text{ for each } 0 \le k \le p. \tag{39}$$

For each $n \in \overline{\mathbb{N}}$ and $i \in I$, under Assumptions 1(d–e), by Fubini's Theorem, Proposition 9.2.2($f3$) in Hernández-Lerma and Lasserre (1999) and the construction of $\widehat{P}^{\pi}_{\boldsymbol{\varphi}^{-i},\widetilde{\nu}_n^{m_0}}$ and $\widetilde{P}^{[\boldsymbol{\varphi}^{-i},\pi]}_{\widetilde{\nu}_n^{m_0}}$, we obtain

$$\widetilde{V}^{k,\boldsymbol{\varphi}}_{i,n}(\widetilde{\nu}_n^{m_0}, \pi) = \widetilde{V}^{k}_{i,n}(\widetilde{\nu}_n^{m_0}, [\boldsymbol{\varphi}^{-i}, \pi]) \text{ for each } k \in S, n \in \overline{\mathbb{N}}, \boldsymbol{\varphi} \in \boldsymbol{\Pi}_s \text{ and } \pi \in \Pi^i_m. \tag{40}$$

For each $\widetilde{\mathscr{M}}^{\boldsymbol{\varphi}}_{i,n}(m_0)$, we then consider the following constrained optimality problem:

$$\begin{aligned} &\text{Minimize} \quad \widetilde{V}^{0,\boldsymbol{\varphi}}_{i,n}(\widetilde{\nu}_n^{m_0}, \pi) \\ &\text{subject to} \quad \widetilde{V}^{k,\boldsymbol{\varphi}}_{i,n}(\widetilde{\nu}_n^{m_0}, \pi) \le \widetilde{d}^{k,m_0}_{i,n} \text{ for each } \pi \in \Pi^i_m \text{ and } k \in S. \end{aligned} \tag{41}$$

A policy $\pi \in \Pi^i_m$ is said to be feasible for $\widetilde{\mathscr{M}}^{\boldsymbol{\varphi}}_{i,n}(m_0)$ if π satisfy (41).

Remark 6 By (40) and the similar proof of Theorem 5.5.1 in Puterman (1994), for each fixed $i \in I$, $n \in \overline{\mathbb{N}}$, $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$ and $\pi \in \Pi^i_h$, there exists a Markov policy $\pi' \in \Pi^i_m$ such that $\widetilde{V}^{k}_{i,n}(\widetilde{\nu}_n^{m_0}, [\boldsymbol{\varphi}^{-i}, \pi]) = \widetilde{V}^{k}_{i,n}(\widetilde{\nu}_n^{m_0}, [\boldsymbol{\varphi}^{-i}, \pi']) = \widetilde{V}^{k,\boldsymbol{\varphi}}_{i,n}(\widetilde{\nu}_n^{m_0}, \pi')$ for each $0 \le k \le p$. Hence, we only consider the constrained optimality problem $\widetilde{\mathscr{M}}^{\boldsymbol{\varphi}}_{i,n}(m_0)$ in the set of Markov policies.

For the convenience of arguments below, we introduce the following notations: for each $i \in I$, $n \in \overline{\mathbb{N}}$ and $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$,

$$\widetilde{Q}^{\boldsymbol{\varphi}^{-i}}_{i,n}(y|x, \varphi') := \int_{K_i} \widetilde{Q}^{\boldsymbol{\varphi}^{-i}}_{i,n}(y|x, a_i)\varphi'(da_i|x) \text{ for each } x, y \in X \text{ and } \varphi' \in \Pi^i_s, \tag{42}$$

$$\widetilde{c}^{k,\boldsymbol{\varphi}^{-i}}_{i,n}(x, \varphi') := \int_{K_i} \widetilde{c}^{k,\boldsymbol{\varphi}^{-i}}_{i,n}(x, a_i)\varphi'(da_i|x) \text{ for each } x \in X, 0 \le k \le p \text{ and } \varphi' \in \Pi^i_s. \tag{43}$$

Given $i \in I$, $n \in \overline{\mathbb{N}}$ and $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$, by Fubini's Theorem, we know that

$$\widetilde{Q}^{\boldsymbol{\varphi}^{-i}}_{i,n}(y|x, \varphi') = \widetilde{Q}_n(y|x, [\boldsymbol{\varphi}^{-i}, \varphi']) \text{ for each } x, y \in X \text{ and } \varphi' \in \Pi^i_s, \tag{44}$$

$$\widetilde{c}^{k,\boldsymbol{\varphi}^{-i}}_{i,n}(x,\varphi') = \widetilde{c}^{k}_{i,n}(x,[\boldsymbol{\varphi}^{-i},\varphi']) \text{ for each } x \in X, 0 \leq k \leq p \text{ and } \varphi' \in \Pi^i_s. \quad (45)$$

The next lemma gives an equivalent characterization of stationary constrained Nash equilibria.

Lemma 3 *Suppose that Assumptions 1(d–e) hold. For each fixed $n \in \overline{\mathbb{N}}$, a stationary multi-strategy $\boldsymbol{\varphi}^* = (\varphi^{*1},\ldots,\varphi^{*N}) \in \boldsymbol{\Pi}_s$ is a constrained Nash equilibrium of $\widetilde{\mathscr{G}}_n(m_0)$ if and only if φ^{*i} is optimal for $\widetilde{\mathscr{M}}^{\boldsymbol{\varphi}^*}_{i,n}(m_0)$ for each $i \in I$.*

Proof By the definition of constrained Nash equilibrium and Remark 6, this lemma can be easily verified. □

In order to analyze the convergent property of constrained Nash equilibria, we introduce occupation measures.

Definition 4 For each fixed $n \in \overline{\mathbb{N}}$, $i \in I$ and $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$, the occupation measure of $\pi' \in \Pi^i_m$ associated with $\widetilde{\mathscr{M}}^{\boldsymbol{\varphi}}_{i,n}(m_0)$ is a p.m. $\eta^{\pi'}_{\boldsymbol{\varphi}^{-i},n}$ on $X \times A_i$, which is defined by

$$\begin{aligned}&\eta^{\pi'}_{\boldsymbol{\varphi}^{-i},n}(x,\Gamma)\\&:= (1-\alpha)\sum_{t=0}^{\infty}\alpha^t \widehat{E}^{\pi'}_{\boldsymbol{\varphi}^{-i},\widetilde{\nu}^{m_0}_n}[I_{\{x\}\times\Gamma}(x_t,a_{i,t})] \text{ for each } x \in X, \Gamma \in \mathscr{B}(A_i),\end{aligned} \quad (46)$$

where the subscript n in $\eta^{\pi'}_{\boldsymbol{\varphi}^{-i},n}$ is used to emphasize $\eta^{\pi'}_{\boldsymbol{\varphi}^{-i},n}$ is defined for the model $\widetilde{\mathscr{G}}_n(m_0)$ and I_D in (46) dentes the indicator function of a set D.

Under Assumptions 1(d)–(e), for each $n \in \overline{\mathbb{N}}$, $i \in I$, $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$, $\pi' \in \Pi^i_m$ and $0 \leq k \leq p$,

$$\widetilde{V}^{k,\boldsymbol{\varphi}}_{i,n}(\widetilde{\nu}^{m_0}_n,\pi') = \int_{K_i}\widetilde{c}^{k,\boldsymbol{\varphi}^{-i}}_{i,n}(x,a_i)\eta^{\pi'}_{\boldsymbol{\varphi}^{-i},n}(x,da_i). \quad (47)$$

Proposition 2 *Suppose that Assumption 1(e) holds. Let $i \in I$, $n \in \overline{\mathbb{N}}$ and $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$. Then the following assertions hold:*

(i) For each $\pi' \in \Pi^i_m$, $\eta^{\pi'}_{\boldsymbol{\varphi}^{-i},n}$ satisfies the following equation and inequality

$$\hat{\eta}^{\pi'}_{\boldsymbol{\varphi}^{-i},n}(x) = (1-\alpha)\widetilde{\nu}^{m_0}_n(x) + \alpha\int_{K_i}\widetilde{Q}^{\boldsymbol{\varphi}^{-i}}_{i,n}(x|y,a_i)\eta^{\pi'}_{\boldsymbol{\varphi}^{-i},n}(y,da_i) \quad \forall x \in X, \quad (48)$$

$$\sum_{x\in X}\omega(x)\hat{\eta}^{\pi'}_{\boldsymbol{\varphi}^{-i},n}(x) \leq \frac{(1-\alpha)}{1-\alpha\beta}\sum_{x\in X_{m_0}}\omega(x)\widetilde{\nu}^{m_0}_{\infty}(x) < \infty, \quad (49)$$

where $\hat{\eta}^{\pi'}_{\boldsymbol{\varphi}^{-i},n}(x) := \eta^{\pi'}_{\boldsymbol{\varphi}^{-i},n}(x,A_i(x))$.

(ii) Conversely, if a p.m. η on K_i such that $(\eta, \boldsymbol{\varphi}^{-i})$ satisfies (48) and (49), then there exists a randomized stationary policy $\varphi' \in \Pi_s^i$ (depending on n and $\boldsymbol{\varphi}^{-i}$) such that $\eta = \eta_{\boldsymbol{\varphi}^{-i},n}^{\varphi'}$, and φ' can be obtained from the following decomposition of η:

$$\eta(x, da_i) = \hat{\eta}(x)\varphi'(da_i|x), \quad \text{where } \hat{\eta}(x) := \eta(x, A_i(x))\; x \in X. \tag{50}$$

Proof (48) follows from Remark 6.3.1 in Hernández-Lerma and Lasserre (1996). Using (19)–(20), by a direct calculation, it is easy to obtain (49). Part (ii) follows from Lemma 4.1 and Proposition 4.2 in Alvarez-Mena and Hernández-Lerma (2002). □

Remark 7 Remark 6 and Proposition 2 actually tell us that for each $n \in \overline{\mathbb{N}}$, $i \in I$ and stationary multi-strategy $\boldsymbol{\varphi}$, the constrained optimality problem of $\widetilde{\mathscr{M}}_{i,n}^{\boldsymbol{\varphi}}(m_0)$ can be restricted to Π_s^i.

Under Assumptions 1(d–e), by Proposition 3.6 in Alvarez-Mena and Hernández-Lerma (2002), for each fixed $n \in \overline{\mathbb{N}}$ and $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$, the constrained optimality problem of $\widetilde{\mathscr{M}}_{i,n}^{\boldsymbol{\varphi}}(m_0)$ is equivalent to the following linear programming ($\widetilde{\mathrm{LP}}_{i,n}^{\boldsymbol{\varphi}}(m_0)$):

$$\widetilde{\mathrm{LP}}_{i,n}^{\boldsymbol{\varphi}}(m_0) : \inf_{\eta} \int_{K_i} \widetilde{c}_{i,n}^{0,\boldsymbol{\varphi}^{-i}}(x, a_i)\eta(x, da_i) \tag{51}$$

$$\text{subject to} \begin{cases} \int_{K_i} \widetilde{c}_{i,n}^{k,\boldsymbol{\varphi}^{-i}}(x, a_i)\eta(x, da_i) \le \widetilde{d}_{i,n}^{k,m_0}, k \in S, \\ \hat{\eta}(x) = (1-\alpha)\widetilde{\nu}_n^{m_0}(x) + \alpha \int_{K_i} \widetilde{Q}_{i,n}^{\boldsymbol{\varphi}^{-i}}(x|y, a_i)\eta(y, da_i), \\ \sum_{x\in X} \omega(x)\hat{\eta}(x) < \infty \text{ and } \eta \in \mathscr{P}(K_i). \end{cases} \tag{52}$$

The p.m. η is called a feasible solution of $\widetilde{\mathrm{LP}}_{i,n}^{\boldsymbol{\varphi}}(m_0)$ if it satisfies (52). The set of all feasible solutions of $\widetilde{\mathrm{LP}}_{i,n}^{\boldsymbol{\varphi}}(m_0)$ is denoted by $\mathbf{F}_{i,n}^{\boldsymbol{\varphi}}$ for each $n \in \overline{\mathbb{N}}$. The p.m. η^* is called optimal for $\widetilde{\mathrm{LP}}_{i,n}^{\boldsymbol{\varphi}}(m_0)$ if $\eta^* \in \mathbf{F}_{i,n}^{\boldsymbol{\varphi}}$ and

$$\int_{K_i} \widetilde{c}_{i,n}^{0,\boldsymbol{\varphi}^{-i}}(x, a_i)\eta^*(x, da_i) = \inf_{\eta\in\mathbf{F}_{i,n}^{\boldsymbol{\varphi}}} \int_{K_i} \widetilde{c}_{i,n}^{0,\boldsymbol{\varphi}^{-i}}(x, a_i)\eta(x, da_i).$$

For finite games, we have the following result.

Lemma 4 *Suppose that Assumption 1 holds. Then $\widetilde{\mathscr{G}}_n(m_0)$ has a constrained Nash equilibrium $\boldsymbol{\varphi}_n^*$ for each $n \in \mathbb{N}$.*

Proof Let $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$ and $i \in I$ be fixed. Since $\widetilde{\mathscr{G}}_\infty(m_0)$ satisfies the slater condition, there exists $\pi \in \Pi_h^i$ such that $\widetilde{V}_{i,\infty}^k(\widetilde{\nu}_\infty^{m_0}, [\boldsymbol{\varphi}^{-i}, \pi]) < \widetilde{d}_{i,\infty}^{k,m_0}$ for each $k \in S$. By the similar proof of Theorem 5.5.1 in Puterman (1994), there exists a Markov policy $\pi' \in \Pi_m^i$ such that

$$\widetilde{V}_{i,\infty}^k(\widetilde{\nu}_\infty^{m_0}, [\boldsymbol{\varphi}^{-i}, \pi']) = \widetilde{V}_{i,\infty}^k(\widetilde{\nu}_\infty^{m_0}, [\boldsymbol{\varphi}^{-i}, \pi]) < \widetilde{d}_{i,\infty}^{k,m_0} \quad \text{for each} \quad k \in S. \tag{53}$$

By Proposition 2(ii), (47) and (53), there exists a stationary policy $\widetilde{\varphi} \in \Pi_s^i$ such that

$$\widetilde{V}_{i,\infty}^k(\widetilde{\nu}_\infty^{m_0}, [\boldsymbol{\varphi}^{-i}, \widetilde{\varphi}]) = \widetilde{V}_{i,\infty}^{k,\boldsymbol{\varphi}}(\widetilde{\nu}_\infty^{m_0}, \widetilde{\varphi}) = \widetilde{V}_{i,\infty}^{k,\boldsymbol{\varphi}}(\widetilde{\nu}_\infty^{m_0}, \pi') < \widetilde{d}_{i,\infty}^{k,m_0}. \tag{54}$$

Under Assumptions 1(a–e), it follows from Proposition 1, (35) and (54) that

$$\widetilde{V}_{i,n}^k(\widetilde{\nu}_n^{m_0}, [\boldsymbol{\varphi}^{-i}, \widetilde{\varphi}]) \le \widetilde{V}_{i,\infty}^k(\widetilde{\nu}_\infty^{m_0}, [\boldsymbol{\varphi}^{-i}, \widetilde{\varphi}]) + \sup_{\boldsymbol{\varphi}' \in \boldsymbol{\Pi}_s} |\widetilde{V}_{i,n}^k(\widetilde{\nu}_n^{m_0}, \boldsymbol{\varphi}') - \widetilde{V}_{i,\infty}^k(\widetilde{\nu}_\infty^{m_0}, \boldsymbol{\varphi}')| < \widetilde{d}_{i,n}^{k,m_0} \tag{55}$$

for each $k \in S$, which implies the slater condition for $\widetilde{\mathscr{G}}_n(m_0)$ for each $n \in \mathbb{N}$. Then according to Assumptions 1(a, b), (12)–(15) and (55), each $\widetilde{\mathscr{G}}_n(m_0)$ ($n \in \mathbb{N}$) satisfies the hypotheses of Theorem 3.6(b) in Alvarez-Mena and Hernández-Lerma (2006). Hence, each $\widetilde{\mathscr{G}}_n(m_0)$ ($n \in \mathbb{N}$) has a stationary constrained Nash equilibrium $\boldsymbol{\varphi}_n^*$. □

In order to obtain our main results, we need the following proposition.

Proposition 3 *Suppose that Assumption 1 holds. Let sequence $\{\boldsymbol{\varphi}_n^*\} \subseteq \boldsymbol{\Pi}_s$, such that $\boldsymbol{\varphi}_n^*$ is the constrained Nash equilibrium of $\widetilde{\mathscr{G}}_n(m_0)$ in Lemma 4 for each $n \in \mathbb{N}$. If $\boldsymbol{\varphi}_n^* \to \boldsymbol{\varphi}_\infty$ weakly, then the following statements hold for each fixed $i \in I$:*

(i) Let η_n be a p.m on K_i such that $(\eta_n, \boldsymbol{\varphi}_n^{*-i})$ satisfying (48)–(49) for each $n \in \mathbb{N}$ and $\eta_n \to \eta$ weakly in $\mathscr{P}(K_i)$, then

$$\lim_{n\to\infty} \int_{K_i} \widetilde{c}_{i,n}^{k,\boldsymbol{\varphi}_n^{*-i}}(x, a_i)\eta_n(x, da_i) = \int_{K_i} \widetilde{c}_{i,\infty}^{k,\boldsymbol{\varphi}_\infty^{-i}}(x, a_i)\eta(x, da_i) \quad \textit{for each} \quad 0 \le k \le p.$$

(ii) Let $\{\varphi_n'\} \subseteq \Pi_s^i$ such that $\varphi_n' \to \varphi_\infty'$ weakly, then $\eta_{\boldsymbol{\varphi}_n^{*-i},n}^{\varphi_n'} \to \eta_{\boldsymbol{\varphi}_\infty^{-i},\infty}^{\varphi_\infty'}$ weakly in $\mathscr{P}(K_i)$.

(iii) If $\eta_n \in \mathbf{F}_{i,n}^{\boldsymbol{\varphi}_n^*}$ for each $n \in \mathbb{N}$, then every accumulation point of $\{\eta_n\}$ in $\mathscr{P}(K_i)$ is a feasible solution of $\widetilde{\mathrm{LP}}_{i,\infty}^{\boldsymbol{\varphi}_\infty}(m_0)$.

(iv) For each $\eta \in \mathbf{F}_{i,\infty}^{\boldsymbol{\varphi}_\infty}$, there exist an integer N and $\eta_n \in \mathbf{F}_{i,n}^{\boldsymbol{\varphi}_n^*}$ for each $n \ge N$, such that $\eta_n \to \eta$ weakly in $\mathscr{P}(K_i)$ and

$$\lim_{n\to\infty} \int_{K_i} \widetilde{c}_{i,n}^{k,\boldsymbol{\varphi}_n^{*-i}}(x, a_i)\eta_n(x, da_i) = \int_{K_i} \widetilde{c}_{i,\infty}^{k,\boldsymbol{\varphi}_\infty^{-i}}(x, a_i)\eta(x, da_i) \text{ for each } 0 \le k \le p.$$

Proof (i) The proof of part (i) is similar to Lemma 4.7(a) in Alvarez-Mena and Hernández-Lerma (2002) for constrained DTMDPs. But our hypotheses here are different from the conditions in Alvarez-Mena and Hernández-Lerma (2002) which assume the transition law and initial distribution of approximating model are both absolutely continuous with respect to some fixed measure. Hence, we give our proof:

Let $0 \leq k \leq p$ be fixed. Under Assumptions 1(d–e), it follows from (49) that

$$\sup_{n \in \mathbb{N}} \int_{K_i} \widetilde{c}_{i,n}^{k,\boldsymbol{\varphi}_n^{*-i}}(x, a_i)\eta_n(x, da_i) < \infty. \tag{56}$$

Let $\{\int_{K_i} \widetilde{c}_{i,n_m}^{k,\boldsymbol{\varphi}_{n_m}^{*-i}}(x, a_i)\eta_{n_m}(x, da_i)\}$ be an arbitrary subsequence of the sequence $\{\int_{K_i} \widetilde{c}_{i,n}^{k,\boldsymbol{\varphi}_n^{*-i}}(x, a_i)\eta_n(x, da_i)\}$ converging to some constant v as $m \to \infty$. By Proposition 2(ii), there exists $\{\varphi'_{n_m}\} \subseteq \Pi_s^i$ such that $\eta_{n_m}(x, da_i) = \hat{\eta}_{n_m}(x)\varphi'_{n_m}(da_i|x)$ for each m. Under Assumption 1(a), by Remark 1(a), there exists a subsequence $\{n_{m_r}\}$ of $\{n_m\}$ such that $\varphi'_{n_{m_r}} \to \varphi' \in \Pi_s^i$ weakly as $r \to \infty$. Moreover, according to $\eta_n \to \eta$ weakly and since X is denumerable, we have that

$$\hat{\eta}_n(x) \to \hat{\eta}(x), \quad \forall x \in X. \tag{57}$$

By the uniqueness of the weak limit, it follows from Lemma 4.6(a) in Alvarez-Mena and Hernández-Lerma (2002) that $\eta(x, da_i) = \hat{\eta}(x)\varphi'(da_i|x)$. By (10) and (14), for each fixed $x \in X$, there exists larger enough N (depend on x) such that

$$\widetilde{c}_{i,n}^k(x, \boldsymbol{a}) = \widetilde{c}_{i,\infty}^k(x, \boldsymbol{a}) = c_i^k(x, \boldsymbol{a}) \text{ for each } n \geq N \text{ and } \boldsymbol{a} \in \boldsymbol{A}(x), \tag{58}$$

$$\text{and } |\widetilde{c}_{i,n}^k(x, \boldsymbol{a})| \leq M\omega(x) \text{ for each } n \in \overline{\mathbb{N}}. \tag{59}$$

By Assumption 1(b) and (58), we see that $\lim_{r\to\infty} \widetilde{c}_{i,n_{m_r}}^k(x, [\boldsymbol{\varphi}_{n_{m_r}}^{*-i}, \varphi'_{n_{m_r}}]) = \widetilde{c}_{i,\infty}^k(x, [\boldsymbol{\varphi}_\infty^{-i}, \varphi'])$ for each $x \in X$ which together with (45) implies that

$$\widetilde{c}_{i,n_{m_r}}^{k,\boldsymbol{\varphi}_{n_{m_r}}^{*-i}}(x, \varphi'_{n_{m_r}}) \to \widetilde{c}_{i,\infty}^{k,\boldsymbol{\varphi}_\infty^{-i}}(x, \varphi') \text{ for each } x \in X. \tag{60}$$

Under Assumptions 1(e,f), by (49) and (57),

$$\sum_{x \in X} \omega(x)\hat{\eta}(x) = \lim_{r\to\infty} \sum_{x \in X} \omega(x)\hat{\eta}_{n_{m_r}}(x) \leq \frac{(1-\alpha)}{1-\alpha\beta} \sum_{x \in X_{m_0}} \omega(x)\widetilde{\nu}_\infty^{m_0}(x) < \infty. \tag{61}$$

Hence, according to Proposition A.4 in Guo and Hernández-Lerma (2009), (57), (59)–(61), we can conclude that

$$\lim_{r\to\infty} \sum_{x \in X} \widetilde{c}_{i,n_{m_r}}^{k,\boldsymbol{\varphi}_{n_{m_r}}^{*-i}}(x, \varphi'_{n_{m_r}})\hat{\eta}_{n_{m_r}}(x) = \sum_{x \in X} \widetilde{c}_{i,\infty}^{k,\boldsymbol{\varphi}_\infty^{-i}}(x, \varphi')\hat{\eta}(x),$$

that is

$$\lim_{r\to\infty}\int\limits_{K_i}\widetilde{c}_{i,n_{m_r}}^{k,\boldsymbol{\varphi}_{n_{m_r}}^{*-i}}(x,a_i)\eta_{n_{m_r}}(x,da_i)=\int\limits_{K_i}\widetilde{c}_{i,\infty}^{k,\boldsymbol{\varphi}_{\infty}^{-i}}(x,a_i)\eta(x,da_i). \tag{62}$$

As the sequence $\{\int_{K_i}\widetilde{c}_{i,n_m}^{k,\boldsymbol{\varphi}_{n_m}^{*-i}}(x,a_i)\eta_{n_m}(x,da_i)\}$ was arbitrary chosen and (by 62) all such sequences have the same limit $\int_{K_i}\widetilde{c}_{i,\infty}^{k,\boldsymbol{\varphi}_{\infty}^{-i}}(x,a_i)\eta(x,da_i)$, we have that

$$\lim_{n\to\infty}\int\limits_{K_i}\widetilde{c}_{i,n}^{k,\boldsymbol{\varphi}_{n}^{*-i}}(x,a_i)\eta_n(x,da_i)=\int\limits_{K_i}\widetilde{c}_{i,\infty}^{k,\boldsymbol{\varphi}_{\infty}^{-i}}(x,a_i)\eta(x,da_i),$$

which completes the proof.

(ii) Let $[\widetilde{Q}_{i,n}^{\boldsymbol{\varphi}_n^{*-i}}(\cdot|x,\varphi_n')]^t$ denote the corresponding t-step transition laws for each $n\in\overline{\mathbb{N}}$. For each fixed $x,y\in X$, by (10) and (12), there exists large enough N (depending on x,y) such that $\widetilde{Q}_n(y|x,\boldsymbol{a})=\widetilde{Q}_\infty(y|x,\boldsymbol{a})$ for each $n\geq N$ and $\boldsymbol{a}\in\boldsymbol{A}(x)$. Hence, it follows from Assumption 1(b) and (15) that, for each $x,y\in X$ and $t\in\mathbb{N}$,

$$\begin{aligned}\lim_{n\to\infty}\widetilde{Q}_{i,n}^{\boldsymbol{\varphi}_n^{*-i}}(y|x,\varphi_n')&=\lim_{n\to\infty}\widetilde{Q}_\infty(y|x,[\boldsymbol{\varphi}_n^{*-i},\varphi_n'])\\&=\widetilde{Q}_\infty(y|x,[\boldsymbol{\varphi}_\infty^{-i},\varphi_\infty'])=\widetilde{Q}_{i,\infty}^{\boldsymbol{\varphi}_\infty^{-i}}(y|x,\varphi_\infty'),\end{aligned} \tag{63}$$

and

$$\begin{aligned}&\lim_{n\to\infty}\sum_{x\in X}\left[\widetilde{Q}_{i,n}^{\boldsymbol{\varphi}_n^{*-i}}(y|x,\varphi_n')\right]^t\widetilde{\nu}_n^{m_0}(x)\\&=\sum_{x\in X}\left[\widetilde{Q}_{i,\infty}^{\boldsymbol{\varphi}_\infty^{-i}}(y|x,\varphi_\infty')\right]^t\widetilde{\nu}_\infty^{m_0}(x),\ \text{ for each } t\in\mathbb{N}\end{aligned} \tag{64}$$

which correspond to the key equalities (4.15) and (4.17) in Alvarez-Mena and Hernández-Lerma (2002), respectively. Then, the remaining proof is exactly the same as the proof of Theorem 4.9(a) in Alvarez-Mena and Hernández-Lerma (2002).

(iii) For convenience, suppose that $\eta_n\to\eta$ weakly in $\mathscr{P}(K_i)$. Under Assumption 1(a), by Remark 1(a) and Proposition 2(ii), there exists a subsequence $\{\varphi_{n_m}'\}$ of $\{\varphi_n'\}\subseteq\Pi_s^i$, such that $\eta_n=\eta_{\boldsymbol{\varphi}_n^{*-i},n}^{\varphi_n'}$ for all $n\geq 1$, and $\varphi_{n_m}'\to\varphi_\infty'\in\Pi_s^i$ weakly. Under Assumption 1(b), part (ii) shows that $\eta_{\boldsymbol{\varphi}_{n_m}^{*-i},n_m}^{\varphi_{n_m}'}\to\eta_{\boldsymbol{\varphi}_\infty^{-i},\infty}^{\varphi_\infty'}$ weakly in $\mathscr{P}(K_i)$ as $m\to\infty$. Moreover, under Assumptions 1(a–e), by part (i) and (35), we have

$$\begin{aligned}\int_{K_i} \widetilde{c}_{i,\infty}^{k,\boldsymbol{\varphi}_\infty^{-i}}(x,a_i)\eta_{\boldsymbol{\varphi}_\infty^{-i},\infty}^{\varphi'_\infty}(x,da_i) &= \lim_{m\to\infty}\int_{K_i} \widetilde{c}_{i,n_m}^{k,\boldsymbol{\varphi}_{n_m}^{*-i}}(x,a_i)\eta_{\boldsymbol{\varphi}_{n_m}^{*-i},n_m}^{\varphi'_{n_m}}(x,da_i)\\ &\le \lim_{m\to\infty}\widetilde{d}_{i,n_m}^{k,m_0}=\widetilde{d}_{i,\infty}^{k,m_0},\end{aligned}$$

for all $k\in S$. Hence, $\eta_{\boldsymbol{\varphi}_\infty^{-i},\infty}^{\varphi'_\infty}\in\mathbf{F}_{i,\infty}^{\boldsymbol{\varphi}_\infty}$. By the uniqueness of the weak limit, we know that $\eta=\eta_{\boldsymbol{\varphi}_\infty^{-i},\infty}^{\varphi'_\infty}$, so part (iii) is true.

(iv) Let $\eta\in\mathbf{F}_{i,\infty}^{\boldsymbol{\varphi}_\infty}$. Under the slater condition for $\widetilde{\mathscr{G}}_\infty(m_0)$, there exists $\pi\in\Pi_h^i$ and some constant $C>0$ such that $\widetilde{V}_{i,\infty}^k(\widetilde{\nu}_\infty^{m_0},[\boldsymbol{\varphi}_\infty^{-i},\pi])\le\widetilde{d}_{i,\infty}^{k,m_0}-C$ for each $k\in S$. Furthermore, using Proposition 2(ii) and the same argument for (54), there exists $\varphi'_\infty,\widetilde{\varphi}_\infty\in\Pi_s^i$, such that $\eta=\eta_{\boldsymbol{\varphi}_\infty^{-i},\infty}^{\varphi'_\infty}$ and

$$\begin{aligned}\widetilde{V}_{i,\infty}^{k,\boldsymbol{\varphi}_\infty}(\widetilde{\nu}_\infty^{m_0},\widetilde{\varphi}_\infty) &= \widetilde{V}_{i,\infty}^k(\widetilde{\nu}_\infty^{m_0},[\boldsymbol{\varphi}_\infty^{-i},\widetilde{\varphi}_\infty])\\ &= \widetilde{V}_{i,\infty}^k(\widetilde{\nu}_\infty^{m_0},[\boldsymbol{\varphi}_\infty^{-i},\pi])\le\widetilde{d}_{i,\infty}^{k,m_0}-C\quad\forall k\in S.\end{aligned}\tag{65}$$

Then, under Assumptions 1(a,b,d–f), by part (ii), we have

$$\eta_{\boldsymbol{\varphi}_n^{*-i},n}^{\varphi'_\infty}\to\eta_{\boldsymbol{\varphi}_\infty^{-i},\infty}^{\varphi'_\infty}\text{ and }\eta_{\boldsymbol{\varphi}_n^{*-i},n}^{\widetilde{\varphi}_\infty}\to\eta_{\boldsymbol{\varphi}_\infty^{-i},\infty}^{\widetilde{\varphi}_\infty}\text{ weakly in }\mathscr{P}(K_i)\text{ as }n\to\infty,$$

and so (by part (i) and (65)), for all $k\in S$,

$$\begin{aligned}&\lim_{n\to\infty}\int_{K_i}\widetilde{c}_{i,n}^{k,\boldsymbol{\varphi}_n^{*-i}}(x,a_i)\eta_{\boldsymbol{\varphi}_n^{*-i},n}^{\varphi'_\infty}(x,da_i)=\int_{K_i}\widetilde{c}_{i,\infty}^{k,\boldsymbol{\varphi}_\infty^{-i}}(x,a_i)\eta_{\boldsymbol{\varphi}_\infty^{-i},\infty}^{\varphi'_\infty}(x,da_i)\le\widetilde{d}_{i,\infty}^{k,m_0},\\ &\lim_{n\to\infty}\int_{K_i}\widetilde{c}_{i,n}^{k,\boldsymbol{\varphi}_n^{*-i}}(x,a_i)\eta_{\boldsymbol{\varphi}_n^{*-i},n}^{\widetilde{\varphi}_\infty}(x,da_i)=\int_{K_i}\widetilde{c}_{i,\infty}^{k,\boldsymbol{\varphi}_\infty^{-i}}(x,a_i)\eta_{\boldsymbol{\varphi}_\infty^{-i},\infty}^{\widetilde{\varphi}_\infty}(x,da_i)\le\widetilde{d}_{i,\infty}^{k,m_0}-C,\end{aligned}$$

which deduce the inequalities (5.3) and (5.4) in Alvarez-Mena and Hernández-Lerma (2002). Then, using the equality (35), the sequences $\{\eta_{\boldsymbol{\varphi}_n^{*-i},n}^{\varphi'_\infty}\}$ and $\{\eta_{\boldsymbol{\varphi}_n^{*-i},n}^{\widetilde{\varphi}_\infty}\}$, we can construct the desired sequence $\{\eta_n\}$ by the same method as in the steps 2–3 of the proof of Lemma 5.1 in Alvarez-Mena and Hernández-Lerma (2002). □

Lemma 5 *Suppose that Assumption 1 holds. Then $\widetilde{\mathscr{G}}_\infty(m_0)$ has a constrained Nash equilibrium.*

Proof Let $\boldsymbol{\varphi}_n^*:=(\varphi_n^{*1},\ldots,\varphi_n^{*N})$ be the constrained Nash equilibrium in Lemma 4 for each $n\in\mathbb{N}$. It follows from Remark 1(a) that there exists $\boldsymbol{\varphi}_\infty\in\boldsymbol{\Pi}_s$ such that $\boldsymbol{\varphi}_n^*\to\boldsymbol{\varphi}_\infty$ weakly. Then, under Assumptions 1(d–e), it follows from Lemma 3 and Proposition 2 that $\boldsymbol{\varphi}_n^*$ is a constrained Nash equilibrium of $\widetilde{\mathscr{G}}_n(m_0)$ ($n\in\mathbb{N}$) if and only if $\eta_{\boldsymbol{\varphi}_n^{*-i},n}^{\varphi_n^{*i}}$ is optimal for $\widetilde{\mathrm{LP}}_{i,n}^{\boldsymbol{\varphi}_n^*}(m_0)$ for each $i\in I$. By Propositions 3(i,iii,iv), the sequence $\{\widetilde{\mathrm{LP}}^{\boldsymbol{\varphi}_n^*}_{i,n}(m_0)\}$ and the "limit" $\widetilde{\mathrm{LP}}^{\boldsymbol{\varphi}_\infty}_{i,\infty}(m_0)$ satisfy the Assumption 2.1 in Alvarez-Mena and Hernández-Lerma (2002) for each $i \in I$. Hence, it follows from Theorem 2.3 in Alvarez-Mena and Hernández-Lerma (2002) that every accumulation point of $\{\eta^{\varphi_n^{*i}}_{\boldsymbol{\varphi}_n^{*-i},n}\}$ is an optimal solution of $\widetilde{\mathrm{LP}}^{\boldsymbol{\varphi}_\infty}_{i,\infty}(m_0)$ for each $i \in I$. Since $\boldsymbol{\varphi}_n^* \to \boldsymbol{\varphi}_\infty$ weakly, by Proposition 3(ii), we have that $\eta^{\varphi_n^{*i}}_{\boldsymbol{\varphi}_n^{*-i},n} \to \eta^{\varphi_\infty^i}_{\boldsymbol{\varphi}_\infty^{-i},\infty}$ weakly in $\mathscr{P}(K_i)$ which implies that $\eta^{\varphi_\infty^i}_{\boldsymbol{\varphi}_\infty^{-i},\infty}$ is an optimal solution of $\widetilde{\mathrm{LP}}^{\boldsymbol{\varphi}_\infty}_{i,\infty}(m_0)$ for each $i \in I$. Thus, by Lemma 3 and Proposition 2, we can see that $\boldsymbol{\varphi}_\infty$ is a constrained Nash equilibrium of $\widetilde{\mathscr{G}}_\infty(m_0)$. □

3.3 The main result

In this subsection, we give the main result in this paper.

Theorem 1 *Suppose that Assumptions 1 and 2 hold. Then the game $\mathscr{G}$ has a constrained Nash equilibrium.*

Proof *Step 1.* Define a sequence $\{\nu_n\}$ of initial distribution by:

$$\nu_n(x) := \begin{cases} \nu(x) & \text{if } x \in X_n \backslash \{n\}, \\ \nu(n) + \nu(X_n^c) & \text{if } x = n, \\ 0 & \text{otherwise}, \end{cases} \tag{66}$$

which satisfies that

$$\lim_{n\to\infty} \nu_n(x) = \nu(x) \text{ for each } x \in X. \tag{67}$$

For each $n \in \mathbb{N}$, with ν_n, we can define an auxiliary constrained game $\widetilde{\mathscr{G}}_\infty(n)$ as in (9) by:

$$\begin{aligned}\widetilde{\mathscr{G}}_\infty(n) := \Big\{&X, \big(A_i, \{A_i(x)|x \in X\}, c_i^0(x, \boldsymbol{a}), \\ &\{(c_i^k(x, \boldsymbol{a}), d_{i,n}^k)|k \in S\}\big)_{i\in I}, Q(\cdot|x, \boldsymbol{a}), \nu_n\Big\},\end{aligned} \tag{68}$$

where for each $i \in I$ and $k \in S$,

$$d_{i,n}^k := d_i^k + \sup_{\boldsymbol{\varphi}\in\boldsymbol{\Pi}_s} |\widetilde{V}_{i,\infty}^k(\nu_n, \boldsymbol{\varphi}) - V_i^k(\nu, \boldsymbol{\varphi})| \to d_i^k \ \text{ (by (18))}, \tag{69}$$

(note that $\widetilde{V}_{i,\infty}^k(\nu_n, \boldsymbol{\varphi}) = \widetilde{V}_{i,\infty}^k(\widetilde{\nu}_\infty^n, \boldsymbol{\varphi})$ for each $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$ because $\nu_n = \widetilde{\nu}_\infty^n$) the other components are the same as in (1). (Compared (68) with (9), we only use the different notations here. This is merely to make the following arguments concise.)

Recall that we have shown that if each auxiliary game $\widetilde{\mathscr{G}}_\infty(n)$ $(n \in \mathbb{N})$ satisfies slater condition with constraints $\{d_{i,n}^k|k \in S, i \in I\}$ (it is worth bearing in mind that

$d^k_{i,n} \equiv \widetilde{d}^{k,n}_{i,\infty}$ for each $k \in S$, $i \in I$ and $n \in \mathbb{N}$), each $\widetilde{\mathscr{G}}_\infty(n)$ has a constrained Nash equilibrium by Lemma 5. Next, we will show that slater condition for each $\widetilde{\mathscr{G}}_\infty(n)$ ($n \in \mathbb{N}$) can be guaranteed by Assumption 2. To do so, let $i \in I$ and $\boldsymbol{\varphi} \in \boldsymbol{\Pi}_s$ be fixed, and π be the corresponding policy as in Assumption 2. By the similar proof of Theorem 5.5.1 in Puterman (1994), there exists a Markov policy $\pi' \in \Pi^i_m$ associated with π such that

$$V^k_i(\nu, [\boldsymbol{\varphi}^{-i}, \pi']) = V^k_i(\nu, [\boldsymbol{\varphi}^{-i}, \pi]) < d^k_i, \text{ for each } k \in S. \tag{70}$$

By introducing occupation measures as in Definition 4 for $\mathscr{G}$ and using the similar argument for (54), we can get a stationary policy $\widetilde{\varphi} \in \Pi^i_s$ such that $V^k_i(\nu, [\boldsymbol{\varphi}^{-i}, \pi]) = V^k_i(\nu, [\boldsymbol{\varphi}^{-i}, \widetilde{\varphi}])$ for each $k \in S$, which together with (69) and (70) implies that

$$\widetilde{V}^k_{i,\infty}(\nu_n, [\boldsymbol{\varphi}^{-i}, \widetilde{\varphi}]) \leq V^k_i(\nu, [\boldsymbol{\varphi}^{-i}, \widetilde{\varphi}]) + \sup_{\boldsymbol{\varphi}' \in \boldsymbol{\Pi}_s} |\widetilde{V}^k_{i,\infty}(\nu_n, \boldsymbol{\varphi}') - V^k_i(\nu, \boldsymbol{\varphi}')| < d^k_{i,n}$$

for each $k \in S$. Hence, $\widetilde{\mathscr{G}}_\infty(n)$ satisfies slater condition with constraints $\{d^k_{i,n} | k \in S, i \in I\}$ for each $n \in \mathbb{N}$ which implies that each $\widetilde{\mathscr{G}}_\infty(n)$ has a constrained Nash equilibrium $\boldsymbol{\varphi}^*_n$.

Step 2. By Remark 1(a), there exists $\boldsymbol{\varphi}_\infty \in \boldsymbol{\Pi}_s$ which is an accumulation point of $\{\boldsymbol{\varphi}^*_n\}$. Now, we will show that $\boldsymbol{\varphi}_\infty$ is a Nash equilibrium of $\mathscr{G}$. To do so, let $Q^{\boldsymbol{\varphi}^{-i}}_i(y|x, a_i)$, $c^{k,\boldsymbol{\varphi}^{-i}}_i(x, a_i)$, $Q^{\boldsymbol{\varphi}^{-i}}_i(y|x, \varphi')$ and $c^{k,\boldsymbol{\varphi}^{-i}}_i(y, \varphi')$ be the functions defined as in (37)–(38) and (42)–(43) with $c^k_i(x, \boldsymbol{a})$ and $Q(y|x, \boldsymbol{a})$ in lieu of $\widetilde{c}^k_{i,n}(x, \boldsymbol{a})$ and $\widetilde{Q}_n(y|x, \boldsymbol{a})$, respectively.

Then, as in (51)–(52), for each $i \in I$ and $n \in \mathbb{N}$, we introduce the following linear programming $\mathrm{LP}^{\boldsymbol{\varphi}^*_n}_{i,n}$ associated with $\widetilde{\mathscr{G}}_\infty(n)$:

$$\mathrm{LP}^{\boldsymbol{\varphi}^*_n}_{i,n} : \inf_\eta \int_{K_i} c^{0,\boldsymbol{\varphi}^{*-i}_n}_i(x, a_i)\eta(x, da_i) \tag{71}$$

$$\text{subject to} \begin{cases} \int_{K_i} c^{k,\boldsymbol{\varphi}^{*-i}_n}_i(x, a_i)\eta(x, da_i) \leq d^k_{i,n}, k \in S, \\ \hat{\eta}(x) = (1-\alpha)\nu_n(x) + \alpha \int_{K_i} Q^{\boldsymbol{\varphi}^{*-i}_n}_i(x|y, a_i)\eta(y, da_i), \\ \sum_{x\in X} \omega(x)\hat{\eta}(x) < \infty \text{ and } \eta \in \mathscr{P}(K_i). \end{cases} \tag{72}$$

Similarly, by replacing $\{d^k_{i,n} | k \in S, i \in I\}$, $\boldsymbol{\varphi}^*_n$ and ν_n with $\{d^k_i | k \in S, i \in I\}$, $\boldsymbol{\varphi}_\infty$ and ν, respectively, in (72), we can also define the linear programming $\mathrm{LP}^{\boldsymbol{\varphi}_\infty}_i$ associated with $\mathscr{G}$ for each $i \in I$.

Furthermore, by (66)–(68) and Assumptions 1(b,f), we have the following statements: Let $i \in I$ be fixed, if $\{\varphi'_n\} \subseteq \Pi^i_s$, such that $\varphi'_n \to \varphi'_\infty$ weakly, then

(a) $\lim_{n\to\infty} c^{k,\boldsymbol{\varphi}^{*-i}_n}_i(x, \varphi'_n) = c^{k,\boldsymbol{\varphi}^{-i}_\infty}_i(x, \varphi'_\infty)$ for each $0 \leq k \leq p$ and $x \in X$;

(b) $\lim_{n\to\infty} Q^{\boldsymbol{\varphi}^{*-i}_n}_i(y|x, \varphi'_n) = Q^{\boldsymbol{\varphi}^{-i}_\infty}_i(y|x, \varphi'_\infty)$ for each $x, y \in X$;

(c) $\lim_{n\to\infty}\sum_{x\in X}[Q_i^{\varphi_n^{*-i}}(y|x,\varphi_n')]^t\nu_n(x)=\sum_{x\in X}[Q_i^{\varphi_\infty^{-i}}(y|x,\varphi_\infty')]^t\nu(x)$ for each $y\in X$ and $t\in\mathbb{N}$, where $[Q_i^{\varphi_n^{*-i}}(\cdot|x,\varphi_n')]^t$ and $[Q_i^{\varphi_\infty^{-i}}(\cdot|x,\varphi_\infty')]^t$ denote the corresponding t-step transition laws for each $t,n\in\mathbb{N}$;

(d) $\sup_{n\in\mathbb{N}}\sum_{x\in X}\omega^2(x)\nu_n(x)\le\sum_{x\in X}\omega^2(x)\nu(x)<\infty$.

Then, using these facts and (69), it can be established in exactly the same way as in the proof of Proposition 3 that the sequence $\{\mathrm{LP}_{i,n}^{\varphi_n^*}\}$ and the "limit" $\mathrm{LP}_i^{\varphi_\infty}$ satisfy the Assumption 2.1 in Alvarez-Mena and Hernández-Lerma (2002). Hence, by the similar proof of Lemma 5, we know that the accumulation point $\boldsymbol{\varphi}_\infty$ is a constrained Nash equilibrium for $\mathscr{G}$. □

Corollary 1 *Suppose that there exists a constant $M'>0$ such that $|c_i^k(x,\boldsymbol{a})|\le M'$ for each $(x,\boldsymbol{a})\in K$. Then, under Assumptions 1(a–b) and 2, the countable-state game in Alvarez-Mena and Hernández-Lerma (2006) has a Nash equilibrium.*

Proof It is obvious that Assumptions 1(a–c,e) are satisfied for the special case that $\omega(x):=1$ for each $x\in X$. Hence, (25) holds. Under Assumptions 1(a–b), 2 and the boundedness hypothesis of the cost functions, Remark 4 here together with Theorem 3.6(c) in Alvarez-Mena and Hernández-Lerma (2006) directly imply the desired result. □

Remark 8 In Alvarez-Mena and Hernández-Lerma (2006), to obtain the existence of Nash equilibria, special hypothesis for transition laws (i.e., Assumption 3.4 therein) is required, which, however, is unnecessary here by our proof.

4 An example

In this section, we illustrate our results with a priority queueing system. We formulate it as a nonzero-sum game.

Example 1 (A priority queueing system) Consider a discrete-time priority queueing system consisting of two queues and a server (i.e., player 1) and an entrance controller (i.e., player 2). The state of the system at time t is given by a vector $\boldsymbol{x}_t:=(x_t^1,x_t^2)$, where x_t^i denotes the number of customers in queue i. The state space is $X:=\mathbb{Z}\times\mathbb{Z}$, where $\mathbb{Z}:=\{0,1,\ldots\}$. At the beginning of each slot $[t,t+1)$, a customer arrives with a fixed probability λ_i to queue i. We assume the arrival processes between queue 1 and queue 2 are independent. The queue 1 is the priority queue containing priority customers, and queue 2 is the nonpriority queue. Player 1 always serves the priority customers if queue 1 is not empty. The arriving nonpriority customer will be admitted or rejected by player 2. At each time t, when $x_t^1\neq0$, player 1 chooses action $a_1\in[\theta_1,\theta_2]$ which means he/she serves a priority customer with service rate a_1, where $0<\theta_1<\theta_2\le1$ are some fixed constants. When $x_t^1=0$ and $x_t^2\neq0$, player 1 chooses action $a_1\in\{0,\theta_1\}$, where $a_1=\theta_1$ means that he/she serves a nonpriority customer with service rate θ_1, and $a_1=0$ means that player 1 is on vacation. If player 1 chooses $a_1=0$, he/she must pay holding cost for nonpriority customers as a penalization. If the system is empty, player 1 is idle and he/she can only choose action $a_1=0$. At each epoch t, player 2 chooses action $a_2\in\{0,1\}$, where action

$a_2 = 1$ means accepting the arriving nonpriority customer to queue 2 and $a_2 = 0$ means rejecting him/her. If queue 2 is empty, player 2 does not reject the arriving customer. We assume that the service time of a customer is one slot, and once the service is completed the customer will leave the system immediately. If the service is unsuccessful, the customer will be served in the next slot. If a customer enters an empty queue, he/she is available for service at the beginning of the following slot. There are a holding cost $c_1^0(x^1, x^2, a_1, a_2)$ and a service cost $c_1^1(x^1, x^2, a_1, a_2)$ for player 1, where the superscript i in x^i is used to indicate the player i. For player 2, he/she has a holding cost $c_2^0(x^1, x^2, a_1, a_2)$ and a rejection cost $c_2^1(x^1, x^2, a_1, a_2)$. Each player i wants to minimize his/her expected discounted cost V_i^0 while keeping his/her associated expected discounted cost V_i^1 bounded above by some constant $d_i > 0$ for each $i = 1, 2$. The initial distribution of this system is denoted by ν.

We next formulate this model as a discrete-time constrained Markov game. The corresponding admission action sets $A_i(x^1, x^2)$, the transition laws $Q(\cdot|x^1, x^2, a_1, a_2)$ and cost functions $c_i^k(x^1, x^2, a_1, a_2)$ $(k = 0, 1)$ for player i are defined as follows:

For player 1, $A_1(x^1, x^2) = [\theta_1, \theta_2]$ for each $x^1 \neq 0$ and $x^2 \in \mathbb{Z}$, $A_1(0, x^2) = \{0, \theta_1\}$ for each $x^2 \geq 1$, and $A_1(0, 0) = \{0\}$. For player 2, $A_2(x^1, x^2) = \{0, 1\}$ if $x^1 \in \mathbb{Z}$ and $x^2 \neq 0$, and $A_2(x^1, 0) = \{1\}$ for each $x^1 \in \mathbb{Z}$.

For convenience, for any variable b, let $\overline{b}$ stand for $1 - b$. Then, for each $x^1 \neq 0$, $x^2 \in \mathbb{Z}$ and $a_1 \in [\theta_1, \theta_2]$,

$$Q(y^1, y^2|x^1, x^2, a_1, 1) = \begin{cases} \lambda_1\lambda_2\overline{a_1} & \text{if } y^1 = x^1 + 1, y^2 = x^2 + 1, \\ \lambda_2[\lambda_1 a_1 + \overline{\lambda_1}\overline{a_1}] & \text{if } y^1 = x^1, y^2 = x^2 + 1, \\ \lambda_2\overline{\lambda_1}a_1 & \text{if } y^1 = x^1 - 1, y^2 = x^2 + 1, \\ \overline{\lambda_2}\lambda_1\overline{a_1} & \text{if } y^1 = x^1 + 1, y^2 = x^2, \\ \overline{\lambda_2}[\lambda_1 a_1 + \overline{\lambda_1}\overline{a_1}] & \text{if } y^1 = x^1, y^2 = x^2, \\ \overline{\lambda_1}a_1\overline{\lambda_2} & \text{if } y^1 = x^1 - 1, y^2 = x^2, \\ 0 & \text{otherwise,} \end{cases} \tag{73}$$

$$Q(y^1, y^2|x^1, x^2, a_1, 0) = \begin{cases} \lambda_1\overline{a_1} & \text{if } y^1 = x^1 + 1, y^2 = x^2, \\ \overline{\lambda_1}\overline{a_1} + \lambda_1 a_1 & \text{if } y^1 = x^1, y^2 = x^2, \\ \overline{\lambda_1}a_1 & \text{if } y^1 = x^1 - 1, y^2 = x^2, \\ 0 & \text{otherwise,} \end{cases} \tag{74}$$

and for each $x^1 = 0$, $x^2 \geq 1$, and $a_1 \in \{0, \theta_1\}$,

$$Q(y^1, y^2|0, x^2, a_1, 1) = \begin{cases} \lambda_1\lambda_2\overline{a_1} & \text{if } y^1 = 1, y^2 = x^2 + 1, \\ \overline{\lambda_1}\lambda_2\overline{a_1} & \text{if } y^1 = 0, y^2 = x^2 + 1, \\ \lambda_1(\overline{\lambda_2}\overline{a_1} + \lambda_2 a_1) & \text{if } y^1 = 1, y^2 = x^2, \\ \overline{\lambda_1}(\overline{\lambda_2}\overline{a_1} + \lambda_2 a_1) & \text{if } y^1 = 0, y^2 = x^2, \\ \lambda_1\overline{\lambda_2}a_1 & \text{if } y^1 = 1, y^2 = x^2 - 1, \\ \overline{\lambda_1}a_1\overline{\lambda_2} & \text{if } y^1 = 0, y^2 = x^2 - 1, \\ 0 & \text{otherwise,} \end{cases} \tag{75}$$

and

$$Q(y^1, y^2|0, x^2, a_1, 0) = \begin{cases} \lambda_1\overline{a_1} & \text{if } y^1 = 1, y^2 = x^2, \\ \overline{\lambda_1}\overline{a_1} & \text{if } y^1 = 0, y^2 = x^2, \\ \lambda_1 a_1 & \text{if } y^1 = 1, y^2 = x^2 - 1, \\ \overline{\lambda_1} a_1 & \text{if } y^1 = 0, y^2 = x^2 - 1, \\ 0 & \text{otherwise,} \end{cases} \tag{76}$$

and

$$Q(y^1, y^2|0, 0, 0, 1) = \begin{cases} \lambda_1\lambda_2 & \text{if } y^1 = 1, y^2 = 1, \\ \overline{\lambda_1}\lambda_2 & \text{if } y^1 = 0, y^2 = 1, \\ \lambda_1\overline{\lambda_2} & \text{if } y^1 = 1, y^2 = 0, \\ \overline{\lambda_1} \cdot \overline{\lambda_2} & \text{if } y^1 = 0, y^2 = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{77}$$

Let $I_D(z)$ denote the indicator function of a set D. Then, the cost functions are defined as follows:

$$c_1^0(x^1, x^2, a_1, a_2) :=p_1 x^1 + p_2 x^2 I_{\{0\}}(a_1), \text{ and } \quad c_1^1(x^1, x^2, a_1, a_2) := h(a_1), \tag{78}$$
$$c_2^0(x^1, x^2, a_1, a_2) :=p_2 x^2 I_{\{b\neq 0\}}(a_1), \text{ and } \quad c_2^1(x^1, x^2, a_1, a_2) := p_3 I_{\{0\}}(a_2), \tag{79}$$

where $p_1 > 0$ ($p_2 > 0$) denotes the fixed holding cost per customer in queue 1 (respectively, queue 2), the function h denotes the service cost for player 1, and p_3 denotes the rejection cost for player 2. Moreover, it is natural to assume that $h(0) = 0$. From (78)–(79), we assume player 2 always pays the holding cost for the nonpriority customers except player 1 is on vacation because the holding cost is payed by player 1 as a penalization in this case. The aim now is to find conditions under which there exists a constrained Nash equilibrium. To do so, we consider the following conditions.

Condition 1 *(a)* $\alpha(1 + 3\lambda_1 + 3\lambda_2 + 2\lambda_1\lambda_2) < 1$*;*
(b) $\sum_{x^1\in\mathbb{Z}}\sum_{x^2\in\mathbb{Z}}(x^1 + x^2)^2\nu(x^1, x^2) < \infty$*;*
(c) The function $h(a_1)$ *is continuous in* $a_1 \in [\theta_1, \theta_2]$*;*
(d) $h(\theta_1) < d_1$.

Proposition 4 *Under Condition 1, the above controlled queueing system satisfies Assumptions 1 and 2, therefore (by Theorem 1), there exists a constrained Nash equilibrium.*

Proof Observe that the state space X can be 'enumerated' in the order

$$(0, 0), (0, 1), (1, 0), (0, 2), (1, 1), (2, 0), \ldots.$$

Let $\omega(x^1, x^2) := x^1+x^2+1$ for all $(x^1, x^2) \in X$. Then the function ω is nondecreasing on X in the above order, and it is a moment function.

We shall first verify Assumption 1(d). By Condition 1(c), there exists a constant $L > 0$ such that $\sup_{a_1 \in [\theta_1, \theta_2]} |h(a_1)| \leq L$. It follows from (78)–(79) that

$$|c_1^0(x^1, x^2, a_1, a_2)| \leq (p_1 + p_2)\omega(x^1, x^2), \text{ and } |c_1^1(x^1, x^2, a_1, a_2)| \leq L\omega(x^1, x^2),$$
$$|c_2^0(x^1, x^2, a_1, a_2)| \leq p_2\omega(x^1, x^2) \text{ and } |c_2^1(x^1, x^2, a_1, a_2)| \leq p_3\omega(x^1, x^2),$$

for each $(x^1, x^2, a_1, a_2) \in K$ which implies Assumption 1(d) with $M := L + p_1 + p_2 + p_3$. By (73)–(77) and Condition 1(a), for each $x^1 \neq 0, x^2 \geq 0$ and $a_1 \in [\theta_1, \theta_2]$,

$$\begin{aligned}
&\sum_{y^1 \in \mathbb{Z}} \sum_{y^2 \in \mathbb{Z}} \omega^2(y^1, y^2) Q(y^1, y^2 | x^1, x^2, a_1, 1) \\
&\quad \leq \omega^2(x^1, x^2) + 3(\lambda_2 + \lambda_1)\omega(x^1, x^2) + 2\lambda_1\lambda_2 \\
&\quad \leq (1 + 3\lambda_1 + 3\lambda_2 + 2\lambda_1\lambda_2)\omega^2(x^1, x^2), \qquad (80) \\
&\sum_{y^1 \in \mathbb{Z}} \sum_{y^2 \in \mathbb{Z}} \omega^2(y^1, y^2) Q(y^1, y^2 | x^1, x^2, a_1, 0) \\
&\quad \leq \omega^2(x^1, x^2) + 2\lambda_1\omega(x^1, x^2) + \lambda_1 \\
&\quad \leq (1 + 3\lambda_1 + 3\lambda_2 + 2\lambda_1\lambda_2)\omega^2(x^1, x^2), \qquad (81)
\end{aligned}$$

and for each $x^2 \geq 1$ and $a_1 \in \{0, \theta_1\}$,

$$\begin{aligned}
&\sum_{y^1 \in \mathbb{Z}} \sum_{y^2 \in \mathbb{Z}} \omega^2(y^1, y^2) Q(y^1, y^2 | 0, x^2, a_1, 1) \\
&\quad \leq \omega^2(0, x^2) + (3\lambda_2 + 3\lambda_1 + 2\lambda_1\lambda_2)\omega(0, x^2) \\
&\quad \leq (1 + 3\lambda_1 + 3\lambda_2 + 2\lambda_1\lambda_2)\omega^2(0, x^2), \qquad (82)
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{y^1 \in \mathbb{Z}} \sum_{y^2 \in \mathbb{Z}} \omega^2(y^1, y^2) Q(y^1, y^2 | 0, x^2, a_1, 0) \\
&\quad \leq \omega^2(0, x^2) + 2(\lambda_1 - a_1)\omega(0, x^2) + \lambda_1 + a_1 \\
&\quad \leq (1 + 3\lambda_1 + 3\lambda_2 + 2\lambda_1\lambda_2)\omega^2(0, x^2), \qquad (83)
\end{aligned}$$

and when $x^1 = x^2 = 0$,

$$\begin{aligned}
&\sum_{y^1 \in \mathbb{Z}} \sum_{y^2 \in \mathbb{Z}} \omega^2(y^1, y^2) Q(y^1, y^2 | 0, 0, 0, 1) \\
&\quad = 1 + 3\lambda_2 + 3\lambda_1 + 2\lambda_1\lambda_2 \\
&\quad \leq (1 + 3\lambda_1 + 3\lambda_2 + 2\lambda_1\lambda_2)\omega^2(0, 0), \qquad (84)
\end{aligned}$$

which together with Condition 1(a) verify Assumption 1(e) with $\beta := (1 + 3\lambda_1 + 3\lambda_2 + 2\lambda_1\lambda_2)^{\frac{1}{2}}$. Then, it follows from (73)–(77) that Assumption 1(c) holds. By Conditions 1(b–c) and the description of the system, Assumptions 1(a,b,f) are satisfied. Let $\tilde{\varphi}^1 \in \Pi_s^1$ and $\tilde{\varphi}^2 \in \Pi_s^2$ be the stationary policy such that $\tilde{\varphi}^1(\{0\}|0,0) = 1$, and $\tilde{\varphi}^1(\{\theta_1\}|x^1,x^2) = 1$ for $(x^1,x^2) \neq (0,0)$, $\tilde{\varphi}^2(\{1\}|x^1,x^2) = 1$ for $(x^1,x^2) \in X$. According to (5) and Condition 1(d), the associated α-discounted costs $V_1^1(\nu,\tilde{\varphi}^1,\varphi^2) \leq h(\theta_1) < d_1$ and $V_2^1(\nu,\varphi^1,\tilde{\varphi}^2) = 0$ for each $(\varphi^1,\varphi^2) \in \Pi_s^1 \times \Pi_s^2$, which imply Assumption 2. Hence, Theorem 1 holds. □

Remark 9 It should be noted that in Example 1, the cost functions c_i^0 $(i = 1,2)$ are allowed to be *unbounded*. Hence, the conditions in Altman and Shwartz (2000) and Alvarez-Mena and Hernández-Lerma (2006) fail to hold. That is because the cost functions in Altman and Shwartz (2000) and Alvarez-Mena and Hernández-Lerma (2006) are both assumed to be bounded.

5 Conclusion

This paper has dealt with nonzero-sum discrete-time constrained stochastic games with respect to discounted cost criteria. Under reasonably mild condition, we have shown the existence of constrained Nash equilibria. In relation to the most closely related works, our results extend those obtained in Altman and Shwartz (2000) and Alvarez-Mena and Hernández-Lerma (2006). Our main idea is to show that the original constrained game can be approximated in some sense by some special games that have Nash equilibria. There remains, however, an interesting issue. The original game in this paper actually is approximated by two steps. We first use a sequence of finite games to approximate an auxiliary game associated with the original game, and then show that the original game can be approximated by a sequence of auxiliary games. Hence, it will be interesting to consider using other special game directly to approximate the original game.

Acknowledgments This research was partially supported by the NSFC and GDUPS. We also thank the anonymous referee for constructive comments.

References

Altman E (1999) Constrained Markov decision processes. Chapman Hall & CRC, Florida

Altman E, Shwartz A (2000) Constrained Markov games: Nash equilibria. Ann Int Soc Dynam Games, vol 5. Birkhäuser, Boston, pp 213–221

Altman E, Avrachenkov K, Bonneau N, Debbah M, EI-Azouzi R, Sadoc Menasche D (2008) Constrained cost-coupled stochastic games with independent state processes. Oper Res Lett 36:160–164

Alvarez-Mena J, Hernández-Lerma O (2002) Convergence of the optimal values of constrained Markov control processes. Math Methods Oper Res 55:461–484

Alvarez-Mena J, Hernández-Lerma O (2006) Existence of Nash equilibria for constrained stochastic games. Math Methods Oper Res 63:261–285

Bäuerle N, Rieder U (2011) Markov decision processes with applications to finance. Springer, Heidelberg

Cavazos-Cadena R (1986) Finite-state approximations for denumerable state discounted Markov decision processes. Appl Math Optim 14:1–26

Durán J (2000) On dynamic programming with unbounded returns. Econom Theory 15:339–352

Federgruen A (1978) On N-person stochastic games with denumerable state space. Adv Appl Probab 10:452–471

Filar J, Vrieze K (1997) Competitive Markov decision processes. Springer, New York

Fink AM (1964) Equilibrium in a stochastic n-person game. J Sci Hiroshima Univ Ser A-I Math 28:89–93

González-Trejo JI, Hernández-Lerma O, Hoyos-Reyes LF (2002) Minimax control of discrete-time stochastic systems. SIAM J Control Optim 41:1626–1659

Guo XP, Hernández-Lerma O (2009) Continuous-time Markov decision processes. Springer, Berlin

Guo XP, Yang J (2008) A new condition and approach for zero-sum stochastic games with average payoffs. Stoch Anal Appl 26:537–561

Hernández-Lerma O, Lasserre JB (1996) Discrete-time Markov control processes. Springer, New York

Hernández-Lerma O, Lasserre JB (1999) Further topics on discrete-time Markov control processes. Springer, New York

Küenle HU (1994) On Nash equilibrium solutions in nonzero-sum stochastic games with complete information. Internat J Game Theory 23:303–324

Maitra AP, Sudderth WD (1996) Discrete gambling and stochastic games. Springer, New York

Nowak AS (1999) Sensitive equilibria for ergodic stochastic games with countable state spaces. Math Methods Oper Res 50:65–76

Nowak AS (2003) On a new class of nonzero-sum discounted stochastic games having stationary Nash equilibrium points. Internat J Game Theory 32:121–132

Nowak AS (2007) On stochastic games in economics. Math Methods Oper Res 66:513–530

Puterman ML (1994) Markov decision processes: discrete stochastic dynamic programming. Wiley, New York

Sennott LI (1994) Nonzero-sum stochastic games with unbounded costs: discounted and average cost cases. Z Oper Res 40:145–162

Sennott LI (1999) Stochastic dynamic programming and the control of queueing systems. Wiley, New York

Sobel MJ (1971) Noncooperative stochastic games. Ann Math Statist 42:1930–1935