STOCHASTIC GAMES WITH TWO NON-ABSORBING STATES

BY

EILON SOLAN*

MEDS Department, Kellogg Graduate School of Management
Northwestern University, 2001 Sheridan Rd., Evanston, IL 60208, USA
e-mail: e-solan@nwu.edu

ABSTRACT

In the present paper we consider recursive games that satisfy an absorbing property defined by Vieille. We give two sufficient conditions for existence of an equilibrium payoff in such games, and prove that if the game has at most two non-absorbing states, then at least one of the conditions is satisfied. Using a reduction of Vieille, we conclude that every stochastic game which has at most two non-absorbing states admits an equilibrium payoff.

1. Introduction

A two-player stochastic game is played in stages. At every stage the game is in one of finitely many states. Each of the players chooses independently of his opponent an action in his action space. The pair of actions, together with the current state, determine the daily payoff for the players and the probability distribution according to which a new state of the game is chosen.

An **equilibrium payoff** is a vector of payoffs $g = (g_s^i)$ (where s is a state and i is a player), such that for every $\epsilon > 0$ there is a strategy profile (which is called an ϵ-**equilibrium profile**), that satisfies for every initial state s:

- If the players follow this strategy profile, then the expected lim inf of the average payoffs of each player i in the infinite game, as well as the expected average payoff in any finite game which is sufficiently long, is at least $g_s^i - \epsilon$.

- If any player i deviates to another strategy, then the expected lim sup of his average payoffs in the infinite game, as well as his expected average payoff in any finite game which is sufficiently long, is at most $g_s^i + \epsilon$.

If the game is zero-sum, then the unique equilibrium payoff is the value of the game. Mertens and Neyman [4] proved that every zero-sum stochastic game has a value. Vrieze and Thuijsman [11] proved that every non zero-sum stochastic game, in which only one state is non-absorbing, has an equilibrium payoff (a state is **absorbing** if the probability to leave it, whatever the players play, is 0; otherwise, it is **non-absorbing**).

Vieille [8, 9] proved that in order to prove existence of equilibrium payoffs in general stochastic games it is sufficient to prove the existence for the class of **positive recursive games with the absorbing property**. In these games the daily payoff for the players is 0 in every non-absorbing state whatever actions they play, the payoff for player 2 in absorbing states is positive, and if player 2 plays a fully mixed stationary strategy then the game eventually reaches an absorbing state with probability 1, whatever player 1 plays.

Following closely Vieille's reduction reveals that he proves even more. Vieille proves that if every positive recursive game with the absorbing property and at most n non-absorbing states has an equilibrium payoff, then every stochastic game with at most n non-absorbing states has an equilibrium payoff.

In the present paper we give two sufficient conditions for existence of an equilibrium payoff in positive recursive games with the absorbing property. Furthermore, we prove that every positive recursive game with the absorbing property, which has at most two non-absorbing states, satisfies at least one of these conditions. By the reduction of Vieille we conclude that every stochastic game which has at most two non-absorbing states has an equilibrium payoff.

The basic difficulty with undiscounted stochastic games is that the undiscounted payoff is not continuous over the strategy space. To overcome this difficulty we note that since player 2 can force absorption, and his absorbing payoff is always positive, it follows that his min-max value is positive. Since the payoff in non-absorbing states is 0, every ϵ-equilibrium strategy profile (if it exists) must be absorbing with high probability (for ϵ sufficiently small).

We define ϵ-approximating games, where player 2 is restricted to fully mixed stationary strategies, and player 1 is not restricted. As $\epsilon \to 0$, the restrictions

on player 2 become weaker. Since the game satisfies the absorbing property, the undiscounted payoff is continuous over the restricted strategy space, and using a standard fixed point theorem one can prove that there exists a stationary equilibrium profile in the ϵ-approximating game. By studying the asymptotic behavior of a sequence of equilibria in the ϵ-approximating games we construct different types of equilibrium payoffs in the original undiscounted game.

Unfortunately, the equilibrium payoff need not be equal to the limit of the equilibrium payoffs of the ϵ-approximating games, hence, as in Vrieze and Thuijsman [11], we cannot generalize the approach for games with more than 2 non-absorbing states. We hope that an approach similar to ours can prove that an equilibrium payoff exists in any positive recursive game with the absorbing property.

The method of studying asymptotic behavior of equilibria in approximating games was used by Vrieze and Thuijsman [11] to prove existence of equilibrium payoff in stochastic games with a single non-absorbing state. In their case the approximating game was the discounted stochastic game. Restricting one of the players to play a fully mixed stationary strategy in order to make the undiscounted payoff continuous, appeared already in Evangelista et al. [2].

Independently, Vieille [10] has proved the existence of an equilibrium payoff in general positive recursive games with the absorbing property. In Vieille's proof, as in our approach, player 1 is not restricted, while player 2 is restricted in his choice of a strategy. Vieille defines for every $\epsilon > 0$ a correspondence (set-valued function) that assigns for each pair of stationary strategies of the two players (i) the set of best reply stationary strategies of player 1 against the strategy used by player 2, and (ii) a collection of fully mixed stationary strategies of player 2 which are almost optimal against the strategy of player 1. Using standard arguments Vieille proves that for every $\epsilon > 0$ this correspondence admits a fixed point and, by studying the asymptotic behavior of a sequence of fixed points, he is able to construct ϵ-equilibrium strategy profiles.

The main difference between the two approaches is that, while we define an approximating game and study equilibria in this game, Vieille defines an approximating best reply correspondence, and studies fixed points of this correspondence. Furthermore, Vieille's definition of the approximating best reply correspondence is more sophisticated than our definition of the approximating games. Vieille's technique was applied successfully to prove the existence of stationary extensive form correlated equilibria in n-player positive recursive games (see Solan and Vieille [7]).

The paper is arranged as follows. In section 2 we give an example of a recursive

game with the absorbing property, and show some of the equilibrium payoffs in this game. In section 3 we give the model of stochastic games and state the main result. In section 4 we state and prove two sufficient conditions for existence of an equilibrium payoff.

Sections 5–7 are devoted to prove that in every positive recursive game with the absorbing property, which has at most two non-absorbing states, at least one of the sufficient conditions hold. In section 5 we give some preliminary results, in section 6 we introduce the ϵ-approximating games, and in section 7 the main result is proven. In section 8 we give an example of a game with more than two non-absorbing states, and show why our approach fails in this game.

2. An example

Consider the following positive recursive game:

state 1

	L	C	R
T	1	1	1/2 **2** 1/2 **3,0** *
B	1	2/3 **2** 1/3 **4,1** *	1

state 2

	L	R
T	0,0 *	1
B	1	0,0 *

Player 1 is the row player, while player 2 is the column player. An asterisked entry means that if this entry is reached then the game moves with probability 1 to an absorbing state which yields the players a payoff as indicated in the entry, while a non-asterisked entry means that if this entry is reached then the game moves to the state that is indicated by the entry (and the players receive no daily payoff). An entry of the form

1/2	**2**
1/2	**3,0** *

means that with probability 1/2 the game moves to an absorbing state, where the payoff for the players is $(3,0)$, and with probability 1/2 the game moves to state **2**.

Note that if player 2 plays a fully mixed stationary strategy then the game is bound to be eventually absorbed, whatever player 1 plays, hence the game satisfies the absorbing property.

One equilibrium payoff is $((2,0),(1,0))$. An ϵ-equilibrium strategy profile (for every $\epsilon > 0$) is:

- In state **1** the players play the mixed actions $(T,(1-\epsilon)L+\epsilon R)$.
- In state **2** both players play the mixed actions $(\frac{1}{2},\frac{1}{2})$.
- If any player plays an action which has probability 0 to be played, then both players play the pure actions (T,L) in both states forever (this part of the strategy serves as a punishment strategy).

It is easy to verify that no player can profit more than ϵ by any deviation, and that this strategy profile yields the players the desired payoff.

Another equilibrium payoff is $((2,4/17),(1,2/17)$. An ϵ-equilibrium strategy profile for this payoff is more complex. Let $n_1 \in \mathbf{N}$ and $\epsilon_1 < \epsilon$ such that $(1-\epsilon_1)^{n_1} = 1/2$. Define the following strategy profile:

- In state **2**, the players play the mixed actions $(\frac{1}{2},\frac{1}{2})$.
- Assume the game moves to state **1**. The players play as follows:
  - The players play the mixed actions $((1-\epsilon_1)T+\epsilon_1 B,(1-\epsilon_1)L+\epsilon_1 C)$. The players play these mixed actions until player 2 played the action C for n_1 times, or until both players played (B,C) at the same stage (and the game leaves state **1**).
  - If player 2 played the action C for n_1 times, then the players play the mixed actions $(T,(1-\epsilon)L+\epsilon R)$ until player 2 plays the action R (and the game leaves state **1**).
  - If any player plays an action which has probability 0 to be played, the players play the pure actions (T,L) in both states forever.

Note that if the players follow this strategy profile, then the game is bound to be eventually absorbed.

Assume that the players follow the above strategy profile, and let $g = (g_s^i)$ be the payoff that the players receive. Clearly no player can deviate and gain in state **2**, and $g_2^i = g_1^i/2$ for $i = 1, 2$.

Moreover, we have

$$g_1 = \frac{1}{2}\left(\frac{1}{3}(4,1)+\frac{2}{3}g_2\right)+\frac{1}{2}\left(\frac{1}{2}(3,0)+\frac{1}{2}g_2\right)$$

and therefore $g_1 = (2,4/17)$.

If player 2 deviates in state **1**, then he cannot gain more than ϵ, since after the punishment begins, his expected payoff is 0. The expected payoff of player 1 is

2, whether the game leaves state 1 through the entry (B, C) or through (T, R). Hence player 1 cannot profit by any deviation. Therefore this strategy profile is an ϵ-equilibrium, as desired.

3. The model and the main result

A **stochastic game** is a 5-tuple $G = (S, A, B, u, w)$ where

- S is a finite set of states.
- A and B are finite sets of actions available for players 1 and 2 respectively in every state.
- $u: S \times A \times B \to \mathbf{R}^2$ is the daily payoff function, $u^i(s, a, b)$ being the payoff for player i in state s when the two players play the actions a and b. We assume w.l.o.g. that $|u|$ is bounded by 1.
- $w: S \times A \times B \to \Delta(S)$ is the transition function, where $\Delta(S)$ is the space of all probability distributions over S.

The game is played as follows. Let $s_1 \in S$ be the initial state. At every stage n, the players are informed of past play including the current state $(s_1, a_1, b_1, s_2, a_2, b_2, \ldots, s_n)$, and player 1 (resp. player 2) chooses an action $a_n \in A$ (resp. $b_n \in B$). Then each player i receives a daily payoff $r_n^i = u^i(s_n, a_n, b_n)$, and a new state s_{n+1} is chosen according to $w(s_n, a_n, b_n)$.

Let $H^n = S \times (A \times B \times S)^n$ be the space of all histories of length n, $H_0 = \bigcup_{n \in \mathbf{N}} H^n$ be the space of all finite histories and $H = S \times (A \times B \times S)^{\mathbf{N}}$ be the space of all infinite histories. H is measurable with the σ-algebra generated by all the finite cylinders.

Definition 3.1: A **behavioral strategy** of player 1 (resp. player 2) is a function $\sigma: H_0 \to \Delta(A)$ (resp. $\tau: H_0 \to \Delta(B)$). A strategy σ is **stationary** if $\sigma(h_0)$ depends only on the last state of h_0. It is **fully mixed** if $\text{supp}(\sigma(h_0)) = A$ for every $h_0 \in H_0$. Symmetric definitions hold for player 2.

A **strategy profile** (or simply a **profile**) is a pair of strategies, one for each player. Any profile (σ, τ) and initial state s induce a probability measure over H. We denote this probability measure by $\text{Pr}_{s,\sigma,\tau}$ and expectation according to this measure by $\mathbf{E}_{s,\sigma,\tau}$.

Definition 3.2: A vector $g = (g_s^i)_{s \in S}^{i=1,2} \in \mathbf{R}^{2|S|}$ is an ϵ-**equilibrium payoff** if there exists a profile (σ, τ) and a positive integer $N \in \mathbf{N}$ such that for every initial state s, every strategy σ' of player 1 and every $n \geq N$,

$$(1) \qquad \mathbf{E}_{s,\sigma,\tau} \left(\frac{r_1^1 + \cdots + r_n^1}{n} \right) \geq g_s^1 - \epsilon \geq \mathbf{E}_{s,\sigma',\tau} \left(\frac{r_1^1 + \cdots + r_n^1}{n} \right) - 2\epsilon,$$

$$(2) \quad \mathbf{E}_{s,\sigma,\tau} \left(\liminf_{n \to \infty} \frac{r_1^1 + \cdots + r_n^1}{n} \right) \geq g_s^1 - \epsilon \geq \mathbf{E}_{s,\sigma',\tau} \left(\limsup_{n \to \infty} \frac{r_1^1 + \cdots + r_n^1}{n} \right),$$

and analogous inequalities hold for player 2, for every strategy τ'. The profile (σ, τ) is an ϵ-**equilibrium profile** for g. The payoff vector g is an **equilibrium payoff** if it is an ϵ-equilibrium payoff for every $\epsilon > 0$.

Definition 3.3: A state $s \in S$ is **absorbing** if $w_s(s, a, b) = 1$ for every pair of actions $(a, b) \in A \times B$. Otherwise it is **non-absorbing**.

The main result of the paper is the following.

THEOREM 3.4: *Every stochastic game with at most two non-absorbing states admits an equilibrium payoff.*

Let $T \subset S$ be the set of all absorbing states and $R = S \smallsetminus T$. Let $\theta = \min\{t \geq 1, s_t \in T\}$ be the absorption stage (the minimum of an empty set is $+\infty$); that is, the first stage in which the play reaches an absorbing state.

Definition 3.5: The game is **positive** if $u^2(s, a, b) > 0$ for every $s \in T$, and every $(a, b) \in A \times B$. It is **recursive** if $u^i(s, a, b) = 0$ for every $s \notin T$, every $(a, b) \in A \times B$ and every player $i = 1, 2$. It satisfies **the absorbing property** if for every fully mixed stationary strategy y of player 2, every strategy σ of player 1 and every state $s \in S$,
$$\Pr_{s,\sigma,y}(\theta < +\infty) = 1.$$

The following theorem follows from Vieille [8, 9]:

THEOREM 3.6: *If every positive recursive game with the absorbing property and at most n non-absorbing states admits an equilibrium payoff, then every stochastic game with at most n non-absorbing states admits an equilibrium payoff.*

Since existence of equilibrium payoffs in stochastic games with one non-absorbing state was solved by Vrieze and Thuijsman [11], to prove Theorem 3.4 it is sufficient to prove the following.

PROPOSITION 3.7: *Every positive recursive game with the absorbing property and two non-absorbing states admits an equilibrium payoff.*

The rest of the paper is devoted to prove this result.

From now on we fix a positive recursive game that satisfies the absorbing property. Note that any absorbing state is equivalent to a repeated game, in which equilibrium payoffs are known to exist. Since we are interested in the existence of equilibrium payoffs, we can assume w.l.o.g. that $u(s, \cdot, \cdot)$ is constant

over each $s \in S$, and we denote this constant value by u_s. Moreover, for this reason, the assumption that the available sets of actions are independent of the state is not restrictive (one can add to each player, if necessary, actions that lead to absorbing states with low absorbing payoff for that player, and high absorbing payoff for his opponent).

Since the game is recursive, $\lim_{n \to \infty} (r_1^i + \cdots + r_n^i)/n$ exists. Moreover, for recursive games, for every fixed pair of strategies (σ, τ) and every initial state s,

$$\mathbf{E}_{s,\sigma,\tau} \left(\lim_{n \to \infty} \frac{r_1^1 + \cdots + r_n^1}{n} \right) = \lim_{n \to \infty} \mathbf{E}_{s,\sigma,\tau} \left(\frac{r_1^1 + \cdots + r_n^1}{n} \right).$$

In particular, condition (1) in Definition 3.2 implies condition (2).

Let $c^1 = (c_s^1)_{s \in S}$ be the min-max value of player 1. This is the first coordinate of the (unique) equilibrium payoff of the zero-sum game that has the payoff function $(u^1, -u^1)$. The min-max value of player 2, $c^2 = (c_s^2)_{s \in S}$, is the second coordinate of the (unique) equilibrium payoff of the zero-sum game that has the payoff function $(-u^2, u^2)$. By Everett [3] or Mertens and Neyman [4], c^1 and c^2 exist.

Note that since the game is positive and satisfies the absorbing property, $c_s^2 > 0$ for every $s \in S$ (player 2, by playing some fully mixed stationary strategy, can guarantee a positive payoff, whatever player 1 plays).

We identify each $a \in A$ (resp. $b \in B$) with the probability distribution in $\Delta(A)$ (resp. $\Delta(B)$) that gives weight 1 to a (resp. b).

Let $X = (\Delta(A))^S$ and $Y = (\Delta(B))^S$. Every $x \in X$ and $y \in Y$ can be interpreted as a stationary strategy. We view each stationary strategy of player 1 as a vector in $\mathbf{R}^{|S| \cdot |A|}$ and each stationary strategy of player 2 as a vector in $\mathbf{R}^{|S| \cdot |B|}$. Whenever we use a norm, it is the maximum norm.

For every subset $C \subseteq S$ and every $(s, a, b) \in S \times A \times B$ we denote $w_C(s, a, b) = \sum_{s' \in C} w_{s'}(s, a, b)$. The multi-linear extension of w is denoted by w.

For every $(\alpha, \beta) \in \Delta(A) \times \Delta(B)$ and every function $g \colon S \to \mathbf{R}^2$ we define

$$(3) \qquad \psi_g(s, \alpha, \beta) = \sum_{s' \in S} w_{s'}(s, \alpha, \beta) g(s').$$

$\psi_g(s, \alpha, \beta)$ is the expected payoff for the players if the game is in state s, they play the mixed actions (α, β), and the continuation payoff is given by g. Note that for every fixed s, the function $\psi_\cdot(s, \cdot, \cdot)$ is multi-linear over $\Delta(A) \times \Delta(B) \times \mathbf{R}^{2|S|}$, and therefore continuous.

4. Sufficient conditions for existence of equilibrium payoff

In this section we provide two sets of sufficient conditions for existence of an equilibrium payoff in positive recursive games with the absorbing property.

Definition 4.1:   Let (x, y) be a stationary profile. A set $C \subseteq S$ is **stable under** (x, y) if $w_C(s, x, y) = 1$ for every $s \in C$.

Definition 4.2:   Let (x, y) be a stationary profile. The stationary profile (x', y') is a **perturbation** of (x, y), if $\text{supp}(x'_s) \supseteq \text{supp}(x_s)$ and $\text{supp}(y'_s) \supseteq \text{supp}(y_s)$. It is an ϵ-perturbation of (x, y) if it is a perturbation of (x, y), $\| x - x' \| < \epsilon$ and $\| y - y' \| < \epsilon$.

Definition 4.3:   Let (x, y) be a stationary profile. A set $C \subseteq R$ is **communicating** w.r.t. (x, y) if for every $s \in C$ there exists a stationary perturbation (x', y') of (x, y) such that C is stable under (x, y) and

$$\text{Pr}_{s', x', y'}(\exists n \in \mathbf{N} \text{ s.t. } s_n = s) = 1 \quad \forall s' \in C.$$

A set C is communicating if the players, by changing their stationary strategies a little, can reach from any state in C any other state in C, without leaving the set. Note that if there exists a perturbation (x', y') that satisfies Definition 4.3, then there also exists an ϵ-perturbation that satisfies it.

We denote by $\mathcal{C}(x, y)$ the collection of all communicating sets w.r.t. (x, y). Define for every communicating set $C \in \mathcal{C}(x, y)$ and every state $s \in C$

$$(4) \qquad
\begin{aligned}
A_s^1(C, y) &= \{a \in A \text{ s.t. } w_C(s, a, y_s) < 1\} \quad \text{and} \\
B_s^1(C, x) &= \{b \in B \text{ s.t. } w_C(s, x_s, b) < 1\}.
\end{aligned}$$

Those are all actions at s that cause the game to leave C with positive probability, when the opponent plays x_s or y_s.

Definition 4.4:   Let (x, y) be a stationary profile and $C \in \mathcal{C}(x, y)$. Every triplet (s, x'_s, y_s), where $s \in C$ and $x'_s \in \Delta(A_s^1(C, y))$, is an **exit of player 1** from C. Every triplet (s, x_s, y'_s), where $s \in C$ and $y'_s \in \Delta(B_s^1(C, x))$, is an **exit of player 2** from C. Every triplet $(s, x'_s, y'_s) \in C \times \Delta(A) \times \Delta(B)$ such that

$$\text{supp}(x_s) \cap \text{supp}(x'_s) = \text{supp}(y_s) \cap \text{supp}(y'_s) = \emptyset$$

is a **joint exit** from C if $w_C(s, x'_s, y'_s) < 1$ while $w_C(s, x_s, y'_s) = w_C(s, x'_s, y_s) = 1$.

A joint exit (s, x'_s, y'_s) is **pure** if $|\text{supp}(x'_s)| = |\text{supp}(y'_s)| = 1$. An exit (s, x'_s, y_s) of player 1 is **pure** if $|\text{supp}(x'_s)| = 1$. An exit (s, x_s, y'_s) of player 2 is **pure** if $|\text{supp}(y'_s)| = 1$.

We denote by $D_C^1(x,y), D_C^2(x,y)$ and $D_C^3(x,y)$ the sets of exits of player 1, player 2 and the joint exits from C respectively. Let

$$E_C(x,y) = D_C^1(x,y) \cup D_C^2(x,y) \cup D_C^3(x,y)$$

be the set of all exits from C and $E_C^0(x,y)$ be the set of all pure exits from C. We denote by $s(e), x(e)$ and $y(e)$ the three coordinates of each exit e.

For simplicity we write $s \in C(x,y)$ whenever $\{s\} \in C(x,y)$. In this case we write $E_s(x,y)$ instead of $E_{\{s\}}(x,y)$.

Recall that R is the set of non-absorbing states, and T is the set of absorbing states.

LEMMA 4.5: *Let (x,y) be a stationary profile such that $R \in C(x,y)$. Assume there exist an exit $e \in E_R(x,y)$ and a vector $g = (g_s)_{s \in S} \in \mathbf{R}^{2|S|}$ such that*

1. $g_s = u_s$ *for every $s \in T$, and $g_s = \psi_g(e)$ for each $s \in R$. In particular, g_s is a constant over R.*
2. $g_s^1 \geq \psi_c^1(s,a,y_s)$ *for every $s \in R$ and $a \in A$.*
3. $g_s^2 \geq \psi_c^2(s,x_s,b)$ *for every $s \in R$ and $b \in B$.*
4. *If $e \in D_R^1(x,y)$ then $g^1 = \psi_g^1(s(e),a,y_{s(e)})$ for every $a \in \mathrm{supp}(x(e))$.*
5. *If $e \in D_R^2(x,y)$ then $g^2 = \psi_g^2(s(e),x_{s(e)},b)$ for every $b \in \mathrm{supp}(y(e))$.*

Then g is an equilibrium payoff.

Note that condition 2 implies that $g_s^1 \geq c_s^1$ for every $s \in S$, and condition 3 implies that $g_s^2 \geq c_s^2$ for every $s \in S$.

The intuition is the following. Our goal is to construct an ϵ-equilibrium profile where the game leaves R (and is absorbed) through the exit e. By condition 1, the expected payoff for the players is $\psi_g(e)$. Since R is communicating, the players can play in such a way so that $s(e)$ is eventually reached with probability 1. By conditions 2 and 3 no player can profit by a detectable deviation. It might be the case that e is an exit of player 1, and that $|\mathrm{supp}(x(e))| > 1$. In this case, if the expected absorbing payoff of player 1 is *different* if he uses different actions in $\mathrm{supp}(x(e))$, he will prefer using some actions in $\mathrm{supp}(x(e))$ over others. Condition 4 asserts that this is not the case, and player 1 is indifferent between all the actions in the support of $x(e)$. Condition 5 is the analogous condition for player 2.

An analogous result was proved in Solan [6, Lemma 5.3] in the context of n-player absorbing games.

Proof: Let $\epsilon > 0$ and $\delta \in (0,\epsilon)$ be sufficiently small, and let (x',y') be an

ϵ-perturbation of (x, y) such that

$$\Pr_{s', x', y'}(\exists n \in \mathbf{N} \text{ s.t. } s_n = s(e)) = 1 \quad \forall s' \in R.$$

Since $R \in \mathcal{C}(x, y)$, such a perturbation exists.

Define a profile σ as follows:

- Whenever the game is in state $s(e)$ the players play the mixed action combination $((1 - \delta)x_{s(e)} + \delta x(e), (1 - \delta)y_{s(e)} + \delta y(e))$.
- Whenever the game is in a state $s \neq s(e)$ the players play the mixed action combination (x'_s, y'_s).

If the players follow σ then the game is bound to exit R through e, and to be absorbed. Hence, by condition 1, the expected payoff for the players is g_s, where s is the initial state.

In order to prevent the players from deviating, we choose $t_1 \in \mathbf{N}$ sufficiently large and add the following statistical tests at each stage t:

1. Both players check whether the realized action of their opponent is compatible with σ.

2. If $e \in D_R^2(x, y)$ and the game visited the state $s(e)$ at least t_1 times, then player 2 checks whether the distribution of the realized actions of player 1, whenever the game is in $s(e)$, is ϵ-close to $x_{s(e)}$. If $e \in D_R^1(x, y)$, then player 1 employs a symmetric test.

3. If $e \in D_R^3(x, y)$ and the game visited the state $s(e)$ at least t_1 times, then player 1 checks whether the realized actions of player 2 whenever the game is in $s(e)$, restricted to $\text{supp}(y(e))$, is ϵ-close to $y(e)$. Player 2 employs a symmetric test.

If a player fails one of these tests, this player is punished by his opponent with an ϵ-min-max strategy forever.

Since player 1 may profit by causing the game never to be absorbed (if $e \in D_R^1(x, y)$ and $\psi_g^1(e) < 0$), we add one more test. Let $t_2 \in \mathbf{N}$ be sufficiently large such that if no deviation is detected then absorption occurs before stage t_2 with probability greater than $1 - \epsilon$. We add the following test to σ:

4. At stage t_2 both players switch to an ϵ-min-max strategy.

The constants δ and t_1 are chosen in the following way. If $e \in D_R^2(x, y)$ then t_1 is chosen sufficiently large such that the probability of false detection of deviation in the second test is bounded by ϵ; that is

$$\Pr \left(\| \bar{X}_t - x(e) \| \quad \forall t > t_1 \right) > 1 - \epsilon/2$$

where $\bar{X}_t = \frac{1}{t} \sum_{j=1}^t X_j$ and $\{X_j\}$ are i.i.d. r.v. with distribution $x(e)$. If $e \in D_R^1(x, y)$ then t_1 is defined analogously.

The constant δ is chosen sufficiently small such that the probability of absorption in t_1 visits to $s(e)$ (i.e., before the second statistical test is employed) is at most ϵ; that is

$$(1 - \delta)^{t_1} > 1 - \epsilon.$$

If $e \in D_R^3(x,y)$ then δ and t_1 are chosen in such a way that the probability of false detection of deviation in the third test, as well as the probability of absorption before this test is employed, given only one player deviates, is at most ϵ. Since, whenever the game is in state $s(e)$, absorption occurs with probability $O(\delta^2)$, while perturbations occur with probability δ, if δ is sufficiently small then such t_1 exists. For a detailed analysis of this choice, one can refer to Solan [6].

Let n be sufficiently large. By conditions 2 and 3, the players cannot increase their expected average payoff in the first n stages by more than ϵ, using any detectable deviation. If ϵ is sufficiently small, then using a non-detectable deviation cannot increase the expected average payoff in the first n stages by more than 2ϵ. Hence g is an equilibrium payoff. ∎

LEMMA 4.6: *Let (x,y) be a stationary profile, and $g = (g_s)_{s \in S} \in \mathbf{R}^{2|S|}$. Assume the following conditions hold:*

1. *$g_s = u_s$ for every $s \in T$.*
2. *$g_s^1 \geq \psi_c^1(s, a, y_s)$ for every $s \in R$ and $a \in A$.*
3. *$g_s^2 \geq \psi_c^2(s, x_s, b)$ for every $s \in R$ and $b \in B$.*
4. *For every $s \in R$ one of the following two conditions hold:*
   (a) *either $s \notin C(x,y)$ and*
       i. *$g_s^1 = \psi_g^1(s, a, y_s)$ for every $a \in \mathrm{supp}(x_s)$,*
       ii. *$g_s^2 = \psi_g^2(s, x_s, b)$ for every $b \in \mathrm{supp}(y_s)$;*
   (b) *or $s \in C(x,y)$ and there exist two exits $e_1, e_2 \in E_s(x,y)$ and $\alpha \in [0,1]$ that satisfy the following:*
       i. *$\psi_g^1(e_j) = g_s^1$ for each $j = 1, 2$, and $g_s^2 = \alpha \psi_g^2(e_1) + (1-\alpha)\psi_g^2(e_2)$,*
       ii. *if $e_j \in D_s^1(x,y)$ then $g_s^1 = \psi_g^1(s, a, y_s)$ for every $a \in \mathrm{supp}(x(e_j))$,*
       iii. *if $e_j \in D_s^2(x,y)$ then*

       $$g_s^2 \geq \psi_g^2(s, x_s, b_1) = \psi_g^2(s, x_s, b_2) \geq \psi_c^2(s, x_s, b_3)$$

       for every $b_1, b_2 \in \mathrm{supp}(y(e_j))$ and $b_3 \in B$,
       iv. *at most one of e_1 and e_2 is an exit of player 2.*
5. *The Markov chain over S whose transition law is induced by (x_s, y_s) for every $s \notin C(x,y)$ and by $\alpha e_1 + (1-\alpha)e_2$ for every $s \in C(x,y)$ is absorbing (i.e. an absorbing state is reached with probability 1).*

Then g is an equilibrium payoff.

Note that condition 4(b).iv is redundant, since, if $e_1, e_2 \in D_s^2(x, y)$, then, by 4(b).i and 4(b).iii, $\psi_g^2(e_1) = \psi_g^2(e_2)$, one can define (with abuse of notations) $e_1' = \alpha e_1 + (1 - \alpha)e_2$, and then condition 4(b) is satisfied with e_1' and $\alpha' = 1$.

The intuition here is as follows. By condition 4, every non-absorbing state is either transient under (x, y), or there exist two exits that satisfy various conditions. We will devise a profile under which, in transient states, the players follow (x, y), whereas in non-transient states, the play leaves the state through these two exits.

By condition 5 the game eventually reaches an absorbing states, and by conditions 1, 4(a) and 4(b).i, the expected payoff for the players is g. By conditions 2 and 3 no player can profit by playing an action he is not supposed to play. By condition 4(a) no player can profit by any deviation in transient states. As we will see, condition 4(b) implies the same in non-transient states.

Proof: Let $\epsilon > 0$ and $\delta \in (0, \epsilon)$ be sufficiently small. For every $s \in \mathcal{C}(x, y)$ we consider the two exits e_1, e_2 of condition 4(b). We assume w.l.o.g. that $\psi_g^2(e_1) \geq \psi_g^2(e_2)$. In particular, by condition 4(b).i, $\psi_g^2(e_1) \geq g_s^2 \geq \psi_g^2(e_2)$, and by condition 4(b).iii it follows that $e_1 \notin D_s^2(x, y)$.

Define a profile σ as follows:

a) Whenever the game is in a state $s \in R$ such that $s \notin \mathcal{C}(x, y)$, the players play (x_s, y_s).

b) Whenever the game is in a state $s \in R$ such that $s \in \mathcal{C}(x, y)$, the players play as follows:

- Play $((1 - \delta)x_s + \delta x(e_1), (1 - \delta)y_s + \delta y(e_1))$ for n stages or until an action combination in the support of $(x(e_1), y(e_1))$ is played, where n satisfies:
  - $(1 - \delta)^n = 1 - \alpha$ if e_1 is a unilateral exit,
  - $(1 - \delta^2)^n = 1 - \alpha$ if e_1 is a joint exit.
- Play $((1-\delta)x_s + \delta x(e_2), (1-\delta)y_s + \delta y(e_2))$ until an action combination in the support of $(x(e_2), y(e_2))$ is played.

If an action combination in the support of $(x(e_j), y(e_j))$ is played, but the game remains in s, the players repeat step (b).

By condition 5, if the players follow σ then the game is bound to be eventually absorbed, and by conditions 1, 4(a) and 4(b).i the expected payoff for the players is g_s, where s is the initial state.

In order to prevent the players from playing actions which are not compatible with σ, the players check, as in Lemma 4.5, that the realized action combination

that is played is compatible with σ.

In order to prevent other deviations, we add the following statistical tests. Assume that the game moves to a state $s \in \mathcal{C}(x,y)$ at stage t_0. Let $t_1, t_2 \in \mathbf{N}$ be sufficiently large, and let e_1, e_2 be the two exits from $\{s\}$ of condition 4(b). Each player checks his opponent behavior as follows at every stage t such that $t_0 < t < t_0 + n$:

1. If $e_1 \in D_s^1(x,y)$ and the game has visited s for at least t_1 times, then player 1 checks whether the distribution of the realized actions of player 2 at stages $t_0, t_0 + 1, \ldots, t - 1$ is ϵ-close to x_s.

2. If $e_1 \in D_s^2(x,y)$, a symmetric test is done by player 2.

3. If $e_1 \in D_s^3(x,y)$ and the game has visited s for at least t_2 times, then both players check whether the realized actions of their opponent at stages $t_0, t_0 + 1, \ldots, t - 1$, restricted to $\mathrm{supp}(x(e_1))$ and $\mathrm{supp}(y(e_1))$, is ϵ-close to $x(e_1)$ and $y(e_1)$ respectively.

If a player fails this test at a stage $t_0 \leq t \leq t_0 + n$, this player is punished with an ϵ-min-max strategy forever.

If no deviation is detected before stage $t_0 + n$, then each player begins to check, in a similar way, if his opponent continues to follow σ, until the game leaves the state s (i.e. replace e_1 by e_2 in the statistical tests).

Since it might be the case that both exits are unilateral exits of player 1, and that $g_s^1 < 0$, so that player 1 gains if the game is never absorbed, we add the following test. Let t_3 be sufficiently large such that if no deviation is detected then leaving s occurs in t_3 stages with probability greater than $1 - \epsilon$. As in the proof of Lemma 4.5, at stage $t_0 + t_3$ both players switch to an ϵ-min-max strategy.

The constants t_1, t_2 and δ are chosen, as in the proof of Lemma 4.5, in such a way that no player can profit more than ϵ by any non-detectable deviation, and the probability of false detection of deviation is bounded by ϵ.

Let $m_n = \#\{t < n: s_t \neq s_{t-1}\}$ be the number of times the state process changes values until stage n. Recall that θ is the stage of absorption. By condition 5 the induced Markov chain is absorbing, hence, if no deviation is detected, there is some $K > 0$ such that $\Pr(\theta < +\infty, m_\theta < K) > 1 - \epsilon$. By condition 4(b), in any visit to a state $s \in \mathcal{C}(x,y)$ the players may profit at most ϵ, while by conditions 2 and 3 no player can profit more than ϵ by deviating in a detectable way. It follows that if n is sufficiently large, no player can increase his expected average payoff by more than $(K + 1)\epsilon$ using any deviation. In particular, g is an equilibrium payoff.   ∎

5. Preliminary results

A stationary profile (x, y) is **absorbing** if $\mathbf{Pr}_{s,x,y}(\theta < +\infty) = 1$ for every $s \in S$; that is, the game eventually reaches an absorbing state with probability 1.

For every state $s \in S$, let $v_s^i(x, y)$ be the expected undiscounted payoff for player i if the initial state is s and the players play the profile (x, y):

$$v_s^i(x, y) = \mathbf{E}\left(1_{\theta < +\infty} u_{s_\theta}\right).$$

The function $v(x, y) = (v_s(x, y))_{s \in S} \in \mathbf{R}^{2|S|}$ is harmonic over S w.r.t. the transition $p_{s,s'} = w_{s'}(s, x_s, y_s)$. If (x, y) is absorbing then $v(x, y)$ is the unique solution of the following system of linear equations:

$$(5) \qquad \begin{aligned} \xi_s &= u_s & \forall s \in T, \\ \xi_s &= \psi_\xi(s, x_s, y_s) & \forall s \in R. \end{aligned}$$

LEMMA 5.1: *Let (x, y) be an absorbing stationary profile. Let $g \colon S \to \mathbf{R}^2$ be such that $\psi_g^2(s, x_s, y_s) \leq g_s^2$ for every $s \in R$ and $g_s^2 = u_s^2$ for every $s \in T$. Then $v_s^2(x, y) \leq g_s^2$ for every $s \in S$*

Proof: $v^2(x, y)$ is a harmonic function and g^2 is a sub-harmonic function over S that have the same values over T. Hence $v^2(x, y) - g^2$ is a super-harmonic function that vanishes over T. Since (x, y) is absorbing, $v^2(x, y) - g^2$ is non-positive. ∎

COROLLARY 5.2: *Let x be a stationary strategy of player 1. Let $g \colon S \to \mathbf{R}^2$ satisfy, for every stationary strategy y of player 2, $g_s^2 \geq \psi_g^2(s, x_s, y_s)$ for every $s \in R$ and $g_s^2 = u_s^2$ for every $s \in T$. Then $c_s^2 \leq g_s^2$ for every $s \in S$.*

Proof: Since the game is a positive recursive game with the absorbing property, the best reply of player 2 against the stationary strategy x is a stationary strategy y such that (x, y) is absorbing. By Lemma 5.1, $v_s^2(x, y) \leq g_s^2$ for every stationary strategy y such that (x, y) is absorbing and $s \in S$. Hence $c_s^2 \leq g_s^2$. ∎

A symmetric proof proves the following lemma:

LEMMA 5.3: *Let y be a fully mixed stationary strategy of player 2. Let $g \colon S \to \mathbf{R}^2$ satisfy, for every stationary strategy x of player 1, $\psi_g^1(s, x_s, y_s) \leq g_s^1$ for every $s \in R$ and $g_s^1 = u_s^1$ for every $s \in T$. Then $c_s^1 \leq g_s^1$ for every $s \in S$.*

6. The ϵ-approximating game

6.1 THE GAME.   Let $\epsilon^\star = 1/|B|$. For every $\epsilon \in (0, \epsilon^\star)$ define the set

$$
(6) \qquad Y_s(\epsilon) = \left\{ y_s \in \Delta(B) \mid \sum_{b \in J} y_s^b \geq \epsilon^{|B| - |J|} \quad \forall J \subseteq B \right\}.
$$

Let $Y(\epsilon) = \times_{s \in S} Y_s(\epsilon)$. Every stationary strategy $y \in Y(\epsilon)$ is fully mixed. Since the game satisfies the absorbing property, the payoff function $v(x, y)$ is continuous over $X \times Y(\epsilon)$.

Define the ϵ-approximating game $G'(\epsilon)$ as a positive recursive game with the absorbing property (S, A, B, w, u), where player 2 is restricted to strategies τ such that $\tau(h) \in Y_s(\epsilon)$, for every finite history h (s is the last state of h), and player 1 is not restricted.

6.2 EXISTENCE OF A STATIONARY EQUILIBRIUM.   Note that X and $Y(\epsilon)$ (for every $\epsilon \in (0, \epsilon^\star)$) are non-empty, convex and compact sets. Define the correspondence $\phi^1_{s, \epsilon} \colon X \times Y(\epsilon) \to \Delta(A)$ by

$$
(7) \qquad \phi^1_{s, \epsilon}(x, y) = \operatorname*{argmax}_{x'_s \in \Delta(A)} \psi^1_{v(x, y)}(s, x'_s, y_s);
$$

that is, player 1 maximizes his payoff locally — in every state s he chooses a mixed action that maximizes his expected payoff if the initial state is s, player 2 plays the mixed action y_s, and the continuation payoff is given by $v(x, y)$. Let $\phi^1_\epsilon = \times_{s \in S} \phi^1_{s, \epsilon}$.

LEMMA 6.1: *The correspondence ϕ^1_ϵ has non-empty convex values, and it is upper semi-continuous.*

Proof: Since $\psi^1_{v(x, y)}(s, x', y_s)$ is linear in x'_s for every fixed (s, x, y), ϕ^1_ϵ has non-empty and convex values. By the continuity of v^1 over the compact set $X \times Y(\epsilon)$ it follows that ϕ^1_ϵ is upper semi-continuous.   ∎

Define the correspondence $\phi^2_{s, \epsilon} \colon X \times Y(\epsilon) \to Y_s(\epsilon)$ by

$$
(8) \qquad \phi^2_{s, \epsilon}(x, y) = \operatorname*{argmax}_{y'_s \in Y_s(\epsilon)} \psi^2_{v(x, y)}(s, x_s, y'_s).
$$

Let $\phi^2_\epsilon = \times_{s \in S} \phi^2_{s, \epsilon}$.   As in Lemma 6.1, since $Y_s(\epsilon)$ is not empty and convex whenever $\epsilon \in (0, \epsilon^\star)$, we have:

LEMMA 6.2: *The correspondence ϕ_ϵ^2 has non-empty convex values, and it is upper semi-continuous.*

Define the correspondence $\phi_\epsilon \colon X \times Y(\epsilon) \to X \times Y(\epsilon)$ by

$$\phi_\epsilon(x, y) = \phi_\epsilon^1(y) \times \phi_\epsilon^2(x).$$

By Lemmas 6.1, 6.2 and by Kakutani's fixed point Theorem we get:

LEMMA 6.3: *For every $\epsilon \in (0, \epsilon^\star)$ there exist $(x(\epsilon), y(\epsilon)) \in X \times Y(\epsilon)$ that is a fixed point of the correspondence ϕ_ϵ.*

6.3 THE BEHAVIOR AS $\epsilon \to 0$. Since the state and action spaces are finite, there exist sequences $\{\epsilon_n\}_{n \in \mathbf{N}}$ of positive real numbers and $\{(x(n), y(n))\}_{n \in \mathbf{N}}$ of stationary profiles such that:

C.1. $\epsilon_n \to 0$, and $(x(n), y(n)) \in X \times Y(\epsilon_n)$ is a fixed point of ϕ_{ϵ_n} for every $n \in \mathbf{N}$.

C.2. For every $s \in S$, $\text{supp}(x_s(n))$ and $\text{supp}(y_s(n))$ are independent of n.

In the sequel, we need that various sequences that depend on $\{x(n)\}$ and $\{y(n)\}$ have a limit. The number of those sequences is finite, hence, by taking a subsequence, we will assume that the limits exist.

Remark: Using the method of Bewley and Kohlberg [1], it can be proven that we can choose for every $\epsilon > 0$ a fixed point $(x(\epsilon), y(\epsilon))$ of ϕ_ϵ such that x and y, as functions of ϵ, are semi-algebraic functions (they have a Taylor expansion in fractional powers of ϵ), hence **C.1** and **C.2** hold for *every* ϵ sufficiently small (and not only for a sequence $\{\epsilon_n\}$). In particular, all the limits that we use in the sequel exist.

We denote for every $n \in \mathbf{N}$ and $s \in S$, $d_s(n) = v_s(x(n), y(n))$. Denote $d_s(\infty) = \lim_{n \to \infty} d_s(n)$, $x_s(\infty) = \lim_{n \to \infty} x_s(n)$ and $y_s(\infty) = \lim_{n \to \infty} y_s(n)$.

LEMMA 6.4: *Let $s_1 \in S$ and $b_1, b_2 \in B$. If $\lim_{n \to \infty} y_{s_1}^{b_1}(n)/y_{s_1}^{b_2}(n) < \infty$, then for every n sufficiently large*

$$\psi_{v(x(n), y(n))}^2(s_1, x_{s_1}(n), b_1) \le \psi_{v(x(n), y(n))}^2(s_1, x_{s_1}(n), b_2).$$

Proof: Assume that the lemma is not true. Then, by taking a subsequence, $\psi_{v(x(n), y(n))}^2(s_1, x_{s_1}(n), b_1) > \psi_{v(x(n), y(n))}^2(s_1, x_{s_1}(n), b_2)$ for every $n \in \mathbf{N}$. Define for every n the stationary strategy $y'(n)$ for player 2 as follows:

$$y_s'^b(n) = \begin{cases} y_s^{b_2}(n)/2, & (s, b) = (s_1, b_2), \\ y_s^{b_1}(n) + y_s^{b_2}(n)/2, & (s, b) = (s_1, b_1), \\ y_s^b(n), & \text{otherwise.} \end{cases}$$

Let us verify that $y'_s(n) \in Y_s(\epsilon_n)$ for every n sufficiently large. Otherwise, by taking a subsequence, there exists a set $J \subseteq B$ such that $b_2 \in J$, $b_1 \notin J$ and

$$(9) \qquad \sum_{b \in J} y'^b_s(n) < \epsilon_n^{|B|-|J|} \quad \forall n \in \mathbf{N}.$$

In particular, $\lim y^{b_2}_s(n)/\epsilon_n^{|B|-|J|} < \infty$. By the assumption, $\lim y^{b_1}_s(n)/\epsilon_n^{|B|-|J|} < \infty$ as well. Since $\sum_{b \in J} y^b_s(n) + y^{b_1}_s(n) \geq \epsilon_n^{|B|-|J|-1}$ it follows that there exists $b \in J \setminus \{b_2\}$ such that $\lim y^b_s(n)/\epsilon_n^{|B|-|J|-1} > 0$ — a contradiction to (9).

However

$$\psi^2_{d(n)}(s, x_s(n), y'_s(n)) - \psi^2_{d(n)}(s, x_s(n), y_s(n)) =$$
$$\left(\psi^2_{d(n)}(s, x_s(n), b_1) - \psi^2_{d(n)}(s, x_s(n), b_2) \right) y^{b_2}_{s_1}(n)/2 > 0,$$

a contradiction to **C.1**. ∎

By applying Lemma 6.4 in both directions, and taking the limit as $n \to \infty$, we conclude that if player 2 plays two actions with approximately the same frequencies then the corresponding limits of his continuation payoffs are equal:

COROLLARY 6.5: *Let $b_1, b_2 \in B$ and $s \in S$. If $\lim_{n \to \infty} y^{b_1}_s(n)/y^{b_2}_s(n) \in (0, \infty)$ then*

$$\psi^2_{d(\infty)}(s, x_s(\infty), b_1) = \psi^2_{d(\infty)}(s, x_s(\infty), b_2).$$

COROLLARY 6.6: *For every $b \in \mathrm{supp}(y_s(\infty))$*

$$\psi^2_{d(\infty)}(s, x_s(\infty), b) = d^2_s(\infty).$$

Proof:

$$d^2_s(\infty) = \lim_{n \to \infty} d^2_s(n)$$
$$= \lim_{n \to \infty} \psi^2_{d(n)}(s, x_s(n), y_s(n))$$
$$= \psi^2_{d(\infty)}(s, x_s(\infty), y_s(\infty))$$
$$= \sum_{b \in \mathrm{supp}(y_s(\infty))} y^b_s(\infty) \psi^2_{d(\infty)}(s, x_s(\infty), b).$$

The result now follows from Corollary 6.5. ∎

By Lemma 6.4, Corollary 6.6 and the continuity of ψ it follows that

$$(10) \qquad \psi^2_{d(\infty)}(s, x_s(\infty), b) \leq d^2_s(\infty) \quad \forall (s, b) \in S \times B.$$

By Corollary 5.2 and (10) it follows that

$$(11) \qquad c^2_s \leq d^2_s(\infty) \quad \forall s \in S.$$

In particular, (10) and (11) yield

$$(12) \qquad \psi_c^2(s, x_s(\infty), b) \le d_s^2(\infty) \quad \forall (s, b) \in S \times B.$$

By **C.1**, we get that for every n

$$(13) \qquad \psi_{d(n)}^1(s, a, y_s(n)) \le d_s^1(n) \quad \forall (s, a) \in S \times A,$$

and equality holds whenever $a \in \text{supp}(x_s(n))$ (which is independent of n by **C.2**). Taking a limit in (13) as $n \to \infty$ we get

$$(14) \qquad \psi_{d(\infty)}^1(s, a, y_s(\infty)) \le d_s^1(\infty) \quad \forall (s, a) \in S \times A,$$

and equality holds whenever $a \in \text{supp}(x_s(n))$.

By Lemma 5.3 and (13), $c_s^1 \le d_s^1(n)$ for every $s \in S$ and every $n \in \mathbf{N}$, and by taking the limit as $n \to \infty$,

$$(15) \qquad c_s^1 \le d_s^1(\infty) \quad \forall s \in S.$$

Therefore,

$$(16) \qquad \psi_c^1(s, a, y_s(\infty)) \le d_s^1(\infty) \quad \forall (s, a) \in S \times A.$$

To summarize, we have asserted that $d_s^i(\infty)$ is greater than the min-max value of player i (Eqs. (11) and (15)), and that no player can receive more than $d_s(\infty)$ by playing any action in any state s and then be punished with his min-max value (Eqs. (12) and (16)).

7. Existence of an equilibrium payoff

In this section we prove Proposition 3.7. This is done by showing that the conditions of either Lemma 4.5 or Lemma 4.6 hold. Denote the two non-absorbing states by $R = \{s_1, s_2\}$.

7.1 EXITS FROM A STATE.   Fix a state $s \in S$ such that $w_s(s, x_s(\infty), y_s(\infty)) = 1$. In particular, $s \in \mathcal{C}(x(\infty), y(\infty))$. Since the game satisfies the absorbing property, $E_s(x(\infty), y(\infty)) \ne \emptyset$.

Let ρ_n^s be the probability distribution over $E_s^0(x(\infty), y(\infty))$ that is induced by $(x(n), y(n))$. Formally, we define for every $e \in E_s^0(x(\infty), y(\infty))$

$$\bar{\rho}_n^s(e) = \begin{cases} x_s^a(n), & e = (s, a, y(\infty)) \in D_s^1(x(\infty), y(\infty)) \\ y_s^b(n), & e = (s, x(\infty), b) \in D_s^2(x(\infty), y(\infty)) \\ x_s^a(n) \times y_s^b(n), & e = (s, a, b) \in D_s^3(x(\infty), y(\infty)) \end{cases}$$

and $\rho_n^s(e) = \bar{\rho}_n^s(e)/\sum_{e\in E_s^0(x(\infty),y(\infty))} \bar{\rho}_n^s(e)$.

Since $y(n)$ is fully mixed and the game satisfies the absorbing property, ρ_n^s is a well defined probability distribution. Define $\rho_\infty^s \overset{\text{def}}{=} \lim_{n\to\infty} \rho_n^s$.

Every pair of actions $(a,b) \in A \times B$ such that $w_s(s,a,b) < 1$ is either in the support of some exit in $E_s^0(x(\infty),y(\infty))$, or there exists a pair (a_0,b_0) which is in the support of some exit in $E_s^0(x(\infty),y(\infty))$ such that

$$\lim_{n\to\infty} \frac{x_s^a(n)y_s^b(n)}{x_s^{a_0}(n)y_s^{b_0}(n)} = 0.$$

Since $d_s(n) = \psi_{d(n)}(s,x_s(n),y_s(n))$, by taking the limit as $n \to \infty$ we have

(17)
$$d_s(\infty) = \sum_{e\in E_s^0(x(\infty),y(\infty))} \rho_\infty^s(e)\psi_{d(\infty)}(e).$$

That is, the average continuation payoff over the exits is equal to $d_s(\infty)$.

Since $d_s^1(n) = \psi_{d(n)}^1(s,a,y_s(n))$, by summing over all $a \in \text{supp}(x_s(\infty))$ and taking the limit as $n \to \infty$ we have

(18)
$$\left(\sum_{b\,:\,(s,x_s(\infty),b)\in D_s^2(x(\infty),y(\infty))} \rho_\infty^s(s,x_s(\infty),b)\right) d_s^1(\infty) =$$
$$\sum_{b\,:\,(s,x_s(\infty),b)\in D_s^2(x(\infty),y(\infty))} \rho_\infty^s(s,x_s(\infty),b)\psi_{d(\infty)}^1(s,x_s(\infty),b).$$

Let $a \notin \text{supp}(x_s(\infty))$ such that $x_s^a(n) > 0$ for every $n > 0$. Then, as in (18),

(19)
$$\left(\sum_{b\,:\,(s,a,b)\in D_s^3(x(\infty),y(\infty))} \rho_\infty^s(s,a,b)\right) d_s^1(\infty) =$$
$$\sum_{b\,:\,(s,a,b)\in D_s^3(x(\infty),y(\infty))} \rho_\infty^s(s,a,b)\psi_{d(\infty)}^1(s,a,b).$$

Eq. (18) means that the average continuation payoff of player 1 over the unilateral exits of player 2 is equal to $d_s^1(\infty)$, whereas Eq. (19) means that for every action $a \notin \text{supp}(x_s(\infty))$, the average continuation payoff of player 1, restricted to pure joint exits e with $x(e) = a$, is equal to $d_s^1(\infty)$. Together, these lemmas imply some indifference property from the point of view of player 1. This should not surprise us, since player 1 is not restricted in the approximating game.

LEMMA 7.1: *There exist two exits $e_1, e_2 \in E_s(x(\infty),y(\infty))$ and $\alpha \in [0,1]$ such that*

1. $d_s^1(\infty) = \psi_{d(\infty)}^1(e_j)$ for $j = 1, 2$.
2. $d_s^2(\infty) = \alpha \psi_{d(\infty)}^2(e_1) + (1 - \alpha)\psi_{d(\infty)}^2(e_2)$.
3. At most one of e_1, e_2 is an exit of player 2.
4. If there exists $e \in \mathrm{supp}(\rho_\infty^s)$ such that $w_T(e) > 0$, then $w_T(e_1) + w_T(e_2) > 0$.
5. If $e_j \in D_s^1(x(\infty), y(\infty))$, then $d_s^1(\infty) = \psi_{d(\infty)}^1(s, a, y_s(\infty))$ for every $a \in \mathrm{supp}(x_s(e_j))$.
6. If $e_j \in D_s^2(x(\infty), y(\infty))$, then $\psi_{d(\infty)}^2(s, x_s(\infty), b) = \psi_{d(\infty)}^2(s, x_s(\infty), b')$ for every $b, b' \in \mathrm{supp}(y_s(e_j))$.

Proof: For every $a \in A$ such that $w_s(s, a, y_s(\infty)) > 0$, define $\tilde{y}_s(a) = y_s(\infty)$. By (14), $\psi_g^1(s, a, \tilde{y}_s(a)) = d_s^1(\infty)$ for every such a.

For every $a \in A$ such that $w_s(s, a, y_s(\infty)) = 0 < w_s(s, a, y_s(n))$ for every $n \in \mathbf{N}$, define $\tilde{y}_s(a) \in \Delta(B)$ by

$$\tilde{y}_s^b(a) = \begin{cases} 0, & w_s(s, a, b) = 1, \\ \rho_\infty^s(s, a, b) / \sum_{b: \, w_s(s, a, b) < 1} \rho_\infty^s(s, a, b), & \text{otherwise.} \end{cases}$$

By (19), $\psi_g^1(s, a, \tilde{y}_s(a)) = d_s^1(\infty)$ for every such a.

If there exist $a_1, a_2 \notin \mathrm{supp}(x_s(\infty))$ for which $\tilde{y}_s(a_1)$ and $\tilde{y}_s(a_2)$ are defined such that $\psi_g^2(s, a_1, \tilde{y}_s(a_1)) \geq d_s^2(\infty) \geq \psi_g^2(s, a_2, \tilde{y}_s(a_2))$, we are done. Indeed, define $e_j = (s, a_j, \tilde{y}_s(a_j))$ for $j = 1, 2$ and choose $\alpha \in [0, 1]$ that satisfies 2. Note that in this case, one can choose such a_1 and a_2 so that (4) holds.

Otherwise, either for every $a \notin \mathrm{supp}(x_s(\infty))$, $w_s(s, a, y_s(n)) = 1$ for every n, or for every $a \notin \mathrm{supp}(x_s(\infty))$, $\psi_g^2(s, a_1, \tilde{y}_s(a_1)) > d_s^2(\infty)$. In both cases, (17) implies that $w_s(s, x_s(\infty), y_s(n)) < 1$ for every n. Hence one can define a probability distribution $\tilde{y}_s \in \Delta(B)$ by

$$\tilde{y}_s^b = \begin{cases} 0, & w_s(s, x_s(\infty), b) = 1, \\ \rho_\infty^s(s, x_s(\infty), b) / \sum_{b: \, w_s(s, x_s(\infty), b) < 1} \rho_\infty^s(s, x_s(\infty), b), & \text{otherwise.} \end{cases}$$

By (10), $\psi_g^2(s, x_s(\infty), \tilde{y}_s) \leq d_s^2(\infty)$.

If $\psi_g^2(s, x_s(\infty), \tilde{y}_s) = d_s^2(\infty)$, then by (18), $e_1 = (s, x_s(\infty), \tilde{y}_s)$ and $\alpha = 1$ satisfy the conclusion.

If $\psi_g^2(s, x_s(\infty), \tilde{y}_s) < d_s^2(\infty)$, then by (17), there exists $a_1 \notin \mathrm{supp}(x_s(\infty))$ such that $\tilde{y}_s(a_1)$ is defined. In particular, $\psi_g^2(s, a_1, \tilde{y}_s(a_1)) > d_s^2(\infty)$. Moreover, if there is a $a \notin \mathrm{supp}(x_s(\infty))$ with $w_T(s, a, \tilde{y}_s(a)) > 0$, we can assume it is a_1. By defining $e_1 = (s, a_1, \tilde{y}_s(a_1))$, $e_2 = (s, x_s(\infty), \tilde{y}_s)$ and $\alpha \in [0, 1]$ properly, the result follows, where conditions 1 and 5 follow from (18) and (19).

In both cases, condition 6 follows from Corollary 6.5. ∎

7.2 PROOF OF PROPOSITION 3.7.   In this section we prove Proposition 3.7.
Consider the following two conditions:

A.1. R is communicating under $(x(\infty), y(\infty))$.

A.2. For every $s \in R$ such that $s \in \mathcal{C}(x(\infty), y(\infty))$, and every $e \in$
$E_s^0(x(\infty), y(\infty))$ such that $\rho_\infty^s(e) > 0$, we have $w_T(e) = 0$.

We will prove that if conditions **A** hold then the conditions of Lemma 4.5 hold,
while if they do not hold then the conditions of Lemma 4.6 hold.

LEMMA 7.2: *If conditions **A** do not hold then the conditions of Lemma 4.6 hold*
w.r.t. $(x(\infty), y(\infty))$.

Proof: Define $g = d(\infty)$.   We prove that the conditions of Lemma 4.6 hold
w.r.t. $(x(\infty), y(\infty))$ and g.   Condition 1 holds since $d_s(n) = u_s$ for every $s \in T$
and every $n \in \mathbf{N}$.   Condition 2 follows from (16) while condition 3 follow from
(12).   Condition 4(a) follows from (14) and Corollary 6.6, whereas condition 4(b)
follows from Lemma 7.1.   Since conditions **A** do not hold, it follows by Lemma
7.1 that condition 5 of Lemma 4.6 holds.   ∎

LEMMA 7.3: *If conditions **A** hold then the conditions of Lemma 4.5 hold w.r.t.*
$(x(\infty), y(\infty))$.

Proof: To prove that the conditions of Lemma 4.5 hold, we need to find an
exit e from R that satisfies various conditions.   In particular, it should give high
payoff for both players.

As in section 7.1, $(x(n), y(n))$ induce a probability distribution over the exits in
$E_R(x(\infty), y(\infty))$, and we could have looked for some exit whose limit probability
is positive.

We choose a different path.   We first identify the state where the payoff for
player 2 is higher: $d_{s_1}^2(n) \geq d_{s_2}^2(n)$ for every n.   We then look for the actions of
player 2 that cause the game to be absorbed from s_1 with positive probability,
and player 2 plays as often as he can.   Since player 2 is restricted, and since in s_1
the payoff of player 2 is higher, if absorption occurs through those actions player
2 gets at least $d_{s_1}^2(\infty)$.   Since player 1 is indifferent between his various actions,
it will follow that the exit that corresponds to those actions of player 2 is the
desired one.   The limit probability of this exit might vanish.

We now turn to the formal proof.

By taking a subsequence and exchanging the names of s_1 and s_2 if necessary,
we can assume that exactly one of the following holds:

B.1. Either $d_{s_1}^2(n) > d_{s_2}^2(n)$ for every $n \in \mathbf{N}$.

B.2. Or $d^2_{s_1}(n) = d^2_{s_2}(n)$ and $d^1_{s_1}(n) > d^1_{s_2}(n)$ for every $n \in \mathbf{N}$.

B.3. Or $d^2_{s_1}(n) = d^2_{s_2}(n)$, $d^1_{s_1}(n) = d^1_{s_2}(n)$ and $w_T(s_1, x_{s_1}(n), y_{s_1}(n)) > 0$ for every $n \in \mathbf{N}$.

Recall that $\operatorname{supp}(x_s(n))$ and $\operatorname{supp}(y_s(n))$ are independent of n for every $s \in S$.

STEP 1:   Definition of e.

Note that $w_T(s_1, x_{s_1}(n), y_{s_1}(n)) > 0$ for every $n \in \mathbf{N}$. Otherwise, it follows that $d_{s_1}(n) = d_{s_2}(n)$ for every n, which contradicts **B.1**, **B.2** and **B.3**.

Let $B^\star$ be the set of all actions $b \in B$ such that

- $w_T(s_1, x_{s_1}(n), b) > 0$ for every n.
- $\lim_{n \to \infty} y^b_{s_1}(n)/y^{b'}_{s_1}(n) > 0$ for every b' such that $w_T(s_1, x_{s_1}(n), b') > 0$.

$B^\star$ contains all actions of player 2 that are absorbing, and played most often. Since $w_T(s_1, x_{s_1}(n), y_{s_1}(n)) > 0$ for every n, it follows that $B^\star \neq \emptyset$. Let $y^\star_{s_1}(n)$ be the probability distribution induced by $y_{s_1}(n)$ over $B^\star$, and let $y^\star_{s_1}(\infty)$ be the limit distribution. By the definition of $B^\star$, $\operatorname{supp}(y^\star_{s_1}(\infty)) = \operatorname{supp}(y_{s_1}(n))$ for every n.

Let $A^\star$ be the set of all actions $a \in \operatorname{supp}(x_{s_1}(n))$ such that $w_T(s_1, a, y^\star_{s_1}(\infty)) > 0$; that is, the actions of player 1 that are absorbing against $y^\star_{s_1}(\infty)$. Since $\operatorname{supp}(y^\star_{s_1}(\infty)) = \operatorname{supp}(y_{s_1}(n))$, $A^\star$ is well defined.

Note that for any $a \in A^\star$, $w_T(s_1, a, y_{s_1}(n)) > 0$, hence $w_T(s_1, a, y^\star_{s_1}(\infty)) > 0$. Moreover, if $a' \notin A^\star$ then $\lim_{n \to \infty} w_T(s_1, a, y_{s_1}(n))/w_T(s_1, a', y_{s_1}(n)) = +\infty$.

Let $x^\star_{s_1}(n)$ be the probability distribution over $A^\star$ induced by $x_{s_1}(n)$. Denote $x^\star_{s_1}(\infty) = \lim_{n \to \infty} x^\star_{s_1}(n)$. Then $w_T(x^\star_{s_1}(\infty), y^\star_{s_1}(\infty)) > 0$, hence $e = (s_1, x^\star_{s_1}(\infty), y^\star_{s_1}(\infty))$ is an exit from R.

STEP 2: $d^2_{s_1}(\infty) \leq \psi^2_{d(\infty)}(e)$.

Assume to the contrary that $d^2_{s_1}(\infty) > \psi^2_{d(\infty)}(e)$. In particular, for n sufficiently large, $d^2_{s_1}(n) > \psi^2_{d(n)}(e)$. By the definition of $A^\star$ and since $d^2_{s_1}(n) \geq d^2_{s_2}(n)$ it follows that $d^2_{s_1}(n) > \psi^2_{d(n)}(s_1, x_{s_1}(n), y^\star_{s_1}(n))$ for n sufficiently large. Since $d^2_{s_1}(n) = \psi^2_{d(n)}(s_1, x_{s_1}(n), y_{s_1}(n))$, it follows that there exists an action $b_0 \in B$ such that $d^2_{s_1}(n) < \psi^2_{d(n)}(s_1, x_{s_1}(n), b_0)$ for n sufficiently large. By Lemma 6.4, $\lim_{n \to \infty} y^{b_0}_{s_1}(n)/y^b_{s_1}(n) = \infty$ for every action $b \in B^\star$, and by the definition of $B^\star$, $b_0 \notin B^\star$. In particular, $w_T(s_1, x_{s_1}(n), b_0) = 0$, which implies that for every n, $d^2_{s_1}(n) \geq \psi^2_{d(n)}(s_1, x_{s_1}(n), b_0)$ — a contradiction.

STEP 3: $d^1_{s_1}(\infty) \leq \psi^1_{d(\infty)}(e)$.

Assume to the contrary that $d^1_{s_1}(\infty) > \psi^1_{d(\infty)}(e)$. In particular, $d^1_{s_1}(n) > \psi^1_{d(n)}(e)$ for n sufficiently large. However, this implies that $d^1_{s_1}(n) < d^1_{s_2}(n)$ and there exist $a \in A^\star$ and $b_0 \in B$ such that $w_{s_2}(s_1, a, b_0) > 0$ and $\lim_{n \to \infty} y^{b_0}_{s_1}(n)/y^b_{s_1}(n) = \infty$

for every $b \in B^{\star}$. Indeed, otherwise it follows by the definition of $B^{\star}$ that for every $a \in A^{\star}$ and n sufficiently large, $d^1_{s_1}(n) > \psi^1_{d(n)}(s_1, a, y(n))$, which contradicts assumption **C.1**.

Hence **B.1** holds, and therefore $d^2_{s_1}(n) > d^2_{s_2}(n)$ for every $n \in \mathbf{N}$. For every $b \in B$ such that $w_{s_1}(s_1, x_{s_1}(n), b) = 1$, $d^2_{s_1}(n) = \psi^2_{d(n)}(s_1, x_{s_1}(n), b)$. For every $b \in B$ such that $w_T(s_1, x_{s_1}(n), b) = 0$ and $w_{s_2}(s_1, x_{s_1}(n), b) > 0$ (such as b_0), $d^2_{s_1}(n) > \psi^2_{d(n)}(s_1, x_{s_1}(n), b)$. By Lemma 6.4, for every $b \in B$ such that $\lim_{n \to \infty} y^{b_0}_{s_1}(n)/y^b_{s_1}(n) = \infty$, $d^2_{s_1}(n) > \psi^2_{d(n)}(s_1, x_{s_1}(n), b)$. It follows that $\psi^2_{d(n)}(s_1, x_{s_1}(n), y_{s_1}(n)) < d^2_{s_1}(n)$ for n sufficiently large — a contradiction.

STEP 4:   Definition of the equilibrium payoff.
Define $g = (g_s)_{s \in S} \in \mathbf{R}^{2|S|}$ by

$$
g_s = \begin{cases} u_s, & s \in T, \\ \dfrac{\sum_{s' \in T} w_{s'}(e) u_{s'}}{w_T(e)}, & s \in R. \end{cases}
$$

STEP 5:   The conditions of Lemma 4.5 hold w.r.t. $(x(\infty), y(\infty))$ and g. Condition 1 of Lemma 4.5 follows from the definition of z. Condition 2 follows from step 3 and (16) while condition 3 follows from step 2, (10) and (11). Condition 4 follows from (14) and condition 5 follows from Corollary 6.5.   ∎

8. More than two non-absorbing states

Why does our approach fail for games with more than two non-absorbing states? The reason is that if conditions **A** hold, then the equilibrium payoff that we construct need not be equal to $d(\infty)$ (see Lemma 7.3), and we run into a similar problem as when trying to generalize the proof of Vrieze and Thuijsman [11] for more than one non-absorbing state.

As an example, consider the game with four non-absorbing states (Figure 1).

Let $(x(n), y(n))$ be the stationary profile indicated in Figure 1. It is easy to verify that $(x(n), y(n))$ is a fixed point of the correspondence $\phi_{1/n}$ defined in section 6. Indeed, both players are indifferent between their actions in states 3 and 4, and in states 1 and 2 player 2 must play each action with a probability at least $1/n$. Hence the strategy of player 2 is a best reply in $Y(1/n)$. The expected payoff for player 1 by this stationary strategy profile is

$$
d^1(n) = \left(\frac{1/n}{2 - 1/n}, \frac{-1/n}{2 - 1/n}, \frac{1}{2 - 1/n}, \frac{-1}{2 - 1/n} \right).
$$

One can verify that player 1 is indifferent between his actions in states 1 and 2, and therefore his strategy is a best reply against $y(n)$.

Let $C = \{s_1, s_2\}$. Note that C is communicating w.r.t. $(x(\infty), y(\infty))$, and that the exits from C w.r.t. $(x(\infty), y(\infty))$ are $(1, B, L), (1, T, R), (2, B, L)$ and $(2, T, R)$. It turns out that ρ_∞ is the uniform distribution over these four exits and therefore

$$d_1(\infty) = d_2(\infty) = \frac{1}{4}(0,6) + \frac{1}{4}(1,1) + \frac{1}{4}(0,5) + \frac{1}{4}(-1,0) = (0,3).$$

However, there does not exist any way to exit from C in such a way that is individually rational for both players and yields the players an expected payoff $(0,3)$.
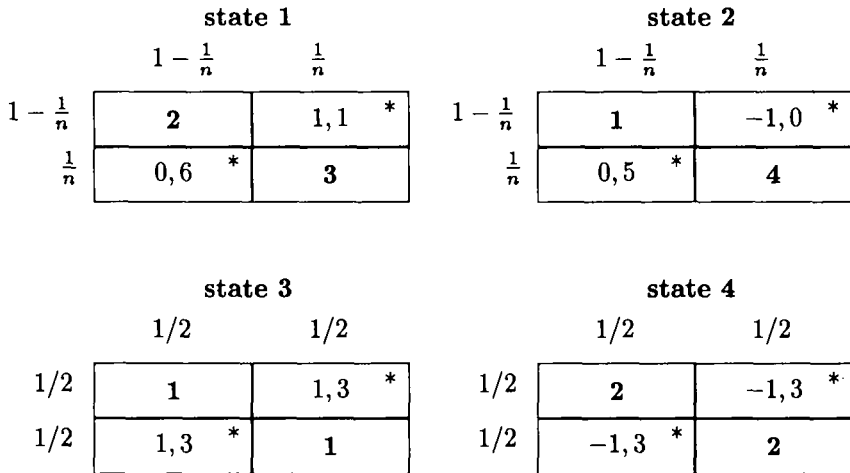

Figure 1

Though our approach fails for games with more than two non-absorbing states, we hope that one can prove the existence of an equilibrium payoff for an arbitrary number of non-absorbing states by finding another payoff function for player 2 or another constraint on his strategy space (or both).

References

[1] T. Bewley and E. Kohlberg, *The asymptotic theory of stochastic games*, Mathematics of Operations Research **1** (1976), 197–208.

[2] F. Evangelista, T. E. S. Raghavan and O. J. Vrieze, *Repeated ARAT games,* in *Statistics, Probability and Game Theory* (T. S. Ferguson and L. S. Shapley, eds.), Institute of Mathematical Statistics Lecture Notes – Monograph Series, **30**, 1996, pp. 13–28.

[3] H. Everett, *Recursive games,* in *Contributions to the Theory of Games,* **3**, Annals of Mathematical Studies, **39**, Princeton University Press, Princeton, N.J., 1957, pp. 47–78.

[4] J. F. Mertens and A. Neyman, *Stochastic games,* International Journal of Game Theory **10** (1981), 53–66.

[5] J. F. Mertens, S. Sorin and S. Zamir, *Repeated games,* Core Discussion Paper 9421, 1994.

[6] E. Solan, *Three-person absorbing games,* Mathematics of Operations Research **24** (1999), 669–698.

[7] E. Solan and N. Vieille, *Correlated equilibrium in stochastic games,* Games and Economic Behavior, to appear.

[8] N. Vieille, *On equilibria in undiscounted stochastic games,* D.P. #9446, CEREMADE, 1994.

[9] N. Vieille, *Two-player stochastic games I: A reduction,* Israel Journal of Mathematics, this volume, pp. 55–91.

[10] N. Vieille, *Two-player stochastic games II: The case of recursive games,* Israel Journal of Mathematics, this volume, pp. 93–126.

[11] O. J. Vrieze and F. Thuijsman, *On equilibria in repeated games with absorbing states,* International Journal of Game Theory **18** (1989), 293–310.