Stationary Markov Nash equilibria for nonzero-sum constrained ARAT Markov games

François Dufour[*]   Tomás Prieto-Rumeau[†]

Abstract

We consider a nonzero-sum Markov game on an abstract measurable state space with compact metric action spaces. The goal of each player is to maximize his respective discounted payoff function under the condition that some constraints on a discounted payoff are satisfied. We are interested in the existence of a Nash or noncooperative equilibrium. Under suitable conditions, which include absolute continuity of the transitions with respect to some reference probability measure, additivity of the payoffs and the transition probabilities (ARAT condition), and continuity in action of the payoff functions and the density function of the transitions of the system, we establish the existence of a constrained stationary Markov Nash equilibrium, that is, the existence of stationary Markov strategies for each of the players yielding an optimal profile within the class of all history-dependent profiles.

Keywords: Nash equilibrium; Nonzero-sum games; Constrained games; ARAT games.
AMS 2020 Subject Classification: 91A10, 91A15.

1 Introduction

Nonzero-sum stochastic games are nowadays a largely developed and yet very active field of research. It started in the 1950's with the pioneering work of J. Nash and it has developed in many directions. It is impossible to give a complete overview here, though we refer the reader to the recent comprehensive survey [16] and the citations therein. The existence of stationary equilibria is a central problem in the theory nonzero-sum stochastic games, and it can be traced back to [13].

Results on the existence of stationary equilibria have then been obtained for different families of games. We can mention [13, 21, 24] for ARAT (Additive Reward and Additive Transition) games, [25] for games with finite action spaces and state independent transition kernel, [20] for stationary correlated equilibria (see [9] for related results), [10] for noisy stochastic games, [22] for a class of games where the transition probability measure is a convex combination of a finite number of probability measures and where the coefficients depend on the state and actions variables, and [11] for a family of games satisfying a general condition called "decomposable coarser transition kernels".

For the particular case of discounted stochastic ARAT games, its analysis goes back to the 1970s and 1980s with results on the existence of stationary equilibria under the assumption that the state space is Borel and the action spaces are finite [13, 24]. These first results were then generalized in [21] in several directions by considering an ARAT game with general measurable state space and compact action spaces, showing the existence of nonrandomized stationary ϵ-equilibria assuming

[*]Institut Polytechnique de Bordeaux; INRIA Bordeaux Sud Ouest, Team: ASTRAL; IMB, Institut de Mathématiques de Bordeaux, Université de Bordeaux, France `francois.dufour@math.u-bordeaux.fr`
[†]Statistics Department, UNED, Madrid, Spain. e-mail: `tprieto@ccia.uned.es`   (Author for correspondence)

that the initial distribution is nonatomic. In [15], the authors assume that the state space is Borel, the action space is finite, and also that the transition probability distribution is absolutely continuous with respect to a nonatomic probability measure, and then show the existence of pure Nash equilibria in the class of stationary almost Markov strategies. Other type of criteria are analyzed in [18, 29] for ARAT stochastic games. More specifically, the expected total cost criterion is studied in [18], while the authors focus on the limiting average criterion in [29].

In the aforementioned references on *unconstrained* games, the usual technique is to reduce the infinite horizon discounted game to a one-shot game with some terminal payoff function v. Each player addresses his associated control problem (with terminal payoff v) by solving the corresponding *dynamic programming equation*. The main idea is to obtain a terminal payoff function v^* satisfying a suitable fixed point property (related to fixed point theorems for correspondences). Nash equilibria for the players are then obtained from this terminal payoff function v^*. To use such fixed point results, a *convexification* step is needed and a final *de-convexification* step becomes also necessary in order to get back to the original problem. This important procedure, obtained in [20], is a key point to prove the existence of stationary equilibria for different types of games or different notions of equilibrium (see [23] for examples arising from economic theory or [20] for stationary equilibria with public randomization). Finally, [11] gives a general de-convexification procedure based on the already mentioned notion of a decomposable coarser transition kernel.

A natural extension of the *unconstrained* game models described above is to impose constraints on the players. More precisely, each player has a constraint function and he has to maximize his own discounted payoff under the condition that the expected discounted payoff associated to the constraint function is above some given level. A Nash equilibrium consists of policies of the players such that the constraints of each player are satisfied and, in addition, no player improves his payoff when unilaterally varying his strategy while still satisfying his own constraint. Constrained games are, by far, less developed than the unconstrained counterpart.

In the reference [3], the authors study a constrained game with finite state and actions spaces. In the same vein as for constrained Markov decision processes, the approach developed in [3] consists in considering the family of occupation measures of the state-action process of each player. From these occupation measures, the corresponding policies of the players are identified by disintegration. Then, when fixing the policies of the remaining players, each player finds his optimal constrained response by solving a linear programming problem stated in, again, the family of occupation measures. In this way, a correspondence on the family of occupation measures is defined, and any fixed point of this correspondence yields a constrained Nash equilibrium. In [4], these results are generalized to a game model with countable state space and compact action spaces which is, roughly speaking, "nearly finite" in each transition of the dynamic system [4, Assumption 3.4]. Additional results have been obtained in, e.g., [2, 28], where the authors consider a finite state and action game model with independent state and action dynamics, in which interaction between players occurs only on the payoff functions.

In this paper, we will consider a constrained nonzero-sum stochastic game with ARAT structure under the discounted payoff criterion. The state space is an abstract measurable space, the action spaces of the players are compact metric spaces, the payoff and constraint functions are bounded Carathéodory (i.e., measurable in state, continuous in action) functions. It is important to emphasize that the results presented in [3] cannot be generalized to the general state space case. Indeed, compared to [3], dealing with constraints and extending the state space from a finite or countable space to a general measurable (not even metric) space entails serious technical difficulties and it is —we believe— far from being straightforward. To work in such a general context, one needs to introduce additional hypotheses including continuity type assumptions of the transition

2

kernel and the ARAT separability condition similar to those used in [21] for unconstrained games. In particular, the ARAT condition is used to uniquely identify a Markov policy of a player starting from a state-action occupation measure and to characterize the corresponding payoffs. We use the weak-strong topology for the space of occupation measures, which is compatible with the so-called narrow topology (also known as the stable topology) of Young measures, and the latter will be used for the spaces of Markov policies of the players. In this fairly general setting, we will be able to establish the existence of a constrained Nash equilibrium —consisting of stationary Markov policies of the players— within the class of all history-dependent policies of the players. Finally, it is worth mentioning that the existence of a stationary Markov equilibrium is an important and difficult issue in the literature on stochastic games. In support of this, we can cite A. Nowak who wrote in [19]: *This question is highly nontrivial and not as yet completely resolved.* As emphasized in [11], this issue still remains an important problem. One aspect of our contribution is to show that this type of existence result holds for a class of constrained stochastic games with general state space.

The rest of the paper is organized as follows. In Section 2 we define the constrained game model and state our assumptions. We introduce the occupation measures of the players in Section 3 and we also state some preliminary results. Finally, the existence of constrained Nash equilibria is addressed in Section 4. We note that we deal here with the two-player game so as to handle a simpler notation. Our results can be easily generalized to an m-player model.

2 Model and assumptions

2.1 Notation and preliminary results

On a measurable space $(\mathbf{\Omega}, \mathcal{F})$ we will consider the set of finite signed measures $\mathcal{M}(\mathbf{\Omega})$, the set of finite nonnegative measures $\mathcal{M}^+(\mathbf{\Omega})$, and the set of probability measures $\mathcal{P}(\mathbf{\Omega})$. On $\mathcal{P}(\mathbf{\Omega})$, the s-topology is the coarsest topology that makes $\mu \mapsto \mu(D)$ continuous for every $D \in \mathcal{F}$. For $\lambda \in \mathcal{P}(\mathbf{\Omega})$, let us consider the set $\mathcal{P}_\lambda(\mathbf{\Omega})$ of probability measures η on $\mathbf{\Omega}$ which are absolutely continuous with respect to λ; in symbols, $\eta \ll \lambda$. The set $\mathcal{P}_\lambda^e(\mathbf{\Omega})$ consists of all probability measures η on $\mathbf{\Omega}$ which are equivalent to λ, that is: $\eta \ll \lambda$ and $\lambda \ll \eta$, written $\eta \sim \lambda$.

For a product of measurable spaces we will always consider the product σ-algebra. Let $(\mathbf{\Omega}, \mathcal{F})$ and $(\mathbf{\Omega}', \mathcal{F}')$ be two measurable spaces. A kernel on $\mathbf{\Omega}'$ given $\mathbf{\Omega}$ is a mapping $Q : \mathbf{\Omega} \times \mathcal{F}' \to \mathbb{R}^+$ such that $\omega \mapsto Q(B|\omega)$ is measurable on $(\mathbf{\Omega}, \mathcal{F})$ for every $B \in \mathcal{F}'$, and $B \mapsto Q(B|\omega)$ is in $\mathcal{M}^+(\mathbf{\Omega}')$ for every $\omega \in \mathbf{\Omega}$. If $Q(\mathbf{\Omega}'|\omega) = 1$ for all $\omega \in \mathbf{\Omega}$ then we say that Q is a *stochastic kernel* and if $Q(\mathbf{\Omega}'|\omega) \leq 1$ for all $\omega \in \mathbf{\Omega}$ then we say that Q is a *substochastic kernel*. Let Q be a kernel on $\mathbf{\Omega}'$ given $\mathbf{\Omega}$ and let f be a bounded measurable function $f : \mathbf{\Omega}' \to \mathbb{R}$. We will denote by $Qf : \mathbf{\Omega} \to \mathbb{R}$ the measurable function

$$Qf(\omega) = \int_{\mathbf{\Omega}'} f(z)Q(dz|\omega) \quad \text{for } \omega \in \mathbf{\Omega}.$$

If Q is a stochastic (or substochastic) kernel on $\mathbf{\Omega}'$ given $\mathbf{\Omega}$ and $\mu \in \mathcal{P}(\mathbf{\Omega})$, we denote by μQ the probability measure (or finite measure) on $(\mathbf{\Omega}', \mathcal{F}')$ given by

$$B \mapsto \mu Q\,(B) = \int_{\mathbf{\Omega}} Q(B|\omega)\mu(d\omega) \quad \text{for } B \in \mathcal{F}'.$$

We will also write $\mu Q\,(d\omega') = \int_{\mathbf{\Omega}} Q(d\omega'|\omega)\mu(d\omega)$. In addition, we define $\mu \otimes Q$ as the probability measure (or finite measure) on the product space $(\mathbf{\Omega} \times \mathbf{\Omega}', \mathcal{F} \otimes \mathcal{F}')$ given by

$$(\mu \otimes Q)(A \times B) = \int_A Q(B|\omega)\mu(d\omega) \quad \text{for } (A, B) \in \mathcal{F} \times \mathcal{F}'.$$

We will write $(\mu \otimes Q)(d\omega, d\omega') = Q(d\omega'|\omega)\mu(d\omega)$. Given $\mu \in \mathcal{M}(\boldsymbol{\Omega} \times \boldsymbol{\Omega}')$, the marginal measures are $\mu^{\boldsymbol{\Omega}} \in \mathcal{M}(\boldsymbol{\Omega})$ and $\mu^{\boldsymbol{\Omega}'} \in \mathcal{M}(\boldsymbol{\Omega}')$ defined by $\mu^{\boldsymbol{\Omega}}(\cdot) = \mu(\cdot \times \boldsymbol{\Omega}')$ and $\mu^{\boldsymbol{\Omega}'}(\cdot) = \mu(\boldsymbol{\Omega} \times \cdot)$.

Throughout this paper, any metric space $\mathbf{S}$ will be endowed with its Borel σ-algebra $\mathfrak{B}(\mathbf{S})$. Also, when considering the product of a finite family of metric spaces, we will consider the product topology (which makes the product again a metric space). We say that $f : \boldsymbol{\Omega} \times \mathbf{S} \to \mathbf{S}'$, where $\mathbf{S}'$ is a metric space, is a *Carathéodory function* if $f(\cdot, s)$ is measurable on $\boldsymbol{\Omega}$ for every $s \in \mathbf{S}$ and $f(\omega, \cdot)$ is continuous on $\mathbf{S}$ for every $\omega \in \boldsymbol{\Omega}$. The family of the so-defined Carathéodory functions is denoted by $Car(\boldsymbol{\Omega} \times \mathbf{S}, \mathbf{S}')$. The family of Carathéodory functions which, in addition, are bounded is denoted by $Car_b(\boldsymbol{\Omega} \times \mathbf{S}, \mathbf{S}')$. When the metric space $\mathbf{S}$ is separable then any $f \in Car(\boldsymbol{\Omega} \times \mathbf{S}, \mathbf{S}')$ is a jointly measurable function on $(\boldsymbol{\Omega} \times \mathbf{S}, \mathcal{F} \otimes \mathfrak{B}(\mathbf{S}))$; see [1, Lemma 4.51].

If $\mathbf{S}$ is a Polish space (a complete and separable metric space), on $\mathcal{M}(\boldsymbol{\Omega} \times \mathbf{S})$ we will consider the *ws*-topology (weak-strong topology) which is the coarsest topology for which the mappings

$$\mu \mapsto \int_{\boldsymbol{\Omega} \times \mathbf{S}} f(\omega, s)\mu(d\omega, ds)$$

for $f \in Car_b(\boldsymbol{\Omega} \times \mathbf{S}, \mathbb{R})$ are continuous. There are other equivalent definitions for this topology as discussed, for instance, in [7, 14, 27]. Note that $\mathcal{P}(\boldsymbol{\Omega} \times \mathbf{S})$ is a closed subset of $\mathcal{M}(\boldsymbol{\Omega} \times \mathbf{S})$. It is clear that convergence in the *ws*-topology in $\mathcal{P}(\boldsymbol{\Omega} \times \mathbf{S})$ implies convergence of the $\boldsymbol{\Omega}$-marginal probability measures in the *s*-topology.

Inequality $\geq$ in $\mathbb{R}^p$ means a componentwise inequality $\geq$, while the inequality $>$ in $\mathbb{R}^p$ is a componentwise strict inequality $>$. Let $\mathbf{1}_p \in \mathbb{R}^p$ be the vector with all components equal to one.

The next result characterizes relative compactness in $\mathcal{P}(\boldsymbol{\Omega} \times \mathbf{S})$. It ressembles to [27, Theorem 3.10] except that, here, $\boldsymbol{\Omega}$ is a measurable space, whereas [27] assumes that $\boldsymbol{\Omega}$ is Borel.

Proposition 2.1 *Let $(\boldsymbol{\Omega}, \mathcal{F})$ be a measurable space and let $\mathbf{S}$ be a Polish space, and consider the ws-topology on $\mathcal{P}(\boldsymbol{\Omega} \times \mathbf{S})$. A necessary and sufficient condition for a set $K \subseteq \mathcal{P}(\boldsymbol{\Omega} \times \mathbf{S})$ to be relatively compact is that for any sequence $\{f_n\}_{n \in \mathbb{N}}$ of functions $f_n \in Car_b(\boldsymbol{\Omega} \times \mathbf{S}, \mathbb{R})$ which decrease pointwise to zero (i.e., $f_n(\omega, s) \downarrow 0$ for each $(\omega, s) \in \boldsymbol{\Omega} \times \mathbf{S}$) we have*

$$\lim_{n \to \infty} \sup_{\mu \in K} \int_{\boldsymbol{\Omega} \times \mathbf{S}} f_n d\mu = 0. \tag{2.1}$$

Proof. The *ws*-topology being the weak topology induced by bounded Carathéodory functions, we use [27, Lemma 3.1] for the necessary condition. For sufficiency, by [7, Theorem 5.2] we have that K is relatively compact if and only if $K^{\boldsymbol{\Omega}} = \{\mu^{\boldsymbol{\Omega}} : \mu \in K\}$ and $K^{\mathbf{S}} = \{\mu^{\mathbf{S}} : \mu \in K\}$ are relatively *s*-compact (for the *s*-topology) and relatively *w*-compact (for weak convergence of probability measures on $\mathbf{S}$), respectively. For relative *s*-compactness, we use [27, Lemma 3.5]. If $\{g_n\}$ is a sequence of bounded measurable functions on $\boldsymbol{\Omega}$ decreasing pointwise to zero, we have $\sup_{\mu \in K} \int_{\boldsymbol{\Omega}} g_n d\mu^{\boldsymbol{\Omega}} = \sup_{\mu \in K} \int_{\boldsymbol{\Omega} \times \mathbf{S}} \bar{g}_n d\mu \to 0$ by hypothesis, where $\bar{g}_n(\omega, s) = g_n(\omega)$, showing that $K^{\boldsymbol{\Omega}}$ is relatively *s*-compact. Similarly, if $\{h_n\}$ is a sequence of bounded continuous functions on $\mathbf{S}$ which decrease pointwise to zero then we have $\sup_{\mu \in K} \int_{\mathbf{S}} h_n d\mu^{\mathbf{S}} = \sup_{\mu \in K} \int_{\boldsymbol{\Omega} \times \mathbf{S}} \bar{h}_n d\mu \to 0$ where $\bar{h}_n(\omega, s) = h_n(s)$. By [27, Lemma 3.2], we obtain that $K^{\mathbf{S}}$ is relatively *w*-compact. $\square$

We now state without proof two results that we will use repeatedly in the following.

Proposition 2.2 *If $(\boldsymbol{\Omega}, \mathcal{F})$ is a measurable space and $\mathbf{S}$ is a Polish space, then $\mathcal{M}(\boldsymbol{\Omega} \times \mathbf{S})$ equipped with the ws-topology is a locally convex Hausdorff topological vector space.*

Lemma 2.3 (Disintegration lemma) *Let $(\mathbf{\Omega}, \mathcal{F})$ be a measurable space and let $\mathbf{S}$ be a Polish space. Let $\varphi : \mathbf{\Omega} \twoheadrightarrow \mathbf{S}$ be a weakly measurable correspondence with nonempty closed values, and let $\mathbf{K}$ be the graph of the correspondence. For every $\mu \in \mathcal{P}(\mathbf{\Omega} \times \mathbf{S})$ such that $\mu(\mathbf{K}) = 1$ there exists a stochastic kernel Q on $\mathbf{S}$ given $\mathbf{\Omega}$ such that*

$$\mu = \mu^{\mathbf{\Omega}} \otimes Q \qquad (2.2)$$

and such that $Q(\varphi(\omega)|\omega) = 1$ for each $\omega \in \mathbf{\Omega}$. Moreover, Q is unique $\mu^{\mathbf{\Omega}}$-almost surely, meaning that if Q and Q' are two stochastic kernels that satisfy (2.2) then for all ω in a set of $\mu^{\mathbf{\Omega}}$-probability one, the probability measures $Q(\cdot|\omega)$ and $Q'(\cdot|\omega)$ coincide.

2.2 Game model and assumptions

Elements of the game model. The elements of the game model are the following.

(a) $(\mathbf{X}, \mathfrak{X})$ is an abstract measurable space, where $\mathbf{X}$ stands for the state space of the system.

(b) The separable metric spaces $\mathbf{A}_1$ and $\mathbf{A}_2$ represent the action sets for player 1 and 2, respectively. For each $x \in \mathbf{X}$ the nonempty measurable sets $\mathbf{A}_1(x) \subseteq \mathbf{A}_1$ and $\mathbf{A}_2(x) \subseteq \mathbf{A}_2$ are the set of feasible actions for players 1 and 2, respectively, when the system is in state $x \in \mathbf{X}$. Let $\mathbf{A} = \mathbf{A}_1 \times \mathbf{A}_2$ and $\mathbf{A}(x) = \mathbf{A}_1(x) \times \mathbf{A}_2(x)$ for each $x \in \mathbf{X}$.

(c) Fix a player $i = 1, 2$. The measurable functions $r_i : \mathbf{X} \times \mathbf{A} \to \mathbb{R}$ and $c_i : \mathbf{X} \times \mathbf{A} \to \mathbb{R}^p$ are the reward and the constraint functions, respectively, of player i. The constraint constant for player i is $\rho_i \in \mathbb{R}^p$. Here, $p \geq 1$ is a fixed integer. Write $\rho = (\rho_1, \rho_2) \in \mathbb{R}^p \times \mathbb{R}^p$.

(e) The transitions of the system are given by a stochastic kernel Q on $\mathbf{X}$ given $\mathbf{X} \times \mathbf{A}$.

(f) The discount factor is $\beta \in (0, 1)$.

(g) The initial distribution is the probability measure $\eta \in \mathcal{P}(\mathbf{X})$.

This game model will be denoted by $\mathcal{G}(\eta, \rho)$. The reason for this notation is that, in the forthcoming, we will need the initial distribution η and the constraint constants (ρ_1, ρ_2) to vary, while the other elements of the game model will remain fixed.

Policies of the players. We define the sets $\mathbf{H}_0 = \mathbf{X}$ and $\mathbf{H}_t = (\mathbf{X} \times \mathbf{A}_1 \times \mathbf{A}_2)^t \times \mathbf{X}$ for $t \geq 1$. The elements of $\mathbf{H}_t$ will be usually denoted by $(x_0, a_0, b_0, \ldots, x_{t-1}, a_{t-1}, b_{t-1}, x_t)$.

Definition 2.4 *(i). A policy for player i (with $i = 1, 2$) is a sequence $\{\pi_i^t\}_{t \in \mathbb{N}}$ of stochastic kernels on $\mathbf{A}_i$ given $\mathbf{H}_t$ such that*

$$\pi_i^t(\mathbf{A}_i(x_t)|x_0, a_0, b_0, \ldots, x_t) = 1 \quad \text{for all } t \geq 0 \text{ and } h_t = (x_0, a_0, b_0, \ldots, x_t) \in \mathbf{H}_t.$$

The family of all policies of player i is denoted by $\mathbf{\Pi}_i$.

(ii). Let $\mathbf{M}_i$ be the family of stochastic kernels π_i on $\mathbf{A}_i$ given $\mathbf{X}$ such that $\pi_i(\mathbf{A}_i(x)|x) = 1$ for each $x \in \mathbf{X}$. We say that $\{\pi_i^t\}_{t \in \mathbb{N}} \in \mathbf{\Pi}_i$ is a stationary Markov policy for player i (with $i = 1, 2$) if there is some $\pi_i \in \mathbf{M}_i$ which verifies

$$\pi_i^t(\cdot|x_0, a_0, b_0, \ldots, x_t) = \pi_i(\cdot|x_t) \quad \text{for all } t \geq 0 \text{ and } h_t = (x_0, a_0, b_0, \ldots, x_t) \in \mathbf{H}_t.$$

The class $\mathbf{\Pi}_i$ is also referred to as the family of *history-dependent policies* for player i. We can identify the class of stationary Markov policies for player i with $\mathbf{M}_i$, and so $\mathbf{M}_i \subseteq \mathbf{\Pi}_i$. In the sequel, we will refer to $\mathbf{M}_i$ as to the family of *stationary Markov policies* for player i. Our conditions below will ensure that these classes of policies are nonempty.

Construction of the state-actions process. We address the existence of a probability space supporting the dynamic system of the game model $\mathcal{G}(\eta, \rho)$. On the product canonical space $\mathbf{H}_\infty = (\mathbf{X} \times \mathbf{A}_1 \times \mathbf{A}_2)^{\mathbb{N}}$, endowed with its product σ-algebra $(\mathfrak{X} \otimes \mathfrak{B}(\mathbf{A}_1) \otimes \mathfrak{B}(\mathbf{A}_2))^{\mathbb{N}}$, define $(X_t, A_t, B_t)_{t \in \mathbb{N}}$ as the coordinates projection operators and write $H_t = (X_0, A_0, B_0, \ldots, X_t)$. Given an initial distribution $\eta \in \mathcal{P}(\mathbf{X})$ and policies $(\pi_1, \pi_2) \in \mathbf{\Pi}_1 \times \mathbf{\Pi}_2$, we can uniquely define a probability measure $\mathbb{P}_{\eta,\pi_1,\pi_2}$ on $\mathbf{H}_\infty$ which verifies the following properties for any $D \in \mathfrak{X}$, $D_1 \in \mathfrak{B}(\mathbf{A}_1)$, and $D_2 \in \mathfrak{B}(\mathbf{A}_2)$, and every $t \geq 0$. Firstly, $\mathbb{P}_{\eta,\pi_1,\pi_2}\{X_0 \in D\} = \eta(D)$. Secondly,

$$\mathbb{P}_{\eta,\pi_1,\pi_2}(A_t \in D_1, B_t \in D_2 | H_t) = \pi_1^t(D_1 | H_t) \cdot \pi_2^t(D_2 | H_t).$$

is the conditional distribution of the actions. Finally, distribution of the next state of the system is $\mathbb{P}_{\eta,\pi_1,\pi_2}(X_{t+1} \in D | X_0, A_0, B_0, \ldots, X_t, A_t, B_t) = Q(D | X_t, A_t, B_t)$. The expectation operator associated to $\mathbb{P}_{\eta,\pi_1,\pi_2}$ will be denoted by $\mathbb{E}_{\eta,\pi_1,\pi_2}$.

Payoffs of the players and equilibrium. Our assumptions below will ensure that the following expressions are well defined and finite. Consider the game model $\mathcal{G}(\eta, \rho)$ and suppose the players use the policies $(\pi_1, \pi_2) \in \mathbf{\Pi}_1 \times \mathbf{\Pi}_2$. The total expected discounted reward for player i is

$$R_i(\eta, \pi_1, \pi_2) = (1 - \beta) \mathbb{E}_{\eta,\pi_1,\pi_2} \Big[\sum_{t=0}^{\infty} \beta^t r_i(X_t, A_t, B_t) \Big] \in \mathbb{R},$$

while the total expected discounted constraint for player i is

$$C_i(\eta, \pi_1, \pi_2) = (1 - \beta) \mathbb{E}_{\eta,\pi_1,\pi_2} \Big[\sum_{t=0}^{\infty} \beta^t c_i(X_t, A_t, B_t) \Big] \in \mathbb{R}^p.$$

We say that the policies $(\pi_1, \pi_2) \in \mathbf{\Pi}_1 \times \mathbf{\Pi}_2$ satisfy the constraint of player i when $C_i(\eta, \pi_1, \pi_2) \geq \rho_i$.

Definition 2.5 *Consider the game model $\mathcal{G}(\eta, \rho)$. The pair of policies $(\pi_1^*, \pi_2^*) \in \mathbf{\Pi}_1 \times \mathbf{\Pi}_2$ is a constrained Nash equilibrium in the class of all history-dependent policies $\mathbf{\Pi}_1 \times \mathbf{\Pi}_2$ when:*

(i) The constraints of both players are satisfied: $C_i(\eta, \pi_1^*, \pi_2^*) \geq \rho_i$ for $i = 1, 2$.

(ii) The following conditions hold:

$$\forall \pi_1 \in \mathbf{\Pi}_1, \ C_1(\eta, \pi_1, \pi_2^*) \geq \rho_1 \ \Rightarrow \ R_1(\eta, \pi_1^*, \pi_2^*) \geq R_1(\eta, \pi_1, \pi_2^*).$$

$$\forall \pi_2 \in \mathbf{\Pi}_2, \ C_2(\eta, \pi_1^*, \pi_2) \geq \rho_2 \ \Rightarrow \ R_2(\eta, \pi_1^*, \pi_2^*) \geq R_2(\eta, \pi_1^*, \pi_2).$$

Assumptions on the game model. First of all, we introduce a Slater-type condition.

Definition 2.6 *We say that the game model $\mathcal{G}(\eta, \rho)$ satisfies the Slater condition if*

$$\forall \pi_1 \in \mathbf{M}_1, \exists \pi_2' \in \mathbf{M}_2, C_2(\eta, \pi_1, \pi_2') > \rho_2 \quad and \quad \forall \pi_2 \in \mathbf{M}_2, \exists \pi_1' \in \mathbf{M}_1, C_1(\eta, \pi_1', \pi_2) > \rho_1.$$

An equivalent formulation of this condition is discussed in Remark 3.4. This is the usual form of the Slater condition for constrained games; see [3, Assumption Π_2] and [4, Assumption 3.3.c]. A stronger uniform Slater condition is imposed in [2, Assumption Π_2] and in [28, Assumption A.2].

Next we state our conditions on the game model. Notice that Assumptions (A.3)–(A.4) impose precisely the additive reward additive transition (ARAT) character of the game.

Assumption A. For an initial distribution $\nu \in \mathcal{P}(\mathbf{X})$ and constraint constants $\theta = (\theta_1, \theta_2)$ in $\mathbb{R}^p \times \mathbb{R}^p$, we suppose that the game model $\mathcal{G}(\nu, \theta)$ satisfies the following conditions.

(A.1) The σ-algebra $\mathfrak{X}$ is countably generated.

(A.2) For each $i = 1, 2$ the action set $\mathbf{A}_i$ is compact and the correspondence from $\mathbf{X}$ to $\mathbf{A}_i$ defined by $x \mapsto \mathbf{A}_i(x)$ is weakly measurable with nonempty compact values.

(A.3) For each player i (with $i = 1, 2$) there exist Carathéodory functions $r_i^j \in \mathcal{C}ar_b(\mathbf{X} \times \mathbf{A}_j, \mathbb{R})$ and $c_i^j \in \mathcal{C}ar_b(\mathbf{X} \times \mathbf{A}_j, \mathbb{R}^p)$ for $j = 1, 2$ such that, for every $(x, a_1, a_2) \in \mathbf{X} \times \mathbf{A}$,

$$r_i(x, a_1, a_2) = r_i^1(x, a_1) + r_i^2(x, a_2) \quad \text{and} \quad c_i(x, a_1, a_2) = c_i^1(x, a_1) + c_i^2(x, a_2).$$

(A.4) There exist a probability measure $\lambda \in \mathcal{P}(\mathbf{X})$ and measurable functions $q_i : \mathbf{X} \times \mathbf{X} \times \mathbf{A}_i \to \mathbb{R}^+$ for $i = 1, 2$ such that for every $D \in \mathfrak{X}$ and every $(x, a_1, a_2) \in \mathbf{X} \times \mathbf{A}$

$$Q(D|x, a_1, a_2) = \int_D [q_1(y, x, a_1) + q_2(y, x, a_2)] \lambda(dy).$$

Also, the following continuity condition holds: for each $i = 1, 2$ and every $x \in \mathbf{X}$ we have

$$\lim_{n \to \infty} \int_{\mathbf{X}} |q_i(y, x, b_n) - q_i(y, x, b)| \lambda(dy) = 0$$

whenever $b_n \to b$ in $\mathbf{A}_i$. Moreover, the probability space $(\mathbf{X}, \mathfrak{X}, \lambda)$ is complete.

(A.5) The initial distribution $\nu \in \mathcal{P}(\mathbf{X})$ and the constraint constants $(\theta_1, \theta_2) \in \mathbb{R}^p \times \mathbb{R}^p$ satisfy the Slater condition in Definition 2.6.

    In the sequel we will suppose that Assumption A holds with no need of an explicit mention. Our main result in this paper —Theorem 4.8 below— establishes that, for the game model $\mathcal{G}(\nu, \theta)$, there exists a pair of stationary Markov policies $(\pi_1^*, \pi_2^*) \in \mathbf{M}_1 \times \mathbf{M}_2$ which is a constrained Nash equilibrium within the class $\mathbf{\Pi}_1 \times \mathbf{\Pi}_2$ of all history-dependent policies. Some comments on Assumption A are given next.

Remark 2.7 *(a). There is no loss of generality in assuming that the initial distribution ν is absolutely continuous with respect to λ. Indeed, we can replace λ in Assumption A with $\bar{\lambda} = (\nu + \lambda)/2$, so that $\nu \in \mathcal{P}_{\bar{\lambda}}(\mathbf{X})$. It is then easily seen that the functions $\bar{q}_i(y, x, a_i) = q_i(y, x, a_i) \frac{d\lambda}{d\bar{\lambda}}(y)$ satisfy Assumption (A.4) and that $(\mathbf{X}, \mathfrak{X}, \bar{\lambda})$ is a complete probability space as well. Hence, from now on we will consider that $\nu \in \mathcal{P}_\lambda(\mathbf{X})$.*

(b). By Assumption (A.2), the correspondences $x \mapsto \mathbf{A}_1(x)$ and $x \mapsto \mathbf{A}_2(x)$ are measurable [1, Lemma 18.2] and they have measurable graph [1, Theorem 18.6]. Therefore,

$$\begin{aligned} \mathbf{K}_i &= \{(x, a_i) : \mathbf{X} \times \mathbf{A}_i : a_i \in \mathbf{A}_i(x)\} \in \mathfrak{X} \otimes \mathfrak{B}(\mathbf{A}_i) \quad \text{for } i = 1, 2 \\ \mathbf{K} &= \{(x, a_1, a_2) : \mathbf{X} \times \mathbf{A}_1 \times \mathbf{A}_2 : (a_1, a_2) \in \mathbf{A}(x)\} \in \mathfrak{X} \otimes \mathfrak{B}(\mathbf{A}_1) \otimes \mathfrak{B}(\mathbf{A}_2). \end{aligned}$$

(c). By the Kuratowski-Ryll-Nardzewski selection theorem [1, Theorem 18.13], there exist measurable selectors for $x \mapsto \mathbf{A}_1(x)$ and $x \mapsto \mathbf{A}_2(x)$. In particular, $\mathbf{M}_1$ and $\mathbf{M}_2$ are nonempty.

3 Occupation measures

3.1 Occupation measures of the policies

Definition of the occupation measures. Given an initial distribution $\eta \in \mathcal{P}(\mathbf{X})$ and policies $(\pi_1, \pi_2) \in \mathbf{\Pi}_1 \times \mathbf{\Pi}_2$ of the players, the occupation measure gives the total expected discounted time spent by the state-action process in a given subset of $\mathbf{X} \times \mathbf{A}$.

Definition 3.1 *Given an initial distribution $\eta \in \mathcal{P}(\mathbf{X})$ and a pair of policies $(\pi_1, \pi_2) \in \mathbf{\Pi}_1 \times \mathbf{\Pi}_2$, the associated occupation measure $\mu_{\eta,\pi_1,\pi_2} \in \mathcal{P}(\mathbf{X} \times \mathbf{A})$ is defined, for $D \in \mathfrak{X}$ and $(D_1, D_2) \in \mathfrak{B}(\mathbf{A}_1) \times \mathfrak{B}(\mathbf{A}_2)$ as*

$$\mu_{\eta,\pi_1,\pi_2}(D \times D_1 \times D_2) = (1 - \beta) \sum_{t=0}^{\infty} \beta^t \mathbb{P}_{\eta,\pi_1,\pi_2}\{X_t \in D, A_t \in D_1, B_t \in D_2\}.$$

The set of all occupation measures is denoted by $\mathcal{O}_\eta = \{\mu_{\eta,\pi_1,\pi_2} : (\pi_1, \pi_2) \in \mathbf{\Pi}_1 \times \mathbf{\Pi}_2\} \subseteq \mathcal{P}(\mathbf{X} \times \mathbf{A})$.

As a direct consequence of Definition 3.1 and Assumption (A.3) we have that the total expected payoffs of the pair of policies $(\pi_1, \pi_2) \in \mathbf{\Pi}_1 \times \mathbf{\Pi}_2$ for the initial distribution $\eta \in \mathcal{P}(\mathbf{X})$ equal

$$R_i(\eta, \pi_1, \pi_2) = \int_{\mathbf{X} \times \mathbf{A}} r_i \, d\mu_{\eta,\pi_1,\pi_2} = \int_{\mathbf{X} \times \mathbf{A}_1} r_i^1 \, d\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X} \times \mathbf{A}_1} + \int_{\mathbf{X} \times \mathbf{A}_2} r_i^2 \, d\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X} \times \mathbf{A}_2} \tag{3.1}$$

for any $i = 1, 2$. Similar equalities hold for the constraints $C_i(\eta, \pi_1, \pi_2)$ of the players.

Occupation measures of stationary Markov policies. Given a pair of stationary Markov policies $(\pi_1, \pi_2) \in \mathbf{M}_1 \times \mathbf{M}_2$, we define the stochastic kernel Q_{π_1,π_2} on $\mathbf{X}$ given $\mathbf{X}$ as

$$Q_{\pi_1,\pi_2}(D|x) = \int_{\mathbf{A}} Q(D|x, a_1, a_2) \pi_1(da_1|x) \pi_2(da_2|x) \tag{3.2}$$

for $x \in \mathbf{X}$ and $D \in \mathfrak{X}$. We denote by Q_{π_1,π_2}^t the t-th composition of Q_{π_1,π_2} with itself and we make the convention that $Q_{\pi_1,\pi_2}^0(\cdot|x) = \delta_x(\cdot)$. It is then easily shown that the $\mathbf{X}$-marginal of the occupation measure of $(\pi_1, \pi_2) \in \mathbf{M}_1 \times \mathbf{M}_2$ for the initial distribution $\eta \in \mathcal{P}(\mathbf{X})$ is

$$\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}} = (1 - \beta) \sum_{t=0}^{\infty} \beta^t \eta Q_{\pi_1,\pi_2}^t = \eta \Big[(1 - \beta) \sum_{t=0}^{\infty} \beta^t Q_{\pi_1,\pi_2}^t \Big]. \tag{3.3}$$

Therefore, we have

$$\mu_{\eta,\pi_1,\pi_2}(dy, da_1, da_2) = \pi_1(da_1|y) \pi_2(da_2|y) \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}}(dy), \tag{3.4}$$

and so $\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X} \times \mathbf{A}_1} = \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}} \otimes \pi_1$ and $\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X} \times \mathbf{A}_2} = \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}} \otimes \pi_2$.

Basic results on occupation measures. Our next result summarizes the main results on the occupation measures of history-dependent policies. We note that statements (i) and (iii)–(iv) below are quite standard, while (ii) makes use of the ARAT nature of the game.

Proposition 3.2 *Let $\eta \in \mathcal{P}(\mathbf{X})$ be an initial distribution and let $(\pi_1, \pi_2) \in \mathbf{\Pi}_1 \times \mathbf{\Pi}_2$ be an arbitrary pair of history-dependent policies.*

(i). The occupation measure μ_{η,π_1,π_2} satisfies the linear equations (written in μ)

$$\mu(\mathbf{K}) = 1 \quad and \quad \mu^{\mathbf{X}} = (1-\beta)\eta + \beta\mu Q \qquad for \ \mu \in \boldsymbol{\mathcal{P}}(\mathbf{X} \times \mathbf{A}).$$

(ii). There exists a pair $(\pi_1', \pi_2') \in \mathbf{M}_1 \times \mathbf{M}_2$ of stationary Markov policies such that

$$\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X} \times \mathbf{A}_1} = \mu_{\eta,\pi_1',\pi_2'}^{\mathbf{X} \times \mathbf{A}_1} \quad and \quad \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X} \times \mathbf{A}_2} = \mu_{\eta,\pi_1',\pi_2'}^{\mathbf{X} \times \mathbf{A}_2}.$$

If $\pi_i \in \mathbf{M}_i$ then we can choose $\pi_i' = \pi_i$.

(iii). If $\eta \in \boldsymbol{\mathcal{P}}_\lambda(\mathbf{X})$ (resp., $\eta \in \boldsymbol{\mathcal{P}}_\lambda^e(\mathbf{X})$) then $\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}} \in \boldsymbol{\mathcal{P}}_\lambda(\mathbf{X})$ (resp., $\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}} \in \boldsymbol{\mathcal{P}}_\lambda^e(\mathbf{X})$).

Proof. (i). It is well known in the theory of MDPs that the occupation measure satisfies the linear constraint in the statement of the proposition; see, for instance, [12, Theorem 6.3.7].

(ii). Given $i = 1, 2$, since the probability measure $\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X} \times \mathbf{A}_i}$ is supported on the set $\mathbf{K}_i$, by Lemma 2.3 it follows that there exists $\pi_i' \in \mathbf{M}_i$ such that

$$\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X} \times \mathbf{A}_i}(dx, da_i) = \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}}(dx)\pi_i'(da_i|x) \tag{3.5}$$

(clearly, if $\pi_i \in \mathbf{M}_i$ then we can let $\pi_i' = \pi_i$). This implies that

$$\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}} = (1-\beta)\eta + \beta\mu_{\eta,\pi_1,\pi_2}Q = (1-\beta)\eta + \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}}Q_{\pi_1',\pi_2'}$$

recalling Assumption (A.4). However, the solution of the equation $\gamma = (1-\beta)\eta + \beta\gamma Q_{\pi_1',\pi_2'}$ for $\gamma \in \boldsymbol{\mathcal{P}}(\mathbf{X})$ is unique and (recall (3.3)) it equals $\mu_{\eta,\pi_1',\pi_2'}^{\mathbf{X}}$. We have thus shown that $\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}} = \mu_{\eta,\pi_1',\pi_2'}^{\mathbf{X}}$ and by (3.4)–(3.5) it follows that $\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X} \times \mathbf{A}_1} = \mu_{\eta,\pi_1',\pi_2'}^{\mathbf{X} \times \mathbf{A}_1}$ and $\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X} \times \mathbf{A}_2} = \mu_{\eta,\pi_1',\pi_2'}^{\mathbf{X} \times \mathbf{A}_2}$, as we wanted to prove.

(iii). The first statement easily follows because $\eta \ll \lambda$ and the transitions are absolutely continuous with respect to λ; recall Assumption (A.4). Regarding the second statement, observe from (3.3) that $(1-\beta)\eta \le \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}}$ and since, by hypothesis, we have $\eta \sim \lambda$, we readily get the result. $\qquad\square$

As a direct consequence of (3.1) and Proposition 3.2(ii) we have the following fact.

Corollary 3.3 *For every initial distribution $\eta \in \boldsymbol{\mathcal{P}}(\mathbf{X})$ and any pair $(\pi_1, \pi_2) \in \boldsymbol{\Pi}_1 \times \boldsymbol{\Pi}_2$, there exist stationary Markov policies $(\pi_1', \pi_2') \in \mathbf{M}_1 \times \mathbf{M}_2$ such that for any $i = 1, 2$ we have*

$$R_i(\eta, \pi_1, \pi_2) = R_i(\eta, \pi_1', \pi_2') \quad and \quad C_i(\eta, \pi_1, \pi_2) = C_i(\eta, \pi_1', \pi_2').$$

If $\pi_1 \in \mathbf{M}_1$ (respectively, $\pi_2 \in \mathbf{M}_2$) then the result holds for $\pi_1' = \pi_1$ (respectively, $\pi_2' = \pi_2$).

Remark 3.4 *By Corollary 3.3, the Slater condition is equivalent to the apparently weaker condition that for any $\pi_1 \in \mathbf{M}_1$ there exists $\pi_2 \in \boldsymbol{\Pi}_2$ with $C_2(\eta, \pi_1, \pi_2) > \rho_2$, and symmetrically for player 1.*

Stationary Markov policies and Young measures. In Proposition 3.2(iii) we have shown that, given any initial distribution $\eta \in \boldsymbol{\mathcal{P}}_\lambda(\mathbf{X})$, the state process $\{X_t\}_{t \ge 0}$ visits any λ-null set with probability zero. Therefore, we can give an alternative definition of the set $\mathbf{M}_i$ of stationary Markov policies for player i by letting $\mathbf{M}_i$ to be the family of stochastic kernels π_i on $\mathbf{A}_i$ given $\mathbf{X}$ satisfying

$$\pi_i(\mathbf{A}_i(x)|x) = 1 \quad \text{for λ-almost every } x \in \mathbf{X}. \tag{3.6}$$

This definition does not alter any of the properties of Markov policies seen so far.

Let $L^1(\mathbf{X}, \mathfrak{X}, \lambda)$ be the family of real-valued measurable functions which are λ-integrable (where, as usual, we identify functions which are equal λ-a.s.). When endowed with the $\|\cdot\|_1$-norm, and as a consequence of Assumption (A.1), we have that $L^1(\mathbf{X}, \mathfrak{X}, \lambda)$ becomes a separable Banach space.

On the set $\mathbf{M}_i$ we define the following equivalence relation. Given $\pi_i, \pi_i' \in \mathbf{M}_i$ we say that $\pi_i \approx \pi_i'$ when $\pi_i(\cdot|x) = \pi_i'(\cdot|x)$ for λ-almost every $x \in \mathbf{X}$. Let $\boldsymbol{\mathcal{Y}}_i$ be the family of equivalence classes of this relation. Each element of $\boldsymbol{\mathcal{Y}}_i$ is referred to as a Young measure. We equip the family of Young measures $\boldsymbol{\mathcal{Y}}_i$ with the narrow (stable) topology, which is the coarsest topology on $\boldsymbol{\mathcal{Y}}_i$ which makes the following mappings continuous:

$$\pi_i \mapsto \int_{\mathbf{X}} \int_{\mathbf{A}_i} f(x, a_i) \pi_i(da_i|x) \lambda(dx),$$

for any $f \in \mathcal{C}ar(\mathbf{X} \times \mathbf{A}_i, \mathbb{R})$ such that for some Φ in $L^1(\mathbf{X}, \mathfrak{X}, \lambda)$ we have $|f(x, a_i)| \leq \Phi(x)$ for every $(x, a_i) \in \mathbf{X} \times \mathbf{A}_i$; see [5, Theorem 2.2]. By [6, Lemma 1], the set $\boldsymbol{\mathcal{Y}}_i$ endowed with the narrow topology becomes a compact metric space.

Proposition 3.5 *Given any two pairs of stationary Markov policies* $(\pi_1, \pi_2), (\pi_1', \pi_2') \in \mathbf{M}_1 \times \mathbf{M}_2$ *and an initial distribution* $\eta \in \mathcal{P}(\mathbf{X})$*, we have the following results.*

(i). If $\pi_i \approx \pi_i'$ *(for* $i = 1, 2$*) and* $\eta \in \mathcal{P}_\lambda(\mathbf{X})$*, then* $\mu_{\eta, \pi_1, \pi_2} = \mu_{\eta, \pi_1', \pi_2'}$*.*

(ii). If $\mu_{\eta, \pi_1, \pi_2} = \mu_{\eta, \pi_1', \pi_2'}$ *and* $\eta \in \mathcal{P}_\lambda^e(\mathbf{X})$*, then* $\pi_i \approx \pi_i'$ *for each* $i = 1, 2$*.*

Proof. (i). Assume that $\eta \ll \lambda$ and $\pi_i \approx \pi_i'$ for $i = 1, 2$. Observe that $Q_{\pi_1, \pi_2}(\cdot|x) = Q_{\pi_1', \pi_2'}(\cdot|x)$ for λ-almost every $x \in \mathbf{X}$, and that $\eta Q_{\pi_1, \pi_2}^t \ll \lambda$ for every $t \geq 0$. Then, it can be easily shown by induction that $\eta Q_{\pi_1, \pi_2}^t = \eta Q_{\pi_1', \pi_2'}^t$ for every $t \geq 0$ implying —by (3.3)— that $\mu_{\eta, \pi_1, \pi_2}^{\mathbf{X}} = \mu_{\eta, \pi_1', \pi_2'}^{\mathbf{X}}$. Once we know that the $\mathbf{X}$-marginals of μ_{η, π_1, π_2} and $\mu_{\eta, \pi_1', \pi_2'}$ coincide, equality of the occupation measures follows from (3.4) because $\pi_i(\cdot|y)$ and $\pi_i'(\cdot|y)$ coincide for every $y \in \mathbf{X}$ on a set of λ-probability one, which is also a set of $\mu_{\eta, \pi_1, \pi_2}^{\mathbf{X}} = \mu_{\eta, \pi_1', \pi_2'}^{\mathbf{X}}$-probability one (recall Proposition 3.2(iii)).

(ii). Suppose now that $\eta \sim \lambda$ and $\mu_{\eta, \pi_1, \pi_2} = \mu_{\eta, \pi_1', \pi_2'}$. We have

$$\mu_{\eta, \pi_1, \pi_2}^{\mathbf{X}}(dy) \pi_i(da_i|y) = \mu_{\eta, \pi_1', \pi_2'}^{\mathbf{X} \times \mathbf{A}_i}(dy, da_i) = \mu_{\eta, \pi_1', \pi_2'}^{\mathbf{X}}(dy) \pi_i'(da_i|y) = \mu_{\eta, \pi_1, \pi_2}^{\mathbf{X}}(dy) \pi_i'(da_i|y),$$

for $i = 1, 2$. Since disintegration of $\mu_{\eta, \pi_1, \pi_2}^{\mathbf{X} \times \mathbf{A}_i}$ is unique up to sets of $\mu_{\eta, \pi_1, \pi_2}^{\mathbf{X}}$-probability zero and, hence, sets of λ-probability zero (recall Proposition 3.2(iii)), it follows that $\pi_i \approx \pi_i'$. $\square$

Proposition 3.5(i) above shows that, whenever $\eta \in \mathcal{P}_\lambda(\mathbf{X})$, stationary Markov policies in the same class of equivalence yield the same occupation measures. Therefore, in case that $\eta \in \mathcal{P}_\lambda(\mathbf{X})$, we will henceforth refer to the sets of Young measures $\boldsymbol{\mathcal{Y}}_1$ and $\boldsymbol{\mathcal{Y}}_2$ as to the stationary Markov policies of the players, with a slight abuse of terminology.

3.2 Some continuity facts

Let $L^\infty(\mathbf{X}, \mathfrak{B}(\mathbf{X}), \lambda)$ be the set of λ-essentially bounded measurable real-valued functions on $\mathbf{X}$ endowed with the weak* topology (we identify functions which are equal λ-a.s.). Let $\|v\|$ be the essential supremum of $v \in L^\infty(\mathbf{X}, \mathfrak{B}(\mathbf{X}), \lambda)$. The next result follows from Assumption (A.4).

Lemma 3.6 *If* $v \in L^\infty(\mathbf{X}, \mathfrak{B}(\mathbf{X}), \lambda)$ *then* $Qv \in \mathcal{C}ar_b(\mathbf{X} \times \mathbf{A}, \mathbb{R})$*.*

Lemma 3.7 *Suppose that $(\pi_{1,n}, \pi_{2,n}) \to (\pi_1, \pi_2)$ in $\boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ and $v_n \overset{*}{\to} v$ in $L^\infty(\mathbf{X}, \mathfrak{X}, \lambda)$ as $n \to \infty$. Under these conditions, for any $t \geq 0$ we have*

$$Q^t_{\pi_{1,n}, \pi_{2,n}} v_n \overset{*}{\to} Q^t_{\pi_1, \pi_2} v \quad as\ n \to \infty\ in\ L^\infty(\mathbf{X}, \mathfrak{X}, \lambda). \tag{3.7}$$

Proof. The result is trivial for $t = 0$. The case $t = 1$ can be shown by using similar arguments as in the proof of Lemma 4.1 in [17]. Once we know that $Q_{\pi_{1,n}, \pi_{2,n}} v_n \overset{*}{\to} Q_{\pi_1, \pi_2} v$, the stated result for any integer $t \geq 2$ follows easily. $\qquad\square$

Proposition 3.8 *The mappings from $\boldsymbol{\mathcal{P}}_\lambda(\mathbf{X}) \times \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ to $\boldsymbol{\mathcal{P}}(\mathbf{X} \times \mathbf{A}_i)$ defined by $(\eta, \pi_1, \pi_2) \mapsto \mu^{\mathbf{X} \times \mathbf{A}_i}_{\eta, \pi_1, \pi_2}$ are continuous for $i = 1, 2$ where $\boldsymbol{\mathcal{P}}_\lambda(\mathbf{X})$ is endowed with the metric of total variation and $\boldsymbol{\mathcal{P}}(\mathbf{X} \times \mathbf{A}_i)$ with the ws-topology. As a consequence, the $\mathbb{R}$- and $\mathbb{R}^p$-valued mappings*

$$(\eta, \pi_1, \pi_2) \mapsto R_i(\eta, \pi_1, \pi_2) \quad and \quad (\eta, \pi_1, \pi_2) \mapsto C_i(\eta, \pi_1, \pi_2)$$

are also continuous on $\boldsymbol{\mathcal{P}}_\lambda(\mathbf{X}) \times \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ for $i = 1, 2$.

Proof: Fix $i = 1, 2$. Since $\boldsymbol{\mathcal{P}}_\lambda(\mathbf{X}) \times \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ is a metric space, we will check continuity by proving sequential continuity. Consider a convergent sequence $(\eta_n, \pi_{1,n}, \pi_{2,n}) \to (\eta, \pi_1, \pi_2)$ in the product space $\boldsymbol{\mathcal{P}}_\lambda(\mathbf{X}) \times \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ and $f \in \mathcal{C}ar_b(\mathbf{X} \times \mathbf{A}_i, \mathbb{R})$. Let us denote by v_n and v the bounded measurable functions on $\mathbf{X}$ defined by $v_n(x) = \int_{\mathbf{A}_i} f(x, a_i) \pi_{i,n}(da_i | x)$ and $v(x) = \int_{\mathbf{A}_i} f(x, a_i) \pi_i(da_i | x)$. Since $\pi_{i,n} \to \pi_i$ in $\boldsymbol{\mathcal{Y}}_i$ then $v_n \overset{*}{\to} v$ in $L^\infty(\mathbf{X}, \mathfrak{X}, \lambda)$. We can now apply Lemma 3.7 to conclude that for any $t \geq 0$ we have $Q^t_{\pi_{1,n}, \pi_{2,n}} v_n \overset{*}{\to} Q^t_{\pi_1, \pi_2} v$ in $L^\infty(\mathbf{X}, \mathfrak{B}(\mathbf{X}), \lambda)$. Moreover, by hypothesis $\eta_n \to \eta$ in total variation, and so it is easily seen that $d\eta_n/d\lambda \to d\eta/d\lambda$ in $L^1(\mathbf{X}, \mathfrak{X}, \lambda)$. Therefore,

$$
\begin{aligned}
\lim_{n \to \infty} \int_{\mathbf{X}} Q^t_{\pi_{1,n}, \pi_{2,n}} v_n(x) \eta_n(dx) &= \lim_{n \to \infty} \int_{\mathbf{X}} Q^t_{\pi_{1,n}, \pi_{2,n}} v_n(x) \frac{d\eta_n}{d\lambda}(x) \lambda(dx) \\
&= \int_{\mathbf{X}} Q^t_{\pi_1, \pi_2} v(x) \frac{d\eta}{d\lambda}(x) \lambda(dx) = \int_{\mathbf{X}} Q^t_{\pi_1, \pi_2} v(x) \eta(dx),
\end{aligned}
$$

by using Proposition 3.13(iv) in [8]. Therefore, since f is bounded, by dominated convergence

$$
\begin{aligned}
\lim_{n \to \infty} \int_{\mathbf{X} \times \mathbf{A}_i} f(x, a_i) \mu^{\mathbf{X} \times \mathbf{A}_i}_{\eta_n \pi_{1,n}, \pi_{2,n}}(dx, da_i) &= (1 - \beta) \sum_{t=0}^{\infty} \beta^t \lim_{n \to \infty} \int_{\mathbf{X}} Q^t_{\pi_{1,n}, \pi_{2,n}} v_n(x) \eta_n(dx) \\
&= (1 - \beta) \sum_{t=0}^{\infty} \beta^t \int_{\mathbf{X}} Q^t_{\pi_1, \pi_2} v(x) \eta(dx) = \int_{\mathbf{X} \times \mathbf{A}_i} f(x, a_i) \mu^{\mathbf{X} \times \mathbf{A}_i}_{\eta, \pi_1, \pi_2}(dx, da_i),
\end{aligned}
$$

which establishes the continuity of $(\eta, \pi_1, \pi_2) \mapsto \mu^{\mathbf{X} \times \mathbf{A}_i}_{\eta, \pi_1, \pi_2}$. Continuity of the reward and constraint functions follows because they are defined by integration of bounded Carathéodory functions. $\quad\square$

3.3 The sets $\mathcal{D}_\eta$ and $\mathcal{D}_{\eta, i}$

Based on the result of Proposition 3.2(i), given an initial distribution $\eta \in \boldsymbol{\mathcal{P}}(\mathbf{X})$ we define

$$\mathcal{D}_\eta = \left\{ \mu \in \boldsymbol{\mathcal{P}}(\mathbf{X} \times \mathbf{A}) : \mu(\mathbf{K}) = 1 \text{ and } \mu^{\mathbf{X}} = (1 - \beta)\eta + \beta \mu Q \right\}.$$

The set of occupation measures satisfies $\boldsymbol{\mathcal{O}}_\eta \subseteq \mathcal{D}_\eta$. Note that any $\mu \in \mathcal{D}_\eta$ can be disintegrated as

$$\mu(dx, da_1, da_2) = \pi(da_1, da_2 | x) \mu^{\mathbf{X}}(dx) \tag{3.8}$$

for some stochastic kernel π on $\mathbf{A}$ given $\mathbf{X}$ which satisfies $\pi(\mathbf{A}(x)|x) = 1$ for all $x \in \mathbf{X}$, although this π might not correspond to a pair $(\pi_1, \pi_2) \in \mathbf{M}_1 \times \mathbf{M}_2$ of stationary Markov policies of the players. The stochastic kernel Q_π on $\mathbf{X}$ given $\mathbf{X}$ is then defined as $Q_\pi(dy|x) = \int_\mathbf{A} Q(dy|x,a)\pi(da|x)$ for $x \in \mathbf{X}$, and Q_π^t for $t \geq 0$ denotes the t-th composition of Q_π, with $Q_\pi^0(dy|x) = \delta_x(dy)$. Clearly,

$$\mu^\mathbf{X} = (1 - \beta) \sum_{t=0}^\infty \beta^t \eta Q_\pi^t. \tag{3.9}$$

For $\eta \in \boldsymbol{\mathcal{P}}(\mathbf{X})$, we define $\mathcal{D}_{\eta,i}$ as the set of $(\mathbf{X} \times \mathbf{A}_i)$-marginals of the measures in $\mathcal{D}_\eta$ for $i = 1, 2$:

$$\mathcal{D}_{\eta,i} = \{\mu^{\mathbf{X} \times \mathbf{A}_i} : \mu \in \mathcal{D}_\eta\} \subseteq \boldsymbol{\mathcal{P}}(\mathbf{X} \times \mathbf{A}_i).$$

In our next result we establish compactness of the above defined sets. Such result is known in the literature when $\mathbf{X}$ is a Borel space whereas, in this paper, $\mathbf{X}$ is an abstract measurable space.

Proposition 3.9 *Given an initial distribution $\eta \in \boldsymbol{\mathcal{P}}(\mathbf{X})$, the sets $\mathcal{D}_\eta$ and $\mathcal{D}_{\eta,i}$ are convex compact metric spaces when endowed with their respective ws-topologies.*

Proof. Let us first show that $\mathcal{D}_\eta$ is relatively compact in $\boldsymbol{\mathcal{P}}(\mathbf{X} \times \mathbf{A})$. To this end, consider a decreasing sequence $\{h_j\}_{j \in \mathbb{N}}$ of functions in $\mathcal{C}ar_b(\mathbf{X} \times \mathbf{A}, \mathbb{R})$ such that $h_j(x, a_1, a_2) \downarrow 0$ for any $(x, a_1, a_2) \in \mathbf{X} \times \mathbf{A}$. By Proposition 2.1, to prove relative compactness we need to show that

$$\lim_{j \to \infty} \sup_{\mu \in \mathcal{D}_\eta} \int_{\mathbf{X} \times \mathbf{A}} h_j d\mu = 0. \tag{3.10}$$

According to (3.8) and (3.9), for any $\mu \in \mathcal{D}_\eta$ there exists a stochastic kernel π_μ on $\mathbf{A}$ given $\mathbf{X}$ satisfying $\pi_\mu(\mathbf{A}(x)|x) = 1$ for all $x \in \mathbf{X}$ such that

$$\int_\mathbf{K} h_j(x, a)\mu(dx, da) \leq \int_\mathbf{X} g_j^\mu(x)\eta(dx) \quad \text{where} \quad g_j^\mu(x) = (1 - \beta) \sum_{t=0}^\infty \beta^t Q_{\pi_\mu}^t f_j(x) \tag{3.11}$$

with $f_j(x) = \sup_{a \in \mathbf{A}} h_j(x, a)$. Combining Lemma 10.1 and Theorem 12.1 in [26], we obtain that $\{f_j\}_{j \in \mathbb{N}}$ is a decreasing sequence of bounded measurable functions defined on $\mathbf{X}$ which satisfies $\lim_{j \to \infty} f_j(x) = \sup_{a \in \mathbf{A}} \lim_{j \to \infty} h_j(x, a) = 0$ for any $x \in \mathbf{X}$. Define the functions $\bar{f}_{j,t}$ for $j, t \in \mathbb{N}$ on $\mathbf{X}$ by

$$\bar{f}_{j,t}(x) = \sup_{a \in \mathbf{A}} Q\bar{f}_{j,t-1}(x, a) \quad \text{and} \quad \bar{f}_{j,0}(x) = f_j(x)$$

for $x \in \mathbf{X}$, $t \geq 1$, and $j \in \mathbb{N}$. Observe that the functions $\{\bar{f}_{j,t}\}_{j,t \in \mathbb{N}}$ do not depend on $\mu \in \mathcal{D}_\eta$. Clearly, we have $(Q_{\pi_\mu})^t f_j(x) \leq \bar{f}_{j,t}(x)$ for any $j, t \in \mathbb{N}$. From (3.11), this implies for any $x \in \mathbf{X}$ and $j \in \mathbb{N}$ that $g_j^\mu(x) \leq (1 - \beta) \sum_{t=0}^\infty \beta^t \bar{f}_{j,t}(x)$. It follows that

$$\sup_{\mu \in \mathcal{D}} \int_{\mathbf{X} \times \mathbf{A}} h_j(x, a)\mu(dx, da) \leq (1 - \beta) \sum_{k=0}^\infty \beta^k \int_\mathbf{X} \bar{f}_{j,k}(x)\eta(dx). \tag{3.12}$$

From Lemma 10.1 and Theorem 12.1 in [26], and recalling Assumption (A.4), it can be shown by induction that for any $k \in \mathbb{N}$, $\{\bar{f}_{j,k}\}_{j \in \mathbb{N}}$ is a decreasing sequence of bounded measurable functions satisfying $\lim_{j \to \infty} \bar{f}_{j,k}(x) = 0$ for any $x \in \mathbf{X}$ and $k \in \mathbb{N}$. Therefore, taking the limit in (3.12) and by using the monotone convergence theorem, we get (3.10) and thus $\mathcal{D}_\eta$ is indeed relatively compact.

Once we know that $\mathcal{D}_\eta$ is relatively compact we use [7, Theorem 5.2] and the fact that $\mathbf{A}$ is a compact metric space to conclude that $\mathcal{D}_\eta^{\mathbf{X}} = \{\mu^{\mathbf{X}} : \mu \in \mathcal{D}_\eta\}$ is relatively compact in the s-topology of $\mathcal{P}(\mathbf{X})$. Since $\mathfrak{X}$ is countably generated, we deduce from [7, Proposition 2.3] that $\mathcal{D}_\eta$ is metrizable.

To conclude the proof of compactness, it remains to show that $\mathcal{D}_\eta$ is closed. To this end, let $\{\mu_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathcal{D}_\eta$ that converges in the ws-topology to $\mu \in \mathcal{P}(\mathbf{X}\times\mathbf{A})$. For any bounded measurable function $f : \mathbf{X} \to \mathbb{R}$ we have

$$\int_{\mathbf{X}} f d\mu_n^{\mathbf{X}} = \int_{\mathbf{X}\times\mathbf{A}} f d\mu_n = (1-\beta)\int_{\mathbf{X}} f d\eta + \beta \int_{\mathbf{X}\times\mathbf{A}} Qf(x, a_1, a_2)\mu_n(dx, da_1, da_2),$$

where we still use the notation f for the function $(x, a_1, a_2) \mapsto f(x)$ which is in $Car_b(\mathbf{X} \times \mathbf{A}, \mathbb{R})$. By Lemma 3.6, we can take the limit as $n \to \infty$ in order to obtain

$$\int_{\mathbf{X}} f d\mu^{\mathbf{X}} = (1-\beta)\int_{\mathbf{X}} f d\eta + \beta \int_{\mathbf{X}\times\mathbf{A}} Qf(x, a_1, a_2)\mu(dx, da_1, da_2)$$

and, in particular, $\mu^{\mathbf{X}} = (1-\beta)\eta + \beta\mu Q$. To show that $\mu \in \mathcal{D}_\eta$ it remains to prove that $\mu(\mathbf{K}) = 1$. The mapping $(x, a_1, a_2) \mapsto -\mathbf{I}_{\mathbf{K}}(x, a_1, a_2)$ is measurable on $\mathbf{X} \times \mathbf{A}$ and it is such that $(a_1, a_2) \mapsto -\mathbf{I}_{\mathbf{K}}(x, a_1, a_2) = -\mathbf{I}_{\mathbf{A}(x)}(a_1, a_2)$ is lower semicontinuous for any fixed $x \in \mathbf{X}$ because $\mathbf{A}(x)$ is compact. Thus, $-\mathbf{I}_{\mathbf{K}}$ is a normal integrand [7, p. 502] and so by [7, Theorem 3.1.(c)] we have $\overline{\lim}_n \mu_n(\mathbf{K}) \le \mu(\mathbf{K})$, implying that $\mu(\mathbf{K}) = 1$. This concludes the proof that $\mu \in \mathcal{D}_\eta$. Finally, observe that convexity of $\mathcal{D}_\eta$ is a straightforward consequence of its definition.

It is easy to check that the mapping from $\mathcal{P}(\mathbf{X} \times \mathbf{A})$ to $\mathcal{P}(\mathbf{X} \times \mathbf{A}_i)$ that associates to each $\mu \in \mathcal{P}(\mathbf{X} \times \mathbf{A})$ its marginal probability measure $\mu^{\mathbf{X}\times\mathbf{A}_i}$ is continuous for the ws-topologies. Hence $\mathcal{D}_{\eta,i}$ is compact in $\mathcal{P}(\mathbf{X} \times \mathbf{A}_i)$. Recalling that $\mathfrak{X}$ is countably generated and noting that the set of $\mathbf{X}$-marginal probability measures of $\mathcal{D}_{\eta,i}$ is precisely $\mathcal{D}_\eta^{\mathbf{X}}$, which has already been shown to be relatively compact for the s-topology, we get from [7, Proposition 2.3] that $\mathcal{D}_{\eta,i}$ is metrizable.

Finally, regarding convexity, let $\hat{\gamma}$ and $\bar{\gamma}$ in $\mathcal{D}_{\eta,i}$ and fix $0 \le \alpha \le 1$. There exist $\hat{\mu}$ and $\bar{\mu}$ in $\mathcal{D}_\eta$ such that $\hat{\gamma} = \hat{\mu}^{\mathbf{X}\times\mathbf{A}_i}$ and $\bar{\gamma} = \bar{\mu}^{\mathbf{X}\times\mathbf{A}_i}$. By convexity of $\mathcal{D}_\eta$ we have $\mu = \alpha\hat{\mu} + (1-\alpha)\bar{\mu} \in \mathcal{D}_\eta$ and so $\alpha\hat{\gamma} + (1-\alpha)\bar{\gamma} = \mu^{\mathbf{X}\times\mathbf{A}_i}$ is in $\mathcal{D}_{\eta,i}$, showing convexity. $\qquad\square$

Proposition 3.10 *Consider an initial distribution $\eta \in \mathcal{P}_\lambda^e(\mathbf{X})$.*

(i). For any $\mu \in \mathcal{D}_\eta$ then there exists a unique $(\pi_1, \pi_2) \in \mathcal{Y}_1 \times \mathcal{Y}_2$ satisfying

$$\mu^{\mathbf{X}\times\mathbf{A}_1} = \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}\times\mathbf{A}_1} \quad and \quad \mu^{\mathbf{X}\times\mathbf{A}_2} = \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}\times\mathbf{A}_2}. \tag{3.13}$$

(ii). We have $\mathcal{D}_{\eta,i} = \{\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}\times\mathbf{A}_i} : (\pi_1, \pi_2) \in \mathcal{Y}_1 \times \mathcal{Y}_2\}$ for each $i = 1, 2$.

(iii). If (π_1, π_2) and (π_1', π_2') in $\mathcal{Y}_1 \times \mathcal{Y}_2$ are such that $\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}\times\mathbf{A}_1} = \mu_{\eta,\pi_1',\pi_2'}^{\mathbf{X}\times\mathbf{A}_1}$ then $\pi_1 = \pi_1'$. A symmetric result holds for the marginal probability measures on $\mathbf{X} \times \mathbf{A}_2$.

Proof. (i). Since $\mu(\mathbf{K}) = 1$ we have $\mu^{\mathbf{X}\times\mathbf{A}_1}(\mathbf{K}_1) = 1$ and $\mu^{\mathbf{X}\times\mathbf{A}_2}(\mathbf{K}_2) = 1$. By Lemma 2.3 there exist $(\pi_1, \pi_2) \in \mathbf{M}_1 \times \mathbf{M}_2$ such that $\mu^{\mathbf{X}\times\mathbf{A}_1} = \mu^{\mathbf{X}} \otimes \pi_1$ and $\mu^{\mathbf{X}\times\mathbf{A}_2} = \mu^{\mathbf{X}} \otimes \pi_2$. Now, μ satisfies

$$\mu^{\mathbf{X}} = (1-\beta)\eta + \beta\mu Q = (1-\beta)\eta + \beta\mu^{\mathbf{X}} Q_{\pi_1,\pi_2}$$

and so $\mu^{\mathbf{X}} = \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}}$. The above stochastic kernels π_i on $\mathbf{A}_i$ given $\mathbf{X}$ are unique $\mu^{\mathbf{X}}$-almost surely. Since $\mu^{\mathbf{X}} = \mu_{\pi_1,\pi_2}^{\mathbf{X}} \sim \lambda$ (Proposition 3.2(iii)), it follows that uniqueness is λ-almost surely. This shows that there is indeed a unique pair $(\pi_1, \pi_2) \in \mathcal{Y}_1 \times \mathcal{Y}_2$ with the above mentioned properties.

(ii). This result directly follows from part (i) and the definition of $\mathcal{D}_{\eta,i}$.

(iii). By using arguments similar to those in (i) above, the result readily follows. $\qquad\square$

4 Main results

In this section we address our results on the existence of Nash equilibria. First of all, in Section 4.1, we will prove the existence of such equilibria for a game model with initial distribution in $\mathcal{P}_\lambda^e(\mathbf{X})$, while in Section 4.2 we will treat the general case a game model $\mathcal{G}(\nu, \theta)$ satisfying Assumption A.

4.1 The case of an initial distribution $\eta \in \mathcal{P}_\lambda^e(\mathbf{X})$

Suppose that the initial distribution of the system $\eta \in \mathcal{P}_\lambda^e(\mathbf{X})$ and that the constraint constants $\rho = (\rho_1, \rho_2) \in \mathbb{R}^p \times \mathbb{R}^p$ that satisfy the Slater condition. In order to prove the existence of a constrained Nash equilibrium for $\mathcal{G}(\eta, \rho)$ we will suitably define a correspondence $\mathcal{D}_{\eta,1} \times \mathcal{D}_{\eta,2} \twoheadrightarrow \mathcal{D}_{\eta,1} \times \mathcal{D}_{\eta,2}$ which will be shown to have a fixed point, from which we will derive equilibrium stationary Markov policies. In fact, we shall construct this correspondence in two steps: as the composition $\mathcal{H}_{\eta,\rho} \circ \mathcal{J}_\eta$ of a function $\mathcal{J}_\eta : \mathcal{D}_{\eta,1} \times \mathcal{D}_{\eta,2} \to \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ and a correspondence $\mathcal{H}_{\eta,\rho} : \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2 \twoheadrightarrow \mathcal{D}_{\eta,1} \times \mathcal{D}_{\eta,2}$.

The function $\mathcal{J}_\eta$. Given a player $i = 1, 2$ we define the function $\mathcal{J}_{\eta,i} : \mathcal{D}_{\eta,i} \to \boldsymbol{\mathcal{Y}}_i$ as follows. For any $\gamma \in \mathcal{D}_{\eta,i}$, let $\pi_i = \mathcal{J}_{\eta,i}(\gamma)$ be the unique $\pi_i \in \boldsymbol{\mathcal{Y}}_i$ satisfying $\gamma = \gamma^{\mathbf{X}} \otimes \pi_i$. The existence and uniqueness of such decomposition is guaranteed by (ii) and (iii) in Proposition 3.10.

Proposition 4.1 *Given $\eta \in \mathcal{P}_\lambda^e(\mathbf{X})$, the functions $\mathcal{J}_{\eta,i} : \mathcal{D}_{\eta,i} \to \boldsymbol{\mathcal{Y}}_i$ are continuous for $i = 1, 2$.*

Proof: For simplicity in the notation, we prove the case $i = 1$. Consider a sequence $\{\gamma_n\}_{n \in \mathbb{N}}$ in $\mathcal{D}_{\eta,1}$ such that $\gamma_n \to \gamma$ for some $\gamma \in \mathcal{D}_{\eta,1}$. Using Proposition 3.10(ii) it follows that for each $n \in \mathbb{N}$ there exist $(\pi_{1,n}, \pi_{2,n}) \in \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ satisfying $\gamma_n = \mu_{\eta,\pi_{1,n},\pi_{2,n}}^{\mathbf{X} \times \mathbf{A}_1}$ and there exist $(\pi_1, \pi_2) \in \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ such that $\gamma = \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X} \times \mathbf{A}_1}$. We have, by definition, $\pi_{1,n} = \mathcal{J}_{\eta,1}(\gamma_n)$ and $\pi_1 = \mathcal{J}_{\eta,1}(\gamma)$. Our goal is to show that $\pi_{1,n} \to \pi_1$. Since $\boldsymbol{\mathcal{Y}}_1$ is a compact metric space, to prove the result it suffices to consider an arbitrary convergent subsequence $\{\pi_{1,n'}\}$ of $\{\pi_{1,n}\}$ with $\pi_{1,n'} \to \pi_1' \in \boldsymbol{\mathcal{Y}}_1$ and to show that, necessarily, $\pi_1' = \pi_1$. But $\boldsymbol{\mathcal{Y}}_2$ being also a compact metric space, there exists a further subsequence —which, without loss of generality, we will denote also by n'— such that $\pi_{2,n'} \to \pi_2'$ for some $\pi_2' \in \boldsymbol{\mathcal{Y}}_2$. By Proposition 3.8 we have

$$\gamma_{n'} = \mu_{\eta,\pi_{1,n'},\pi_{2,n'}}^{\mathbf{X} \times \mathbf{A}_1} \to \mu_{\eta,\pi_1',\pi_2'}^{\mathbf{X} \times \mathbf{A}_1} = \gamma.$$

Recalling Proposition 3.10(iii), this implies that $\pi_1 = \pi_1' = \mathcal{J}_{\eta,1}(\gamma)$, and the proof is complete. $\square$

Based on this result, we can now write $\mathcal{J}_\eta = (\mathcal{J}_{\eta,1}, \mathcal{J}_{\eta,2})$ which is therefore a continuous function from $\mathcal{D}_{\eta,1} \times \mathcal{D}_{\eta,2}$ to $\boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ for any $\eta \in \mathcal{P}_\lambda^e(\mathbf{X})$.

The correspondence $\mathcal{H}_{\eta,\rho}$. Recall that we are considering $\eta \in \mathcal{P}_\lambda^e(\mathbf{X})$. Given any $\pi_2 \in \boldsymbol{\mathcal{Y}}_2$ let

$$\mathcal{L}_{\eta,1}(\pi_2) = \left\{ \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X} \times \mathbf{A}_1} : \pi_1 \in \boldsymbol{\mathcal{Y}}_1 \right\} \subseteq \mathcal{D}_{\eta,1} \subseteq \mathcal{P}(\mathbf{X} \times \mathbf{A}_1),$$

which is the set of $(\mathbf{X} \times \mathbf{A}_1)$-marginals of the occupation measures for the initial distribution η and pairs of stationary Markov policies $(\pi_1, \pi_2) \in \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ when the policy $\pi_2 \in \boldsymbol{\mathcal{Y}}_2$ of player 2 remains fixed and the policy $\pi_1 \in \boldsymbol{\mathcal{Y}}_1$ of player 1 varies. Similarly, for any $\pi \in \boldsymbol{\mathcal{Y}}_1$ we define

$$\mathcal{L}_{\eta,2}(\pi_1) = \left\{ \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X} \times \mathbf{A}_2} : \pi_2 \in \boldsymbol{\mathcal{Y}}_2 \right\} \subseteq \mathcal{D}_{\eta,2} \subseteq \mathcal{P}(\mathbf{X} \times \mathbf{A}_2).$$

Proposition 4.2 *Given an initial distribution $\eta \in \mathcal{P}_\lambda^e(\mathbf{X})$ and any $(\pi_1, \pi_2) \in \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$, the sets $\mathcal{L}_{\eta,1}(\pi_2)$ and $\mathcal{L}_{\eta,2}(\pi_1)$ are convex and compact when endowed with their respective ws-topologies.*

14

Proof. We only make the proof for $\mathcal{L}_{\eta,1}(\pi_2)$. To prove convexity, we fix two measures $\tilde{\gamma}, \bar{\gamma}$ in $\mathcal{L}_{\eta,1}(\pi_2)$ and some $0 \leq \alpha \leq 1$. We want to prove that $\alpha\tilde{\gamma} + (1-\alpha)\bar{\gamma} \in \mathcal{L}_{\eta,1}(\pi_2)$. By definition of $\mathcal{L}_{\eta,1}(\pi_2)$, there exist $\tilde{\pi}_1, \bar{\pi}_1 \in \boldsymbol{\mathcal{Y}}_1$ such that $\tilde{\gamma} = \mu_{\eta,\tilde{\pi}_1,\pi_2}^{\mathbf{X}\times\mathbf{A}_1}$ and $\bar{\gamma} = \mu_{\eta,\bar{\pi}_1,\pi_2}^{\mathbf{X}\times\mathbf{A}_1}$. By convexity of $\mathcal{D}_\eta$ (recall Proposition 3.9), we have that μ defined by $\mu = \alpha\mu_{\eta,\tilde{\pi}_1,\pi_2} + (1-\alpha)\mu_{\eta,\bar{\pi}_1,\pi_2}$ lies in $\mathcal{D}_\eta$. Therefore, by Proposition 3.10(i), there exists a unique $(\pi_1^*, \pi_2^*) \in \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ such that $\mu^{\mathbf{X}\times\mathbf{A}_1} = \mu_{\eta,\pi_1^*,\pi_2^*}^{\mathbf{X}\times\mathbf{A}_1}$ and $\mu^{\mathbf{X}\times\mathbf{A}_2} = \mu_{\eta,\pi_1^*,\pi_2^*}^{\mathbf{X}\times\mathbf{A}_2}$. Clearly, we have

$$\alpha\tilde{\gamma} + (1-\alpha)\bar{\gamma} = \mu_{\eta,\pi_1^*,\pi_2^*}^{\mathbf{X}\times\mathbf{A}_1}. \tag{4.1}$$

To get convexity, we need to show that $\pi_2^* = \pi_2$. Observe that

$$\alpha\tilde{\gamma} + (1-\alpha)\bar{\gamma} = \left[\alpha\mu_{\eta,\tilde{\pi}_1,\pi_2}^{\mathbf{X}} + (1-\alpha)\mu_{\eta,\bar{\pi}_1,\pi_2}^{\mathbf{X}}\right] \otimes \pi_2 = \mu^{\mathbf{X}} \otimes \pi_2 = \mu_{\eta,\pi_1^*,\pi_2^*}^{\mathbf{X}} \otimes \pi_2 \tag{4.2}$$

$$\mu_{\eta,\pi_1^*,\pi_2^*}^{\mathbf{X}\times\mathbf{A}_1} = \mu_{\eta,\pi_1^*,\pi_2^*}^{\mathbf{X}} \otimes \pi_2^*. \tag{4.3}$$

However, $\mu_{\eta,\pi_1^*,\pi_2^*}^{\mathbf{X}} \sim \lambda$ from Proposition 3.2(iii) and so, by uniqueness of the disintegration, we deduce from (4.1)–(4.3) that $\pi_2 = \pi_2^*$. Hence, $\mathcal{L}_{\eta,1}(\pi_2)$ is convex.

Since $\mathcal{L}_{1,\eta}(\pi_2) \subseteq \mathcal{D}_{1,\eta}$, which is a compact metric space, in order to prove compactness of $\mathcal{L}_{\eta,1}(\pi_2)$ is suffices to show that it is closed. To this end, consider a sequence $\{\gamma_n\}$ in $\mathcal{L}_{\eta,1}(\pi_2)$ converging to some $\gamma \in \mathcal{D}_{\eta,1}$. We have $\gamma_n = \mu_{\eta,\pi_{1,n},\pi_2}^{\mathbf{X}\times\mathbf{A}_1}$ for some $\pi_{1,n}$. For some subsequence n' and some $\pi_1 \in \boldsymbol{\mathcal{Y}}_1$ we have $\pi_{1,n'} \to \pi_1$ (recall that $\boldsymbol{\mathcal{Y}}_1$ is compact) and, by Proposition 3.8, this implies

$$\gamma_{n'} = \mu_{\eta,\pi_{1,n'},\pi_2}^{\mathbf{X}\times\mathbf{A}_1} \to \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}\times\mathbf{A}_1} = \gamma.$$

This shows that $\gamma \in \mathcal{L}_{\eta,1}(\pi_2)$ and the proof that $\mathcal{L}_{\eta,1}(\pi_2)$ is compact is complete. $\square$

We explain how the payoff of the players for the policies $(\pi_1, \pi_2) \in \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ relates to $\mathcal{L}_1(\pi_2)$.

Remark 4.3 *Recalling (3.1) observe that for an initial distribution $\eta \in \mathcal{P}_\lambda^e(\mathbf{X})$, when player 2 fixes a policy $\pi_2 \in \boldsymbol{\mathcal{Y}}_2$, the family of total expected discounted payoffs $R_1(\eta, \cdot, \pi_2)$ when player 1 uses stationary Markov policies in $\boldsymbol{\mathcal{Y}}_1$ is given by*

$$\int_{\mathbf{X}\times\mathbf{A}_1} r_1^1 d\gamma + \int_{\mathbf{X}\times\mathbf{A}_2} r_1^2 d(\gamma^{\mathbf{X}} \otimes \pi_2) \tag{4.4}$$

when γ varies in $\mathcal{L}_{\eta,1}(\pi_2)$. The important fact is that (4.4) is linear in γ. The same result holds for the constraint $C_1(\eta, \cdot, \pi_2)$ and, symmetrically, for the payoffs $R_2(\eta, \pi_1, \cdot)$ and $C_2(\eta, \pi_1, \cdot)$ of player 2 when player 1 fixes his policy $\pi_1 \in \boldsymbol{\mathcal{Y}}_1$.

Based on this remark, we define the following sets. Given the initial distribution $\eta \in \mathcal{P}_\lambda^e(\mathbf{X})$, the constraint constant $\rho_1 \in \mathbb{R}^p$, and a stationary Markov policy $\pi_2 \in \boldsymbol{\mathcal{Y}}_2$, let $\mathcal{A}_{\eta,\rho_1,1}(\pi_2)$ be the set of $(\mathbf{X} \times \mathbf{A}_1)$-marginals of the occupation measures induced by the policies $(\pi_1, \pi_2) \in \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ —as π_1 varies and π_2 remains fixed— which satisfy the constraint of player 1, that is,

$$\mathcal{A}_{\eta,\rho_1,1}(\pi_2) = \left\{\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}\times\mathbf{A}_1} : \text{for } \pi_1 \in \boldsymbol{\mathcal{Y}}_1 \text{ such that } C_1(\eta, \pi_1, \pi_2) \geq \rho_1\right\}$$

$$= \left\{\gamma \in \mathcal{L}_{\eta,1}(\pi_2) : \int_{\mathbf{X}\times\mathbf{A}_1} c_1^1 d\gamma + \int_{\mathbf{X}\times\mathbf{A}_2} c_1^2 d(\gamma^{\mathbf{X}} \otimes \pi_2) \geq \rho_1\right\} \subseteq \mathcal{D}_{\eta,1}. \tag{4.5}$$

Similarly, for $\eta \in \mathcal{P}_\lambda^e(\mathbf{X})$, $\rho_2 \in \mathbb{R}^p$, and $\pi_1 \in \boldsymbol{\mathcal{Y}}_1$ we define

$$\mathcal{A}_{\eta,\rho_2,2}(\pi_1) = \left\{\mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}\times\mathbf{A}_2} : \text{for } \pi_2 \in \boldsymbol{\mathcal{Y}}_2 \text{ such that } C_2(\eta, \pi_1, \pi_2) \geq \rho_2\right\}$$

$$= \left\{\gamma \in \mathcal{L}_{\eta,2}(\pi_1) : \int_{\mathbf{X}\times\mathbf{A}_1} c_2^1 d(\gamma^{\mathbf{X}} \otimes \pi_1) + \int_{\mathbf{X}\times\mathbf{A}_2} c_2^2 d\gamma \geq \rho_2\right\} \subseteq \mathcal{D}_{\eta,2},$$

15

which is the set of $(\mathbf{X} \times \mathbf{A}_2)$-marginals of the occupation measures of the policies $(\pi_1, \pi_2) \in \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ —as π_1 is fixed and π_2 varies— which satisfy the constraint of player 2. We will need the following result in which we will consider an initial distribution $\zeta \in \boldsymbol{\mathcal{P}}_\lambda(\mathbf{X})$, but not necessarily in $\boldsymbol{\mathcal{P}}_\lambda^e(\mathbf{X})$.

Proposition 4.4 *Consider an initial distribution $\zeta \in \boldsymbol{\mathcal{P}}_\lambda(\mathbf{X})$ and constraint constants $(\rho_1, \rho_2) \in \mathbb{R}^p \times \mathbb{R}^p$ that satisfy the Slater condition, and let $\{\eta_n\}_{n \in \mathbb{N}}$ be an arbitrary sequence in $\boldsymbol{\mathcal{P}}_\lambda^e(\mathbf{X})$ converging to $\zeta \in \boldsymbol{\mathcal{P}}_\lambda(\mathbf{X})$ in total variation. Then the following assertions hold.*

(i). Fix an arbitrary pair $(\pi_1, \pi_2) \in \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ such that $C_1(\zeta, \pi_1, \pi_2) \geq \rho_1$. For any sequence $\{\pi_{2,n}\}_{n \in \mathbb{N}} \subseteq \boldsymbol{\mathcal{Y}}_2$ converging to π_2, there exists a sequence $\{\gamma_n\}_{n \in \mathbb{N}}$ in $\boldsymbol{\mathcal{P}}(\mathbf{X} \times \mathbf{A}_1)$ such that $\gamma_n \to \mu_{\zeta, \pi_1, \pi_2}^{\mathbf{X} \times \mathbf{A}_1}$ and such that, for some $N \in \mathbb{N}$, we have $\gamma_n \in \mathcal{A}_{\eta_n, \rho_1, 1}(\pi_{2,n})$ for every $n \geq N$.

(ii). Fix an arbitrary pair $(\pi_1, \pi_2) \in \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ such that $C_2(\zeta, \pi_1, \pi_2) \geq \rho_2$. For any sequence $\{\pi_{1,n}\}_{n \in \mathbb{N}} \subseteq \boldsymbol{\mathcal{Y}}_1$ converging to π_1, there exists a sequence $\{\gamma_n\}_{n \in \mathbb{N}}$ in $\boldsymbol{\mathcal{P}}(\mathbf{X} \times \mathbf{A}_2)$ such that $\gamma_n \to \mu_{\zeta, \pi_1, \pi_2}^{\mathbf{X} \times \mathbf{A}_2}$ and such that, for some $N \in \mathbb{N}$, we have $\gamma_n \in \mathcal{A}_{\eta_n, \rho_2, 2}(\pi_{1,n})$ for every $n \geq N$.

Proof: We prove (i). From Proposition 3.8 we have $\lim_{n \to \infty} C_1(\eta_n, \pi_1, \pi_{2,n}) = C_1(\zeta, \pi_1, \pi_2) \geq \rho_1$. Therefore, there exist a sequence $\{\epsilon_n\}_{n \in \mathbb{N}}$ taking values in the interval $(0, 1)$ with $\epsilon_n \to 0$ and some index n_0 for which

$$C_1(\eta_n, \pi_1, \pi_{2,n}) \geq \rho_1 - \epsilon_n \quad \text{for all } n \geq n_0. \tag{4.6}$$

On the other hand, since ζ and (ρ_1, ρ_2) satisfy the Slater condition, we can find $\bar{\pi}_1 \in \boldsymbol{\mathcal{Y}}_1$ and $\delta > 0$ such that $C_1(\zeta, \bar{\pi}_1, \pi_2) > \rho_1 + \delta$. So, again from Proposition 3.8, there is some $n_1 \geq n_0$ such that

$$C_1(\eta_n, \bar{\pi}_1, \pi_{2,n}) > \rho_1 + \delta \quad \text{for all } n \geq n_1. \tag{4.7}$$

Observe that for any $n \in \mathbb{N}$ both $\mu_{\eta_n, \pi_1, \pi_{2,n}}^{\mathbf{X} \times \mathbf{A}_1}$ and $\mu_{\eta_n, \bar{\pi}_1, \pi_{2,n}}^{\mathbf{X} \times \mathbf{A}_1}$ are in $\mathcal{L}_{\eta_n, 1}(\pi_{2,n})$, which is a convex set (see Proposition 4.2). Hence, there is some $\gamma_n \in \mathcal{L}_{\eta_n, 1}(\pi_{2,n})$ and $\pi_{1,n} \in \boldsymbol{\mathcal{Y}}_1$ such that

$$\gamma_n = \mu_{\eta_n, \pi_{1,n}, \pi_{2,n}}^{\mathbf{X} \times \mathbf{A}_1} = (1 - \sqrt{\epsilon_n}) \mu_{\eta_n, \pi_1, \pi_{2,n}}^{\mathbf{X} \times \mathbf{A}_1} + \sqrt{\epsilon_n} \mu_{\eta_n, \bar{\pi}_1, \pi_{2,n}}^{\mathbf{X} \times \mathbf{A}_1} \tag{4.8}$$

with, as a consequence, $\mu_{\eta_n, \pi_{1,n}, \pi_{2,n}}^{\mathbf{X}} = (1 - \sqrt{\epsilon_n}) \mu_{\eta_n, \pi_1, \pi_{2,n}}^{\mathbf{X}} + \sqrt{\epsilon_n} \mu_{\eta_n, \bar{\pi}_1, \pi_{2,n}}^{\mathbf{X}}$, and so

$$\mu_{\eta_n, \pi_{1,n}, \pi_{2,n}}^{\mathbf{X} \times \mathbf{A}_2} = \mu_{\eta_n, \pi_{1,n}, \pi_{2,n}}^{\mathbf{X}} \otimes \pi_{2,n} = (1 - \sqrt{\epsilon_n}) \mu_{\eta_n, \pi_1, \pi_{2,n}}^{\mathbf{X} \times \mathbf{A}_2} + \sqrt{\epsilon_n} \mu_{\eta_n, \bar{\pi}_1, \pi_{2,n}}^{\mathbf{X} \times \mathbf{A}_2}. \tag{4.9}$$

It is clear from (4.8) and Proposition 3.8 that $\lim_{n \to \infty} \gamma_n = \lim_{n \to \infty} \mu_{\eta_n, \pi_1, \pi_{2,n}}^{\mathbf{X} \times \mathbf{A}_1} = \mu_{\zeta, \pi_1, \pi_2}^{\mathbf{X} \times \mathbf{A}_1}$. We deduce from (4.8)–(4.9) that for all $n \geq n_1$

$$
\begin{aligned}
C_1(\eta_n, \pi_{1,n}, \pi_{2,n}) &= \int_{\mathbf{X} \times \mathbf{A}_1} c_1^1 d\mu_{\eta_n, \pi_{1,n}, \pi_{2,n}}^{\mathbf{X} \times \mathbf{A}_1} + \int_{\mathbf{X} \times \mathbf{A}_1} c_1^2 d\mu_{\eta_n, \pi_{1,n}, \pi_{2,n}}^{\mathbf{X} \times \mathbf{A}_2} \\
&= (1 - \sqrt{\epsilon_n}) C_1(\eta_n, \pi_1, \pi_{2,n}) + \sqrt{\epsilon_n} C_1(\eta_n, \bar{\pi}_1, \pi_{2,n}) \\
&\geq \rho_1 + \sqrt{\epsilon_n} \big[\delta - (1 - \sqrt{\epsilon_n}) \sqrt{\epsilon_n} \big],
\end{aligned} \tag{4.10}
$$

where the last inequality is derived from (4.6) and (4.7). Consequently, there exists some $N \geq n_1$ such that $C_1(\eta_n, \pi_{1,n}, \pi_{2,n}) \geq \rho_1$ for $n \geq N$. Since by definition $\gamma_n = \mu_{\eta_n, \pi_{1,n}, \pi_{2,n}}^{\mathbf{X} \times \mathbf{A}_1}$, this establishes precisely that $\gamma_n \in \mathcal{A}_{\eta_n, \rho_1, 1}(\pi_{2,n})$ for all $n \geq N$. This completes the proof of (i). The rationale for using the coefficient $\sqrt{\epsilon_n}$ in (4.8) is to mix the measures at a rate slower than the bound in (4.6) in order to satisfy the constraint as in (4.10). $\square$

Proposition 4.5 *Let $\eta \in \boldsymbol{\mathcal{P}}_\lambda^e(\mathbf{X})$ and $(\rho_1, \rho_2) \in \mathbb{R}^p \times \mathbb{R}^p$ satisfy the Slater condition. Then, the correspondences $\mathcal{A}_{\eta, \rho_1, 1} : \boldsymbol{\mathcal{Y}}_2 \twoheadrightarrow \mathcal{D}_{\eta, 1}$ defined by $\pi_2 \mapsto \mathcal{A}_{\eta, \rho_1, 1}(\pi_2)$ and $\mathcal{A}_{\eta, \rho_2, 2} : \boldsymbol{\mathcal{Y}}_1 \twoheadrightarrow \mathcal{D}_{\eta, 2}$ given by $\pi_1 \mapsto \mathcal{A}_{\eta, \rho_2, 2}(\pi_1)$ are both continuous with nonempty convex and compact values.*

Proof. We make the proof for the correspondence $\mathcal{A}_{\eta,\rho_1,1}$. Given $\pi_2 \in \boldsymbol{\mathcal{Y}}_2$, we have that $\mathcal{A}_{\eta,\rho_1,1}(\pi_2)$ is nonempty by Assumption (A.5). Since $\mathcal{L}_{\eta,1}(\pi_2)$ is convex (recall Proposition 4.2), the fact that $\mathcal{A}_{\eta,\rho_1,1}(\pi_2)$ is given by a linear constraint in γ —see (4.5)— yields that it is indeed a convex set.

We prove compactness of $\mathcal{A}_{\eta,\rho_1,1}(\pi_2)$ along with upper semicontinuity of $\pi_2 \mapsto \mathcal{A}_{\eta,\rho_1,1}(\pi_2)$. To this end, we use the Closed Graph Theorem in [1, Theorem 17.11]. Indeed, since the range of the correspondence is the compact metric space $\mathcal{D}_{\eta,1}$, the correspondence has closed graph if and only if it is upper semicontinuous and closed-valued (hence, in our case, compact-valued). So, it suffices to show that $\pi_2 \mapsto \mathcal{A}_{\eta,\rho_1,1}(\pi_2)$ has closed graph. Suppose that we have a convergent sequence $(\pi_{2,n}, \gamma_n)$ in the graph of $\mathcal{A}_{\eta,\rho_1,1}$. This means that $\pi_{2,n} \to \pi_2$ in $\boldsymbol{\mathcal{Y}}_2$ and that $\gamma_n \in \mathcal{A}_{\eta,\rho_1,1}(\pi_{2,n})$ are such that $\gamma_n \to \gamma \in \mathcal{D}_{\eta,1}$. Our goal is to prove that $\gamma \in \mathcal{A}_{\eta,\rho_1,1}(\pi_2)$. We note that for each $n \in \mathbb{N}$ there exists $\pi_{1,n} \in \boldsymbol{\mathcal{Y}}_1$ such that

$$\gamma_n = \mu_{\eta,\pi_{1,n},\pi_{2,n}}^{\mathbf{X}\times\mathbf{A}_1} \quad \text{and} \quad C_1(\eta,\pi_{1,n},\pi_{2,n}) \geq \rho_1 \tag{4.11}$$

for all $n \in \mathbb{N}$. There exists a subsequence $\{\pi_{1,n'}\}$ of $\{\pi_{1,n}\}_{n\in\mathbb{N}}$ that converges to some $\pi_1 \in \boldsymbol{\mathcal{Y}}_1$. Using Proposition 3.8 in (4.11), we have $\gamma_{n'} \to \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}\times\mathbf{A}_1} = \gamma$, which establishes $\gamma \in \mathcal{L}_{\eta,1}(\pi_2)$, and $C_1(\eta,\pi_1,\pi_2) \geq \rho_1$, from which $\gamma \in \mathcal{A}_{\eta,\rho_1,1}(\pi_2)$ follows. Lower semicontinuity of $\mathcal{A}_{\eta,\rho_1,1}$ is a direct consequence of Proposition 4.4(i) applied to the constant sequence $\eta_n \equiv \eta$ and the sequential characterization of lower semicontinuity of correspondences given in [1, Theorem 17.21]. $\square$

Based on Remark 4.3, given $\pi_1 \in \boldsymbol{\mathcal{Y}}_1$ or, equivalently, given $\gamma \in \mathcal{L}_{\eta,1}(\pi_2)$ we have

$$R_1(\eta,\pi_1,\pi_2) = \int_{\mathbf{X}\times\mathbf{A}_1} r_1^1 d\gamma + \int_{\mathbf{X}\times\mathbf{A}_2} r_1^2 d(\gamma^{\mathbf{X}} \otimes \pi_2).$$

For a fixed policy of player 2, therefore, the goal of player 1 is to maximize the above integral over all $\gamma \in \mathcal{A}_{\eta,\rho_1,1}(\pi_2)$, that is, over all measures $\gamma \in \mathcal{L}_{\eta,1}(\pi_2)$ which satisfy his constraint. This leads to the definition of the correspondence $\mathcal{H}_{\eta,\rho_1,1} : \boldsymbol{\mathcal{Y}}_2 \twoheadrightarrow \mathcal{D}_{\eta,1}$ given by

$$\mathcal{H}_{\eta,\rho_1,1}(\pi_2) = \underset{\gamma \in \mathcal{A}_{\eta,\rho_1,1}(\pi_2)}{\arg\max} \left\{ \int_{\mathbf{X}\times\mathbf{A}_1} r_1^1 d\gamma + \int_{\mathbf{X}\times\mathbf{A}_2} r_1^2 d(\gamma^{\mathbf{X}} \otimes \pi_2) \right\}$$

for $\pi_2 \in \boldsymbol{\mathcal{Y}}_2$, $\rho_1 \in \mathbb{R}^p$, and $\eta \in \boldsymbol{\mathcal{P}}_\lambda^e(\mathbf{X})$. Similarly, we define $\mathcal{H}_{\eta,\rho_2,2} : \boldsymbol{\mathcal{Y}}_1 \twoheadrightarrow \mathcal{D}_2$ as

$$\mathcal{H}_{\eta,\rho_2,2}(\pi_1) = \underset{\gamma \in \mathcal{A}_{\eta,\rho_2,2}(\pi_1)}{\arg\max} \left\{ \int_{\mathbf{X}\times\mathbf{A}_1} r_2^1 d(\gamma^{\mathbf{X}} \otimes \pi_1) + \int_{\mathbf{X}\times\mathbf{A}_2} r_2^2 d\gamma \right\}$$

for $\pi_1 \in \boldsymbol{\mathcal{Y}}_1$, $\rho_2 \in \mathbb{R}^p$, and $\eta \in \boldsymbol{\mathcal{P}}_\lambda^e(\mathbf{X})$

Proposition 4.6 *Let $\eta \in \boldsymbol{\mathcal{P}}_\lambda^e(\mathbf{X})$ and $(\rho_1,\rho_2) \in \mathbb{R}^p \times \mathbb{R}^p$ satisfy the Slater condition. The correspondences $\mathcal{H}_{\eta,\rho_1,1}$ and $\mathcal{H}_{\eta,\rho_2,2}$ are upper semicontinuous with nonempty compact and convex values.*

Proof. We prove the result only for $\mathcal{H}_{\eta,\rho_1 1}$. We have that

$$(\pi_2,\gamma) \mapsto f(\pi_2,\gamma) = \int_{\mathbf{X}\times\mathbf{A}_1} r_1^1 d\gamma + \int_{\mathbf{X}\times\mathbf{A}_2} r_1^2 d(\gamma^{\mathbf{X}} \otimes \pi_2)$$

is continuous on the graph of $\mathcal{A}_{\eta,\rho_1,1}$. Indeed, suppose that $\{\pi_{2,n}\} \subseteq \boldsymbol{\mathcal{Y}}_2$ converges to some $\pi_2 \in \boldsymbol{\mathcal{Y}}_2$ and that $\gamma_n \in \mathcal{A}_{\eta,\rho_1,1}(\pi_{2,n})$ is such that $\gamma_n \to \gamma$ with, necessarily (the correspondence $\mathcal{A}_{\eta,\rho_1,1}$ being closed) $\gamma \in \mathcal{A}_{\eta,\rho_1,1}(\pi_2)$. First of all note that $r_1^1 \in \mathcal{C}ar_b(\mathbf{X} \times \mathbf{A}_1, \mathbb{R})$ and so by definition of

17

the ws-convergence we have $\int_{\mathbf{X}\times\mathbf{A}_1} r_1^1 d\gamma_n \to \int_{\mathbf{X}\times\mathbf{A}_1} r_1^1 d\gamma$. On the other hand, since the sequence $\int r_1^2 d(\gamma_n^{\mathbf{X}} \otimes \pi_{2,n})$ is bounded, in order to prove the convergence

$$\int_{\mathbf{X}\times\mathbf{A}_2} r_1^2 d(\gamma_n^{\mathbf{X}} \otimes \pi_{2,n}) \to \int_{\mathbf{X}\times\mathbf{A}_2} r_1^2 d(\gamma^{\mathbf{X}} \otimes \pi_2) \qquad (4.12)$$

it suffices to show that the above limit holds through any convergent subsequence $\int r_1^2 d(\gamma_{n'}^{\mathbf{X}} \otimes \pi_{2,n'})$. For $\gamma_{n'}$ there exists some $\pi_{1,n'} \in \boldsymbol{\mathcal{Y}}_1$ such that $\gamma_{n'} = \mu_{\eta,\pi_{1,n'},\pi_{2,n'}}^{\mathbf{X}\times\mathbf{A}_1}$. Without loss of generality we can assume that, for some $\pi_1 \in \boldsymbol{\mathcal{Y}}_1$, we have $\pi_{1,n'} \to \pi_1$. In particular, using Proposition 3.8, this implies that $\gamma_{n'} \to \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}\times\mathbf{A}_1} = \gamma$. But we also have $\gamma_{n'}^{\mathbf{X}} \otimes \pi_{2,n'} = \mu_{\eta,\pi_{1,n'},\pi_{2,n'}}^{\mathbf{X}\times\mathbf{A}_2} \to \mu_{\eta,\pi_1,\pi_2}^{\mathbf{X}\times\mathbf{A}_2} = \gamma^{\mathbf{X}} \otimes \pi_2$ by using again Proposition 3.8. Thus, since $r_1^2 \in \mathcal{C}ar_b(\mathbf{X} \times \mathbf{A}_2, \mathbb{R})$, the convergence (4.12) through n' follows. This completes the proof of the continuity of the function f on the graph of $\mathcal{A}_{\eta,\rho_1,1}$.

Now we are in position to apply Berge's Maximum Theorem [1, Theorem 17.31] to the function f on the graph of $\mathcal{A}_{\eta,\rho_1,1}$. Indeed, we have already shown that f is continuous and, besides, we have that $\mathcal{A}_{\eta,\rho_1,1}$ is a continuous correspondence with nonempty compact values (Proposition 4.5). By the maximum theorem we conclude that the arg max correspondence (that is, $\mathcal{H}_{\eta,\rho_1,1}$) is upper semicontinuous with nonempty compact values.

Finally, from Proposition 4.5 we also have that $\mathcal{A}_{\eta,\rho_1,1}(\pi_2)$ is convex and since the mapping $\gamma \mapsto f(\pi_2,\gamma) = \int_{\mathbf{X}\times\mathbf{A}_1} r_1^1 d\gamma + \int_{\mathbf{X}\times\mathbf{A}_2} r_1^2 d(\gamma^{\mathbf{X}} \otimes \pi_2)$ is linear in γ for fixed π_2, it follows that its set of maxima is a convex set. This proves that $\mathcal{H}_{\eta,\rho_1,1}$ is also convex-valued. $\qquad\square$

Hence, given an initial distribution $\eta \in \boldsymbol{\mathcal{P}}_\lambda^e(\mathbf{X})$ and constraint constants $\rho = (\rho_1,\rho_2) \in \mathbb{R}^p \times \mathbb{R}^p$ satisfying the Slater condition, let us define the correspondence $\mathcal{H}_{\eta,\rho} : \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2 \twoheadrightarrow \mathcal{D}_{\eta,1} \times \mathcal{D}_{\eta,2}$ as

$$\mathcal{H}_{\eta,\rho}(\pi_1,\pi_2) = \mathcal{H}_{\eta,\rho_1,1}(\pi_2) \times \mathcal{H}_{\eta,\rho_2,2}(\pi_1) \quad \text{for } (\pi_1,\pi_2) \in \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2.$$

As a consequence of the previous results and using [1, Theorem 17.28], we conclude that the correspondence $\mathcal{H}_{\eta,\rho}$ is upper semicontinuous with nonempty compact and convex values.

Nash equilibrium for an initial distribution $\eta \in \boldsymbol{\mathcal{P}}_\lambda^e(\mathbf{X})$.

Theorem 4.7 *Let $\eta \in \boldsymbol{\mathcal{P}}_\lambda^e(\mathbf{X})$ and $\rho = (\rho_1,\rho_2) \in \mathbb{R}^p \times \mathbb{R}^p$ satisfy the Slater condition. Under Assumption A the following results hold.*

  (i) The correspondence $\mathcal{H}_{\eta,\rho} \circ \mathcal{J}_\eta : \mathcal{D}_{\eta,1} \times \mathcal{D}_{\eta,2} \twoheadrightarrow \mathcal{D}_{\eta,1} \times \mathcal{D}_{\eta,2}$ has a fixed point $(\gamma_1^*,\gamma_2^*) \in \mathcal{D}_{\eta,1} \times \mathcal{D}_{\eta,2}$.

  (ii) The pair of stationary Markov policies $(\pi_1^*,\pi_2^*) = \mathcal{J}_\eta(\gamma_1^*,\gamma_2^*) \in \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ is a constrained equilibrium in the class of history-dependent policies $\boldsymbol{\Pi}_1 \times \boldsymbol{\Pi}_2$ for the game model $\mathcal{G}(\eta,\rho)$.

Proof: (i). We can use [1, Theorem 17.23] to prove that $\mathcal{H}_{\eta,\rho} \circ \mathcal{J}_\eta$ is an upper semicontinuous correspondence (it is the composition of the continuous function $\mathcal{J}_\eta$ and the upper semicontinuous correspondence $\mathcal{H}_{\eta,\rho}$) with nonempty compact and convex values. By the Closed Graph Theorem [1, Theorem 17.11], the correspondence $\mathcal{H}_{\eta,\rho} \circ \mathcal{J}_\eta$ is closed. On the other hand, $\mathcal{D}_{\eta,1} \times \mathcal{D}_{\eta,2}$ is a nonempty compact convex subset of the locally convex Hausdorff space $\mathcal{M}(\mathbf{X} \times \mathbf{A}_1) \times \mathcal{M}(\mathbf{X} \times \mathbf{A}_2)$; recall Proposition 2.2. The Kakutani-Fan-Glicksberg fixed point theorem [1, Corollary 17.55] yields the existence of a fixed point for $\mathcal{H}_{\eta,\rho} \circ \mathcal{J}_\eta$.

(ii). Denote by $(\gamma_1^*,\gamma_2^*) \in \mathcal{D}_{\eta,1} \times \mathcal{D}_{\eta,2}$ a fixed point of $\mathcal{H}_{\eta,\rho} \circ \mathcal{J}_\eta$ and let $(\pi_1^*,\pi_2^*) \in \boldsymbol{\mathcal{Y}}_1 \times \boldsymbol{\mathcal{Y}}_2$ be the stationary Markov policies given by $(\pi_1^*,\pi_2^*) = \mathcal{J}_\eta(\gamma_1^*,\gamma_2^*)$, so that $(\gamma_1^*,\gamma_2^*) \in \mathcal{H}_{\eta,\rho}(\pi_1^*,\pi_2^*)$. Observe that $\gamma_1^* \in \mathcal{L}_{\eta,1}(\pi_2^*)$ and so $\gamma_1^* = \mu_{\eta,\pi_1,\pi_2^*}^{\mathbf{X}\times\mathbf{A}_1}$ for some $\pi_1 \in \boldsymbol{\mathcal{Y}}_1$. But then we have $\pi_1 = \mathcal{J}_{\eta,1}(\gamma_1^*) = \pi_1^*$. Hence, using the symmetric argument for γ_2^* we can conclude that

$$\gamma_1^* = \mu_{\eta,\pi_1^*,\pi_2^*}^{\mathbf{X}\times\mathbf{A}_1} \quad \text{and} \quad \gamma_2^* = \mu_{\eta,\pi_1^*,\pi_2^*}^{\mathbf{X}\times\mathbf{A}_1}. \qquad (4.13)$$

It follows that $C_1(\eta, \pi_1^*, \pi_2^*) = \int_{\mathbf{X} \times \mathbf{A}_1} c_1^1 d\gamma_1^* + \int_{\mathbf{X} \times \mathbf{A}_2} c_1^2 d(\gamma_1^{*\mathbf{X}} \otimes \pi_2^*) \geq \rho_1$ since $\gamma_1^* \in \mathcal{A}_{\eta, \rho_1, 1}(\pi_2^*)$. Similarly, we can show that $C_2(\eta, \pi_1^*, \pi_2^*) \geq \rho_2$ and so, the stationary Markov policies (π_1^*, π_2^*) satisfy the constraints of both players (see item (i) in Definition 2.5).

Suppose now that player 1 varies his policy from $\pi_1^* \in \mathcal{Y}_1$ to some history-dependent policy $\pi_1 \in \mathbf{\Pi}_1$ which satisfies his own constraint, that is, $C_1(\eta, \pi_1, \pi_2^*) \geq \rho_1$. We want to check that $R_1(\eta, \pi_1, \pi_2^*) \leq R_1(\eta, \pi_1^*, \pi_2^*)$. By Corollary 3.3 we can suppose without loss of generality that $\pi_1 \in \mathcal{Y}_1$. Hence, for this Markov policy $\pi_1 \in \mathcal{Y}_1$ we have

$$C_1(\eta, \pi_1, \pi_2^*) = \int_{\mathbf{X} \times \mathbf{A}_1} c_1^1 d\mu_{\eta, \pi_1, \pi_2^*}^{\mathbf{X} \times \mathbf{A}_1} + \int_{\mathbf{X} \times \mathbf{A}_2} c_1^2 d\mu_{\eta, \pi_1, \pi_2^*}^{\mathbf{X} \times \mathbf{A}_2} \geq \rho_1. \tag{4.14}$$

With this in mind, (4.14) implies that $\mu_{\eta, \pi_1, \pi_2^*}^{\mathbf{X} \times \mathbf{A}_1} \in \mathcal{A}_{\eta, \rho_1, 1}(\pi_2^*)$ but by the definition of $\gamma_1^* \in \mathcal{H}_{\eta, \rho_1, 1}(\pi_2^*)$ we have

$$
\begin{aligned}
R_1(\eta, \pi_1, \pi_2^*) &= \int_{\mathbf{X} \times \mathbf{A}_1} r_1^1 d\mu_{\eta, \pi_1, \pi_2^*}^{\mathbf{X} \times \mathbf{A}_1} + \int_{\mathbf{X} \times \mathbf{A}_2} r_1^2 d(\mu_{\eta, \pi_1, \pi_2^*}^{\mathbf{X}} \otimes \pi_2^*) \\
&\leq \int_{\mathbf{X} \times \mathbf{A}_1} r_1^1 d\gamma_1^* + \int_{\mathbf{X} \times \mathbf{A}_2} r_1^2 d(\gamma_1^{*\mathbf{X}} \otimes \pi_2^*) = R_1(\eta, \pi_1^*, \pi_2^*),
\end{aligned}
$$

where we have made use of (4.13). We proceed similarly for player 2 and we conclude that the stationary Markov policies $(\pi_1^*, \pi_2^*) \in \mathcal{Y}_1 \times \mathcal{Y}_2$ are indeed a constrained Nash equilibrium. $\qquad\square$

4.2   Existence of a Nash equilibrium for an initial distribution $\nu \in \mathcal{P}_\lambda(\mathbf{X})$

Theorem 4.7 establishes the existence of a Nash equilibrium for an initial distribution in $\mathcal{P}_\lambda^e(\mathbf{X})$. To obtain our main result in this paper, we now drop this condition and replace it with $\nu \in \mathcal{P}_\lambda(\mathbf{X})$.

Theorem 4.8 *Consider the game model $\mathcal{G}(\nu, \theta)$ and suppose that Assumption A is satisfied for the initial distribution $\nu \in \mathcal{P}_\lambda(\mathbf{X})$. There exist stationary Markov policies $(\pi_1^*, \pi_2^*) \in \mathbf{M}_1 \times \mathbf{M}_2$ which are a constrained equilibrium for the players in the class of history-dependent policies $\mathbf{\Pi}_1 \times \mathbf{\Pi}_2$.*

Proof. Let $\{\eta_n\}_{n \in \mathbb{N}} \subseteq \mathcal{P}_\lambda^e(\mathbf{X})$ be the sequence defined by $\eta_n = \frac{n}{n+1}\nu + \frac{1}{n+1}\lambda$ for $n \in \mathbb{N}$. Clearly, $\{\eta_n\}_{n \in \mathbb{N}}$ converges to ν in total variation. Recalling (3.3)–(3.4) and since the constraint function c_i is bounded, there is some constant $\mathfrak{c} > 0$ such that $|C_i(\eta_n, \pi_1, \pi_2) - C_i(\nu, \pi_1, \pi_2)| \leq \mathfrak{c}/(n+1)$ for all $i = 1, 2$ and $(\pi_1, \pi_2) \in \mathcal{Y}_1 \times \mathcal{Y}_2$. Hence, the initial distribution $\eta_n \in \mathcal{P}_\lambda^e(\mathbf{X})$ and the constraint constants $\rho_n = (\rho_{1,n}, \rho_{2,n})$ given by $\rho_{i,n} = \theta_i - \frac{\mathfrak{c}}{n+1}\mathbf{1}_p$ for $n \in \mathbb{N}$ satisfy the Slater condition in Definition 2.6. We are thus in position to apply Theorem 4.7 to the game model $\mathcal{G}(\eta_n, \rho_n)$ to obtain the existence of a Nash equilibrium. That is, for each $n \in \mathbb{N}$ there exist stationary Markov policies $\{(\pi_{1,n}^*, \pi_{2,n}^*)\}_{n \in \mathbb{N}}$ in $\mathcal{Y}_1 \times \mathcal{Y}_2$ satisfying

$$C_i(\eta_n, \pi_{1,n}^*, \pi_{2,n}^*) \geq \rho_{i,n} \quad \text{for } i = 1, 2. \tag{4.15}$$

and

$$\forall \pi_1 \in \mathbf{\Pi}_1, \ C_1(\eta_n, \pi_1, \pi_{2,n}^*) \geq \rho_{1,n} \Rightarrow R_1(\eta_n, \pi_{1,n}^*, \pi_{2,n}^*) \geq R_1(\eta_n, \pi_1, \pi_{2,n}^*),$$

$$\forall \pi_2 \in \mathbf{\Pi}_2, \ C_2(\eta_n, \pi_{1,n}^*, \pi_2) \geq \rho_{2,n} \Rightarrow R_2(\eta_n, \pi_{1,n}^*, \pi_{2,n}^*) \geq R_2(\eta_n, \pi_{1,n}^*, \pi_2).$$

By compactness of $\mathcal{Y}_1$ and $\mathcal{Y}_2$, there exists a convergent subsequence of $\{(\pi_{1,n}^*, \pi_{2,n}^*)\}_{n \in \mathbb{N}}$. Without loss of generality, we will assume that the whole sequence converges to some $(\pi_1^*, \pi_2^*) \in \mathcal{Y}_1 \times \mathcal{Y}_2$. Our goal now is to show that (π_1^*, π_2^*) is a Nash equilibrium for the game model $\mathcal{G}(\nu, \theta)$.

From Proposition 3.8, we get by taking the limit in (4.15) $C_i(\nu, \pi_1^*, \pi_2^*) \geq \theta_i$ for $i = 1, 2$. On the other hand, let π_1 be an arbitrary policy in $\mathbf{\Pi}_1$ satisfying $C_1(\nu, \pi_1, \pi_2^*) \geq \theta_1$. Without loss

19

of generality, it can be assumed that $\pi_1 \in \mathcal{Y}_1$ (recall Corollary 3.3). Consider the associated measure $\gamma = \mu_{\nu,\pi_1,\pi_2^*}^{\mathbf{X} \times \mathbf{A}_1}$. We can apply Proposition 4.4 to deduce the existence of a sequence $\gamma_n \to \gamma$ with $\gamma_n \in \mathcal{A}_{\eta_n,\theta_1,1}(\pi_{2,n}^*)$ for large enough $n \geq N$. For such large n, let $\pi_{1,n} \in \mathcal{Y}_1$ be such that $\gamma_n = \mu_{\eta_n,\pi_{1,n},\pi_{2,n}^*}^{\mathbf{X} \times \mathbf{A}_1}$. We have that $C_1(\eta_n, \pi_{1,n}, \pi_{2,n}^*) \geq \theta_1 \geq \rho_{1,n}$ for any $n \geq N$. But $(\pi_{1,n}^*, \pi_{2,n}^*)$ being a Nash equilibrium for the game model $\mathcal{G}(\eta_n, \rho_n)$, for such $n \geq N$

$$R_1(\eta_n, \pi_{1,n}^*, \pi_{2,n}^*) \geq R_1(\eta_n, \pi_{1,n}, \pi_{2,n}^*) = \int_{\mathbf{X} \times \mathbf{A}_1} r_1^1 d\gamma_n + \int_{\mathbf{X} \times \mathbf{A}_2} r_1^2 d(\gamma_n^{\mathbf{X}} \otimes \pi_{2,n}^*).$$

Observe that since $\gamma_n \to \gamma$, we have

$$\lim_{n \to \infty} \int_{\mathbf{X} \times \mathbf{A}_1} r_1^1 d\gamma_n = \int_{\mathbf{X} \times \mathbf{A}_1} r_1^1 d\gamma = \int_{\mathbf{X} \times \mathbf{A}_1} r_1^1 d\mu_{\nu,\pi_1,\pi_2^*}^{\mathbf{X} \times \mathbf{A}_1}.$$

We want to show that $\int_{\mathbf{X} \times \mathbf{A}_2} r_1^2 d(\gamma_n^{\mathbf{X}} \otimes \pi_{2,n}^*) \to \int_{\mathbf{X} \times \mathbf{A}_2} r_1^2 d\mu_{\nu,\pi_1,\pi_2^*}^{\mathbf{X} \times \mathbf{A}_2}$. Since r_1^2 is bounded, it suffices to show that the limit holds through any convergent subsequence $\{n'\}$ of $\int_{\mathbf{X} \times \mathbf{A}_2} r_1^2 d(\gamma_n^{\mathbf{X}} \otimes \pi_{2,n}^*)$. Without loss of generality, we can assume that $\pi_{1,n'} \to \bar{\pi}_1$ for some $\bar{\pi}_1 \in \mathcal{Y}_1$. By Proposition 3.8 we have

$$\gamma_{n'}^{\mathbf{X}} \otimes \pi_{2,n'}^* = \mu_{\eta_{n'},\pi_{1,n'},\pi_{2,n'}^*}^{\mathbf{X} \times \mathbf{A}_2} \to \mu_{\nu,\bar{\pi}_1,\pi_2^*}^{\mathbf{X} \times \mathbf{A}_2} = \mu_{\nu,\bar{\pi}_1,\pi_2^*}^{\mathbf{X}} \otimes \pi_2^*$$

and, in particular, $\gamma_{n'}^{\mathbf{X}} \to \mu_{\nu,\bar{\pi}_1,\pi_2^*}^{\mathbf{X}}$ in the s-topology of $\mathcal{P}(\mathbf{X})$. On the other hand, we have $\gamma_{n'}^{\mathbf{X}} \to \gamma^{\mathbf{X}} = \mu_{\nu,\pi_1,\pi_2^*}^{\mathbf{X}}$ is the s-topology of $\mathcal{P}(\mathbf{X})$. This implies that $\mu_{\nu,\pi_1,\pi_2^*}^{\mathbf{X}} = \mu_{\nu,\bar{\pi}_1,\pi_2^*}^{\mathbf{X}}$. We conclude that

$$\gamma_{n'}^{\mathbf{X}} \otimes \pi_{2,n'}^* \to \mu_{\nu,\pi_1,\pi_2^*}^{\mathbf{X}} \otimes \pi_2^* = \mu_{\nu,\pi_1,\pi_2^*}^{\mathbf{X} \times \mathbf{A}_2}$$

and the desired convergence follows. Using again Proposition 3.8 we can take the limit in

$$R_1(\eta_n, \pi_{1,n}^*, \pi_{2,n}^*) \geq \int_{\mathbf{X} \times \mathbf{A}_1} r_1^1 d\gamma_n + \int_{\mathbf{X} \times \mathbf{A}_2} r_1^2 d(\gamma_n^{\mathbf{X}} \otimes \pi_{2,n}^*).$$

and obtain $R_1(\nu, \pi_1^*, \pi_2^*) \geq R_1(\nu, \pi_1, \pi_2^*)$. Summarizing, if $\pi_1 \in \mathcal{Y}_1$ satisfies $C_1(\nu, \pi_1, \pi_2^*) \geq \theta_1$ then, necessarily, we have $R_1(\nu, \pi_1^*, \pi_2^*) \geq R_1(\nu, \pi_1, \pi_2^*)$.

Proceeding symmetrically for player 2, it follows that (π_1^*, π_2^*) is a Nash equilibrium. $\qquad \square$

References

[1] Charalambos D. Aliprantis and Kim C. Border. *Infinite dimensional analysis*. Springer, Berlin, third edition, 2006. A hitchhiker's guide.

[2] Eitan Altman, Konstantin Avrachenkov, Nicolas Bonneau, Merouane Debbah, Rachid El-Azouzi, and Daniel Sadoc Menasche. Constrained cost-coupled stochastic games with independent state processes. *Oper. Res. Lett.*, 36(2):160–164, 2008.

[3] Eitan Altman and Adam Shwartz. Constrained Markov games: Nash equilibria. In *Advances in dynamic games and applications (Kanagawa, 1996)*, volume 5 of *Ann. Internat. Soc. Dynam. Games*, pages 213–221. Birkhäuser Boston, Boston, MA, 2000.

[4] Jorge Alvarez-Mena and Onésimo Hernández-Lerma. Existence of Nash equilibria for constrained stochastic games. *Math. Methods Oper. Res.*, 63(2):261–285, 2006.

[5] Erik J. Balder. Generalized equilibrium results for games with incomplete information. *Math. Oper. Res.*, 13(2):265–276, 1988.

[6] Erik J. Balder. On Cournot-Nash equilibrium distributions for games with differential information and discontinuous payoffs. *Econom. Theory*, 1(4):339–354, 1991.

[7] Erik J. Balder. On ws-convergence of product measures. *Math. Oper. Res.*, 26(3):494–518, 2001.

[8] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.

[9] D. Duffie, J. Geanakoplos, A. Mas-Colell, and A. McLennan. Stationary Markov equilibria. *Econometrica*, 62(4):745–781, 1994.

[10] John Duggan. Noisy stochastic games. *Econometrica*, 80(5):2017–2045, 2012.

[11] Wei He and Yeneng Sun. Stationary Markov perfect equilibria in discounted stochastic games. *J. Econom. Theory*, 169:35–61, 2017.

[12] Onésimo Hernández-Lerma and Jean-Bernard Lasserre. *Discrete-time Markov control processes: Basic optimality criteria*, volume 30 of *Applications of Mathematics*. Springer-Verlag, New York, 1996.

[13] C. J. Himmelberg, T. Parthasarathy, T. E. S. Raghavan, and F. S. Van Vleck. Existence of *p*-equilibrium and optimal stationary strategies in stochastic games. *Proc. Amer. Math. Soc.*, 60:245–251 (1977), 1976.

[14] Jean Jacod and Jean Mémin. Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en probabilité. In *Seminar on Probability, XV (Univ. Strasbourg, Strasbourg, 1979/1980) (French)*, volume 850 of *Lecture Notes in Math.*, pages 529–546. Springer, Berlin-New York, 1981.

[15] Anna Jaśkiewicz and Andrzej S. Nowak. On pure stationary almost Markov Nash equilibria in nonzero-sum ARAT stochastic games. *Math. Methods Oper. Res.*, 81(2):169–179, 2015.

[16] Anna Jaśkiewicz and Andrzej S. Nowak. Nonzero-sum stochastic games. In *Handbook of dynamic game theory*, pages 281–344. Springer, Cham, [2018] ©2018.

[17] Anna Jaśkiewicz and Andrzej S. Nowak. Constrained Markov decision processes with expected total reward criteria. *SIAM J. Control Optim.*, 57(5):3118–3136, 2019.

[18] Heinz-Uwe Küenle. Equilibrium strategies in stochastic games with additive cost and transition structure. *Int. Game Theory Rev.*, 1(2):131–147, 1999.

[19] A. S. Nowak. Existence of equilibrium stationary strategies in discounted noncooperative stochastic games with uncountable state space. *J. Optim. Theory Appl.*, 45(4):591–602, 1985.

[20] A. S. Nowak and T. E. S. Raghavan. Existence of stationary correlated equilibria with symmetric information for discounted stochastic games. *Math. Oper. Res.*, 17(3):519–526, 1992.

[21] Andrzej S. Nowak. Nonrandomized strategy equilibria in noncooperative stochastic games with additive transition and reward structure. *J. Optim. Theory Appl.*, 52(3):429–441, 1987.

[22] Andrzej S. Nowak. On a new class of nonzero-sum discounted stochastic games having stationary Nash equilibrium points. *Internat. J. Game Theory*, 32(1):121–132, 2003.

[23] Andrzej S. Nowak. On stochastic games in economics. *Math. Methods Oper. Res.*, 66(3):513–530, 2007.

[24] T. Parthasarathy. Existence of equilibrium stationary strategies in discounted stochastic games. *Sankhyā Ser. A*, 44(1):114–127, 1982.

[25] T. Parthasarathy and S. Sinha. Existence of stationary equilibrium strategies in non-zero-sum discounted stochastic games with uncountable state space and state-independent transitions. *Internat. J. Game Theory*, 18(2):189–194, 1989.

[26] Manfred Schäl. Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 32(3):179–196, 1975.

[27] Manfred Schäl. On dynamic programming: compactness of the space of policies. *Stochastic Processes Appl.*, 3(4):345–364, 1975.

[28] Vikas Vikram Singh and N. Hemachandra. A characterization of stationary Nash equilibria of constrained stochastic games with independent state processes. *Oper. Res. Lett.*, 42(1):48–52, 2014.

[29] Frank Thuijsman and Thirukkannamangai E. S. Raghavan. Perfect information stochastic games and related classes. *Internat. J. Game Theory*, 26(3):403–408, 1997.