Int J Game Theory (2002) 31:91–121

International Journal of Game Theory

Correlated equilibrium payoffs and public signalling in absorbing games

Eilon Solan* and Rakesh V. Vohra†‡

* Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, *and* School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel (e-mail: eilons@post.tau.ac.il)
† Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston IL 60208 (e-mail: r-vohra@nwu.edu)

Received: April 2001/Revised: June 4, 2002

Abstract. An absorbing game is a repeated game where some action combinations are absorbing, in the sense that whenever they are played, there is a positive probability that the game terminates, and the players receive some terminal payoff at every future stage.

We prove that every multi-player absorbing game admits a correlated equilibrium payoff. In other words, for every $\varepsilon > 0$ there exists a probability distribution p_ε over the space of pure strategy profiles that satisfies the following. With probability at least $1 - \varepsilon$, if a pure strategy profile is chosen according to p_ε and each player is informed of his pure strategy, no player can profit more than ε in any sufficiently long game by deviating from the recommended strategy.

Key words: Stochastic games, Absorbing games, correlated equilibrium uniform equilibrium, public signalling

1. Introduction

There are many ways to formulate the notion of Nash equilibrium in undiscounted stochastic games. The strongest of these is *uniform ε-equilibrium*. A strategy profile is a **uniform ε-equilibrium** if for any n sufficiently large, no player could increase his expected average payoff in the first n periods by more than ε by deviating. A payoff vector is a **uniform equilibrium payoff** if it is the limit (as ε goes to 0) of the payoffs that correspond to a sequence of uniform ε-equilibrium strategy profiles. Arguments in favor of this formulation of Nash equilibria can be found in Aumann and Maschler (1995).

‡ We thank Jerome Renault and an anonymous referee who identified several inaccuracies in a previous version of the paper, and suggested ways to overcome them.

Existence of uniform equilibrium payoffs in multi-player undiscounted stochastic games while suspected is still not proven. Progress on this question has been slow and hard won. A major step was made by Mertens and Neyman (1981) who proved that every two-player *zero-sum* stochastic game admits a uniform value. Subsequently Vrieze and Thuijsman (1989) proved the existence of a uniform equilibrium payoff in two-player *non zero-sum absorbing* games. A decade and a half after the paper by Mertens and Neyman, Vieille (2000a,b) proved the existence of a uniform equilibrium payoff in *two-player non zero-sum* stochastic games. The argument is arduous and extending it to more than two players appears difficult. Some progress in this direction is described in Solan (1999) where existence of uniform equilibrium payoffs is established for *three-player absorbing* games, and in Solan and Vieille (2001b) where existence of uniform equilibrium payoffs is established for a class of *multi-player quitting* games.

While Nash equilibrium is the most popular solution concept for a game it is not the only one. For games in strategic form, Aumann (1974) proposes the notion of **correlated equilibria**, which are probability distributions over the space of strategy profiles, such that if a strategy profile is chosen according to this distribution, no player can profit by not following the strategy chosen for him.

For finite games in strategic form, correlated equilibria have a number of appealing properties. They are computationally tractable. Existence is verified by checking a system of linear inequalities rather than a fixed point. The set of correlated equilibria is closed and convex. Aumann (1987) argues that it is the solution concept consistent with the Bayesian perspective on decision making. Nor does one need to assume that the correlation device is a *deux et machina* in the game. In Foster and Vohra (1998) it is argued that players can use the history of past plays as a correlation device. Finally, our colleague Roger Myerson has been quoted as saying:

> 'If there is intelligent life on other planets, in a majority of them, they would have discovered correlated equilibrium before Nash equilibrium.'

An equivalent formulation of correlated equilibria for games in strategic form is to consider an extended game that includes a correlation device. The device chooses a signal for each player before start of play, and reveals to each player the signal chosen for him. The game then proceeds as before, but each player may base his choice of strategy on the signal he received. In this formulation, a uniform correlated ε-equilibrium is a uniform ε-equilibrium in an extended game. A uniform correlated equilibrium payoff is a limit, as ε goes to 0, of the payoffs that correspond to a sequence of uniform correlated ε-equilibria. It is this form of correlated equilibrium that is the focus of the paper.

An absorbing game is a repeated game where some of the action combinations are absorbing, in the sense that whenever they are played, the game terminates with positive probability, and the players receive some terminal payoff at every future stage. We show that *every* absorbing game admits a uniform correlated equilibrium payoff.[1] The proof uses the ideas in Solan

[1] This generalizes Solan and Vohra (2001) which considers quitting games, a special case of absorbing games.

(1999). First an auxiliary game is defined with non-absorbing payoffs that differ from those in the original game. Then we consider the limit of discounted stationary equilibria in this auxiliary game. The asymptotic properties of this sequence suggest the form that a uniform correlated equilibrium must take.

Another generalization of correlated equilibrium for sequential games involves a correlation device that sends to each player a signal before the start of *each* round. The signals can depend on the history of past signals but not on past play. This way the correlation device is independent of the play. In contrast with the problem of existence of uniform equilibrium payoffs, existence of a uniform correlated equilibrium of this kind was proved for every multi-player stochastic game with finitely many states and actions by Solan and Vieille (2001a).

Another related result is Nowak (1994), which studies multi-player stochastic games with measurable state space, compact action spaces and the average payoff criterion, that satisfy the assumption of uniform geometric ergodicity. Nowak proves in this model the existence of stationary correlated equilibrium with public signalling; that is, at every stage all players observe a public signal, which is drawn at every stage according to the same distribution.

There are two additional aspects in the paper that may interest the reader. First, the approach that we take in solving the problem is a development of the approach introduced in Solan (1999), of studying the asymptotic behavior of a sequence of discounted equilibrium in a *modified* game. Solan (1999) defined the daily payoff of each player in the auxiliary game as the minimum between his original daily payoff and his min-max level. This definition is not sufficient for our purposes, and we have to see what are the necessary properties needed for the approach to work. Thus, the proof here illuminates the properties of the modified payoff function that are required for this approach.

Second, some of the results we prove here can be used in the study of equilibria in multi-player stochastic games (see the results in section 9.2).

We start in section 2 with some examples that illustrate the main ideas the proof relies on. We then provide the model and the main result in section 3. In section 4 we present some preliminary results; we study how players can use their actions to transmit information, and we claim that in every absorbing game there exists a mixed action profile that satisfies one of a set of desirable properties. In the following four sections we prove that if those desirable properties hold, the game admits a correlated equilibrium payoff. The proof of the claim appears in section 9.

2. Examples and main ideas

We provide a series of examples that illustrate the main ideas of the proof.

A **quitting game** is a sequential game where each player has two actions: to quit (Q) or to continue (C). The game continues as long as all players decide to continue. The moment any player decides to quit, the game terminates. The terminal payoff depends on the subset of players that quit at the terminating stage. If the game continues forever, the payoff to the players is some fixed payoff vector. Quitting games are a special case of absorbing games.

2.1. Example 1

Consider first, the following three-player quitting game that was studied by Flesch et al. (1997).

Player 3: C

	C	Q
C	0, 0, 0	0, 1, 3*
Q	1, 3, 0*	1, 0, 1*

Player 3: Q

	C	Q
C	3, 0, 1*	1, 1, 0*
Q	0, 1, 1*	0, 0, 0*

In this game player 1 chooses a row, player 2 a column, and player 3 a matrix. Every absorbing entry, which corresponds to at least one player quitting, is denoted with an asterisk. Flesch et al. prove that the following profile is a uniform equilibrium.

- At stage $3n+1$, the players play $\left(\frac{1}{2}C+\frac{1}{2}Q, C, C\right)$.
- At stage $3n+2$, the players play $\left(C, \frac{1}{2}C+\frac{1}{2}Q, C\right)$.
- At stage $3n+3$, the players play $\left(C, C, \frac{1}{2}C+\frac{1}{2}Q\right)$.

Here $n = 0, 1, \ldots$. The corresponding uniform equilibrium payoff is $(1, 2, 1)$.

In a quitting game each pure strategy can be associated with an element $t \in \mathbf{N} \cup \{\infty\}$ that specifies the first period in which the player quits. If $t = \infty$, it means that the player never quits. A profile of pure strategies would be a tuple (a_1, a_2, a_3) where a_j is the period in which player j quits.

The uniform equilibrium that Flesch et al. identify corresponds to a probability distribution $p = p^{(1)} \otimes p^{(2)} \otimes p^{(3)}$ over the space of pure strategy profiles given by

$$p^{(i)}(3n+i) = 1/2^n \quad \forall n = 0, 1, 2, \ldots, i = 1, 2, 3.$$

Note that neither this distribution nor the uniform equilibrium payoff are symmetric. In fact, Flesch et al. prove that the game possesses no symmetric uniform equilibrium payoff, even though the payoff matrix is symmetric.

The probability distribution p that is defined by

$$p(1, \infty, \infty) = p(\infty, 1, \infty) = p(\infty, \infty, 1) = 1/3 \tag{1}$$

is a uniform correlated equilibrium with payoff

$$(4/3, 4/3, 4/3) = \frac{1}{3}(1, 3, 0) + \frac{1}{3}(0, 1, 3) + \frac{1}{3}(3, 0, 1).$$

Our interpretation of the equilibrium is that a correlation device chooses one of the players uniformly at random (the chosen one) and is told to quit in the first stage. The other two players are told never to quit. Suppose player 1 is informed that he was chosen. Notice that if player 1 alone disobeys the instructions by never quitting his payoff will be 0. If player 1 quits at some later stage, this does not increase his payoff.

Consider now a player not chosen, say, player 3. He does not know the identity of the chosen one; its as likely to be player 1 as it is player 2. So, if he follows his instructions to play C, his expected payoff will be 1.5. On the other hand, if player 3 quits in the first round, his expected payoff will be 1/2. He cannot know whether he can profit by deviating and quitting at the first stage, and therefore he should not deviate.

The construction described above is sensitive to two things. The first is the incentives that the chosen player has to not quitting at stage 1. The second is the payoff to an unchosen player from two players quitting at the same stage. If this were large enough, in our example above, player 3 would want to quit at the first stage.

The second of these can be accomodated by masking the stage at which the chosen player quits. For example, the chosen player is told to quit in each stage with probability $\varepsilon > 0$. Now player 3 is ignorant of who the first player is to quit as well as the stage at which they will quit. In fact with high probability any stage that player 3 chooses to quit in, he will be the only player to be quitting. The joint probability distribution p consistent with this formulation is:

$$p(n, \infty, \infty) = p(\infty, n, \infty) = p(\infty, \infty, n) = \varepsilon(1-\varepsilon)^{n-1}/3 \quad \forall n \in \mathbf{N}. \tag{2}$$

Dissuading the chosen player from quitting at a stage other than that prescribed by the device, or continuing indefinitely, is more difficult. The next example shows that this is a real possibility.

2.2. Example 2

Consider a slight modification of Example 1, where only the non-absorbing payoffs are changed.

C

	C	*Q*
C	2, 2, 0	0, 1, 3*
Q	1, 3, 0*	1, 0, 1*

Q

	C	*Q*
C	3, 0, 1*	1, 1, 0*
Q	0, 1, 1*	0, 0, 0*

The correlated equilibrium proposed for the first example does not apply here. Players 1 and 2 get higher payoffs in the non-absorbing entry. Thus, if player 1 is the chosen one, why should he quit? The other two players don't know that he is the chosen one. To deal with this possibility we will ensure that one of the unchosen players can punish player 1 for his deviation. The idea is to instruct the unchosen players to play C for a certain number of rounds and then play Q. To force compliance by player 1, the payoff to player 1 by continuing forever should be at most 1.

In this example each player i has a **single punisher** – a player $j \neq i$ that by quitting yields player i a low payoff. Player 1 is the punisher of player 3, player 2 is the punisher of player 1 and player 3 is the punisher of 2. A simple modification of the previous equilibrium scheme suggests itself: the device chooses a player uniformly at random to quit at the first stage, and informs his punisher that he should quit at the second stage if the chosen one has not quit at the first stage.

The flaws are obvious. First, the punisher knows who the chosen one is, and might profit by quitting on the first period too. This problem does not arise in this example. Second, the player who is neither the chosen one nor the punisher receives some information too. If player 3 is neither the chosen one nor the punisher, he can deduce that player 1 is the chosen one. Therefore player 3 would rather quit at the first stage.

To avoid these flaws the device must inform the punisher while masking the identity of the chosen one. One way of doing this is described below.

Define the following joint probability distribution over the space of pure strategy profiles. A player i is chosen with the uniform distribution. W.l.o.g. assume that player 1 is the chosen one. Denote by (n_1, n_2, n_3) a pure strategy profile. Since player 1 is the chosen one, n_1 is uniformly distributed in $\{1, \ldots, M\}$, where $M > 1/\varepsilon^2$. Player 2 is the punisher of 1, so n_2 is uniformly distributed in $\{M+1, \ldots, 2M\}$. Finally, $n_3 = n_2 + 1$.

Let us verify that with high probability no player can profit by not quitting at the stage recommended by the device.

The chosen player knows that he was chosen, since his quitting stage is at most M, whereas the quitting stages of the other two exceed M. If the chosen player does not quit, he will be punished and get 0. Moreover, the probability he will correctly guess the quitting stage of his punisher is low. Hence he has no reason to disobey the recommendation. With high probability the punisher and the third player received a signal in $\{M+1, \ldots, 2M\}$. In this case, the conditional probability that each is a punisher is 1/2, so they have no reason to deviate also. Thus, this joint probability distribution is a uniform correlated ε-equilibrium, provided ε is sufficiently small.

2.3. Example 3

Absorbing games can be viewed as quitting games where the players have more than one 'quitting' action and more than one 'continue action'. Thus a player may be able to punish two different players with different 'quitting' actions. For example, player i punishes player j_1 with a quitting action Q_1 and he punishes player j_2 with a quitting action Q_2. If the correlation device instructs him to use Q_1 instead of Q_2, he is in a position to infer the identity of the chosen one. This problem is solved by assuming that the game is generic, i.e. the payoffs in all the entries are different. We then consider only punishing actions which maximize the payoff of the punisher amongst his quitting actions. When a player has two continue actions then, by playing one or the other continue actions in various stages, he can send public signals to the other players. This feature can be used to construct a correlated equilibrium different from the one constructed before. This is illustrated in our next example.

We modify example 2 by adding one more action, C_2, for player 1.

	C: C	C: Q	Q: C	Q: Q
C	$2, 2, 0$	$0, 1, 3^*$	$3, 0, 1^*$	$1, 1, 0^*$
Q	$1, 3, 0^*$	$1, 0, 1^*$	$0, 1, 1^*$	$0, 0, 0^*$
C_2	$2, 2, 0$	$0, 4, 4^*$	$0, 4, 4^*$	$0, 4, 4^*$

Any correlated equilibrium payoff of Example 2 is also a correlated equilibrium payoff here. We use this example to illustrate the use of public signalling in constructing correlated equilibria.

To describe the correlated equilibrium profiles it will be convenient to use a correlation device that sends signals to the players in an arbitrary signal space. It is easily verified that the signal space that we use is equivalent to the space of strategy profiles.

Since there are only 3 players, the construction below could be simplified, but we present the construction for an arbitrary number of players.

The correlation device does the following.

1. The device chooses a player i uniformly at random. This player is informed that he should quit in the first M stages, where $M \in \mathbf{N}$ is sufficiently large.
2. The device chooses a verification key v, uniformly from the set $\{1, \ldots, M\}$.
3. The device chooses an encryption key k, uniformly from the set $\{1, \ldots, M\}$.
4. Each player $j \neq i$ receives v.
5. If $i \neq 1$, player 1 receives k, and all other players receive $k + i \bmod M$.

The players play as follows in the first M stages.

6. Each player $j \neq i$ continues in all M stages.
7. Player i chooses at random a stage $t \in \{1, 2, \ldots, M\}$. He continues in all stages but t, and quits at stage t.

If no player quit in the first M stages, the identity of i is revealed.

8. If $i \neq 1$, player 1 publicly announces v. Recall that player 1 knows v if and only if he was *not* chosen.

One possible way for player 1 to publicly announce an integer $v \in \{1, \ldots, M\}$ requires M^2 stages and is described below. Players 2 and 3 play C in all the M^2 stages. Player 1 plays C_2 in one of the stages $(v-1)M, \ldots, vM-1$, and C in all other $M^2 - 1$ stages.

If player 1 chooses the stage in which he plays C_2 at random, and if M is sufficiently large, no player can profit too much by deviating. If M is sufficiently large, the chance that player 1 can correctly guess v when he is the chosen one is arbitrarily small.

Call v' the actual message sent by player 1. If $v' \neq v$, player 1 is declared the deviator, and is punished. If $v' = v$, with high probability player 1 is not the chosen one. The play then proceeds as follows.

9. If $i \neq 1$ player 1 publicly announces k.

Now all players except player 1 can calculate the identity of the chosen one. Note that when there are only three players, once player 1 has correctly announced v, player $j \neq i, 1$ can deduce the identity of the chosen one.

Now that the identity of the chosen one was revealed, he should be punished by his punisher.

10. In one of the next M stages, the punisher j_i, provided $j_i \neq 1$, quits, and punishes player i.

11. If after M stages no one has punished the chosen one, player 1 deduces that he is the punisher, so he quits in one of the subsequent M stages.

3. The model and the main result

In this section we introduce notation and state the main result.

Definition 3.1. *A* **multi-player absorbing game** *G is given by* $(I, (A^i, r^i, u^i)_{i \in I}, w)$ *where:*

- *I is a non-empty finite set of players.*
- *A^i is a non-empty finite set of actions available for player i. Let $A = \times_{i \in I} A^i$.*
- *$r^i : A \to \mathbf{R}$ for $i \in I$. For every $a \in A$, $r^i(a)$ is the daily (non-absorbing) payoff for player i.*
- *$w : A \to [0, 1]$. For every $a \in A$, $w(a)$ is the probability the game is absorbed if the action combination a is played by the players.*
- *$u^i : A \to \mathbf{R}$ for $i \in I$. Given the game was absorbed by action combination $a \in A$, $u^i(a)$ is the constant payoff player i receives at every future stage.*

The game is played as follows. At every stage $n \in \mathbf{N}$ each player $i \in I$ chooses, independently of his opponents, an action $a_n^i \in A^i$. The action combination $a_n = (a_n^i)_{i \in I}$ determines a daily payoff $r(a_n)$ and a probability of absorption $w(a_n)$. With probability $1 - w(a_n)$ the game continues to the next stage, and with probability $w(a_n)$ the game is absorbed, and the players receive the absorbing payoff $u(a_n)$ at *every* future stage. We assume standard monitoring and perfect recall, so at every stage all the moves played upto that stage are known to all players.

For every finite set K, $\Delta(K)$ is the set of all probability distributions over K. For every $\mu \in \Delta(K)$ and every $k \in K$, $\mu[k]$ is the probability of k under μ. For every subset K' of K, $\mu[K'] = \sum_{k \in K'} \mu[k]$. We identify each $k \in K$ with the probability distribution in $\Delta(K)$ that gives weight 1 to k.

Denote $X^i = \Delta(A^i)$ and $X = \times_{i \in I} X^i$, the set of mixed-action profiles. For every subset $L \subseteq I$ of players, we denote $A^L = \times_{i \in L} A^i$ and $A^{-L} = \times_{i \notin L} A^i$. Each action $a^i \in A^i$ is identified with the probability distribution in X^i that gives weight 1 to a^i.

Let $H_n = A^n$ be the space of all histories of length n, and $H = \bigcup_{n \geq 0} H_n$ be the space of all finite histories.

A **(behavioral) strategy** for player i is a function $\sigma^i : H \to X^i$. A **profile** is a vector of strategies, one for each player. A stationary strategy can be identified with an element $x^i \in X^i$, and a stationary profile with a vector $x = (x^i)_{i \in I} \in X$. The mixed extension of w to X is still denoted by w. A mixed action profile $x \in X$ will be called **absorbing** if $w(x) > 0$ and **non-absorbing** otherwise. For every $x \in X$, every $a \in A$ and every $i \in I$, $x^i[a^i]$ is the per-stage probability to play a^i according to x^i, and $x[a] = \prod_{i \in I} x^i[a^i]$ is the per-stage probability that action combination a is played under x.

A strategy σ^i of player i is **pure** if $\sigma^i(h) \in A^i$ for every finite history $h \in H$. A profile $\sigma = (\sigma^i)$ is pure if each σ^i is pure. Let $\mathscr{S}^i$ denote the space of pure strategies of player i, and $\mathscr{S} = \times_{i \in I} \mathscr{S}^i$ the space of pure strategy profiles.

We endow $\mathscr{S}^i$ with the σ-algebra generated by finite cylinders: for every n and every vector of actions $\vec{a^i} = (a^i(h)) \in (A^i)^{H_0 \cup H_1 \cup \cdots \cup H_n}$, the set

$\{\sigma^i \in \mathcal{S}^i \mid \sigma^i(h) = a^i(h), \forall h \in H_0 \cup \cdots \cup H_n\}$ is measurable. $\mathcal{S}$ is endowed with the product σ-algebra.

Every profile σ induces a probability measure over the space of infinite plays. We denote by $\mathbf{E}_\sigma$ the corresponding expectation operator. In particular, every profile σ defines an expected payoff during the first n stages:

$$\gamma_n(\sigma) = \mathbf{E}_\sigma\left[\frac{1}{n}(r(a_1) + r(a_2) + \cdots + r(a_\theta) + 1_{\theta<n}(n-\theta)u(a_\theta))\right],$$

where θ denotes the absorption stage.

Definition 3.2. *Let $\varepsilon > 0$. A payoff vector $\gamma \in \mathbf{R}^{|I|}$ is a (uniform)* **correlated ε-equilibrium payoff** *if there exists a positive integer $n_0 \in \mathbf{N}$ and a probability measure p_ε over $\mathcal{S}$ such that for every player $i \in I$ and every measurable function $f : \mathcal{S}^i \to \mathcal{S}^i$,*

$$\mathbf{E}_{p_\varepsilon}[\gamma_n^i(\sigma)] \geq \gamma^i - \varepsilon \geq \mathbf{E}_{p_\varepsilon}[\gamma_n^i(\sigma^{-i}, f(\sigma^i))] - 2\varepsilon, \quad \forall n \geq n_0.$$

The probability measure p_ε is a (uniform) **correlated ε-equilibrium**.

A payoff vector $\gamma \in \mathbf{R}^{|I|}$ is a (uniform) **correlated equilibrium payoff** *if it is the limit, as ε goes to 0, of correlated ε-equilibrium payoffs.*

The payoff vector $\gamma \in \mathbf{R}^{|I|}$ is a (uniform) **equilibrium payoff** *if it is a correlated equilibrium payoff, and for every $\varepsilon > 0$ the probability measure p_ε is a product measure $p_\varepsilon = \bigotimes_{i \in I} p_\varepsilon^i$, where each p_ε^i is a probability measure over $\mathcal{S}^i$.*

Intuitively, a probability measure p_ε over $\mathcal{S}$ is a correlated ε-equilibrium if there is only a small probability under p_ε that given the pure strategy chosen for him, a player can profit a lot by disobeying the recommendation.

The main result of the paper is:

Theorem 3.3. *Every multi-player absorbing game admits a correlated equilibrium payoff.*

We assume w.l.o.g. that $0 \leq r, u \leq 1$, and that every player has at least two actions: $|A^i| \geq 2$ for every $i \in I$. Since payoffs are bounded, if for every $\varepsilon > 0$ there exists a correlated ε-equilibrium then a correlated equilibrium payoff exists. Moreover, if p is a correlated ε-equilibrium for some absorbing game, it is a correlated 3ε-equilibrium for any game where the payoffs differ by at most ε. In particular, we may assume w.l.o.g. that the function u is generic; that is, for every player $i \in I$ and every two action combinations $a, b \in A$, $u^i(a) \neq u^i(b)$.

As every three player absorbing game admits an equilibrium payoff, we assume throughout the paper that $|I| > 3$ (we will only use the fact that $|I| \geq 3$).

3.1. Correlation devices

It will be more convenient to consider an equivalent formulation of correlated equilibria using **correlation devices**.

Definition 3.4. *A* **correlation device** *is a pair $\mathcal{D} = (S, p)$ where $S = \times_{i \in I} S^i$ is a measurable space of signals and $p \in \Delta(S)$ is a probability distribution.*

Given a correlation device we define an extended game $G(\mathscr{D})$ as follows. A signal $s = (s^i)_{i \in I} \in S$ is chosen according to p (which is common knowledge). Each player i is informed of s^i. The game now proceeds as the original game, but each player can use his private signal to choose an action at every stage.

In this formulation, γ is a correlated ε-equilibrium payoff of G if and only if there is a probability distribution p over $\mathscr{S}$ such that γ is an ε-equilibrium payoff of $G(\mathscr{D})$, where $\mathscr{D} = (\mathscr{S}, p)$. This formulation is more general than the one we presented above, but it is more convenient to work with. In our construction, the signal space S is (equivalent to) the space of pure strategy profiles $\mathscr{S}$.

The information available to each player i at stage n is an element of $S^i \times H_{n-1}$. Thus, a strategy for player i in the extended game is a function $\sigma^i : S^i \times H \to X^i$. All previous definitions (e.g. profiles, induced payoff) can be analogously defined for the extended game.

4. Preliminaries

4.1. On exits and individual rationality

Definition 4.1. *The real number $v^i \in \mathbf{R}$ is the (uniform)* **min-max value** *of player i if for every $\varepsilon > 0$ there exists a positive integer $n_0 \in \mathbf{N}$ such that for every profile σ^{-i} there exists a strategy σ^i of player i that satisfies:*

$$\gamma_n^i(\sigma^{-i}, \sigma^i) \geq v^i - \varepsilon \quad \forall n \geq n_0,$$

and there is a profile σ_ε^{-i} of $I \backslash \{i\}$ such that for every strategy σ^i of player i,

$$\gamma_n^i(\sigma_\varepsilon^{-i}, \sigma^i) \leq v^i + \varepsilon \quad \forall n \geq n_0.$$

The profile σ_ε^{-i} is an ε-**min-max punishment profile** *against player i.*

Thus, players $I \backslash \{i\}$ can reduce the payoff of i to v^i, but they cannot reduce it any more.

Existence of the min-max value was proved by Mertens and Neyman (1981) for two-player stochastic games, and by Neyman (2002) for multi-player stochastic games. Moreover, Neyman (2002) proves that the min-max value is the limit, as the discount factor goes to zero, of the discounted min-max values.

Remark: In our construction, a deviator is punished with the min-max value and not the max-min value. There are two reasons for that. First, we would like to reduce the amount of correlation needed by the players. Second, results that are proven here might be useful in the study of equilibrium payoffs in multi-player stochastic games.

The multi-linear extension of r to X is still denoted by r. Define an extension of u to X by

$$u^i(x) = \sum_{a \in A} x[a] w(a) u^i(a) / w(x)$$

whenever $w(x) > 0$, and $u^i(x) = 0$ otherwise. Note that $w(x)u^i(x)$ is multi-linear, but u^i is *not* multi-linear; it is the expected absorbing payoff if the players play the mixed action x (given absorption occurs with positive probability).

Definition 4.2. *Let $\gamma \in \mathbf{R}^N$ be a payoff vector. A mixed action combination x is* **individually rational for** *γ if $\gamma^i \geq v^i$ for all for $i \in I$ and for every action $a^i \in A^i$,*

$$\gamma^i \geq u^i(x^{-i}, a^i).$$

Usually, deviations can be followed by punishment with the min-max level, hence one gets a stronger definition of individual rationality (see Solan (1999)). In our context players may not know the identity of the deviator, hence the deviator may deviate several times without being detected.

In absorbing games it is sometimes the case that absorption requires coordinated action on the part of a group of two or more players. For every non-absorbing mixed action $x \in X$ we will be interested in the minimal subsets of players who can force the game to be absorbed with positive probability. In other words sets $L \subseteq I$ and vectors of actions $a^L \in A^L$ such that $w(x^{-L}, a^L) > 0$, but $w(x^{-L'}, a^{L'}) = 0$ for every proper subset L' of L.

Definition 4.3. *Let $x \in X$ be a non-absorbing profile. An* **exit** *(w.r.t. x) is a vector $a^L \in A^L$ such that* (i) $\emptyset \subset L \subseteq I$, (ii) $w(x^{-L}, a^L) > 0$, *and* (iii) $w(x^{-L'}, a^{L'}) = 0$ *for every proper subset L' of L.*

If $L = \{i\}$, a singleton, denote the exit simply by a^i, and call it a **unilateral exit** of player i. If $|L| \geq 2$ the exit is a **joint exit**. Denote by $E(x)$ the set of all exits w.r.t. x.

4.2. Signalling

Since players do not have an explicit signalling device, they rely on their strategy choices to signal information. To construct an equilibrium where players will signal to each other one must ensure that no player has the incentive to deviate during a signalling phase.

Definition 4.4. *Let $x \in X$ be a non-absorbing profile. Player $i \in I$ is a* **signaller** *w.r.t. x if either* (i) $|\text{supp}(x^i)| \geq 2$, *or* (ii) *there is $a^i \notin \text{supp}(x^i)$ such that $w(x^{-i}, a^i) = 0$.*

We claim that if i is a signaller w.r.t. x then for every finite message set M and every $\varepsilon > 0$ there exists a vector of strategies of player i, $\sigma^i = (\sigma^i_m)_{m \in M}$, a positive integer n_0 and a partition $\mathscr{P} = (P_m)_{m \in M}$ of H_{n_0} such that

- $\|\sigma^i_m(h) - x^i\|_\infty < \varepsilon$ for every finite history h with length at most n_0 and every $m \in M$.
- $\mathbf{P}_{x^{-i}, \sigma^i_m}(P_m) > 1 - \varepsilon$ for every $m \in M$.
- $w(x^{-i}, \sigma^i_m(h)) = 0$ for every finite history h with length at most n_0.

Thus, the players can associate with each message a unique set of non-absorbing histories. If the realized history at stage n_0 is $h \in H_{n_0}$, and if P_m is the unique element in $\mathscr{P}$ that contains h, all players understand that message

m was sent. The first condition is needed to make deviations during the signalling phase non-profitable. The second condition ensures that with high probability m was the message the signaller intended to transmit. The third condition ensures that if all players follow the signalling mechanism, absorption does not occur during the signalling phase.

To prove the claim, fix an $\varepsilon > 0$. Choose $n_1 > 1/\varepsilon^2$ and $n_0 = |M|n_1$.

If (i) in Definition 4.4 holds, let $y^i \in X^i$ such that $\varepsilon/2 < \|x^i - y^i\| < \varepsilon$ and $\text{supp}(x^i) = \text{supp}(y^i)$. If (ii) in Definition 4.4 holds, let $y^i = \left(1 - \frac{\varepsilon}{2}\right)x^i + \frac{\varepsilon}{2}a^i$. Define σ^i_m as follows. At all stages $(m-1)n_1 \le j < mn_1$, play y^i, and at all other stages play x^i.[2,3]

The definition of P_m is as follows. P_m contains all histories h such that the average of the realized play of player i at stages $(m-1)n_1, \dots, mn_1 - 1$ is $\varepsilon/4$-close to y^i, and for every $l \neq m$, the average of the realized play of player i at stages $(l-1)n_1, \dots, ln_1 - 1$ is $\varepsilon/4$-close to x^i. If n_1 is sufficiently large the second condition holds. The histories that are not in any P_m have low probability under every σ^i_m, hence can be included in any of the sets in the partition.

Note that σ^i_m depends on the message set M, as well as on x^i and ε. M, x^i and ε also determine the number of periods n_0 required to transmit a message. From now on, whenever we specify in a profile that a signaller i sends a message m, we mean that player i plays for n_0 stages the strategy σ^i_m, and any other player $j \neq i$ plays the mixed action x^j. It will be clear from the context which mixed action profile x is to be used.

During the signalling period, players who are not signallers may deviate in two ways. Either they can alter the frequency with which they play actions in $\text{supp}(x^i)$, or they can play actions outside $\text{supp}(x^i)$. The second type of deviation is detected immediately and can be punished with the min-max value. If x is individually rational for the expected payoff of the players conditioned on the message sent, this type of deviation can be deterred. The first type of deviation does not change the message that is sent, since $\mathscr{P}$ depends only on the actions of the signaller.

We conclude this section with a definition of weak-signallers:

Definition 4.5. *Let x be a non-absorbing profile that admits one signaller i_1. A player $i_2 \neq i_1$ is a* **weak-signaller** *w.r.t. x if he is not a signaller, and there exist $a^{i_1} \in A^{i_1}$ and $a^{i_2} \notin \text{supp}(x^{i_2})$ such that $w(x^{-i_1}, a^{i_1}) = w(x^{-i_1,i_2}, a^{i_1}, a^{i_2}) = 0$.*

Since i_2 is not a signaller w.r.t. x, $w(x^{-i_2}, a^{i_2}) > 0$.

A weak-signaller cannot transmit information, since he is not a signaller. However, as we show later, with the help of the signaller he can transmit information.

4.3. Classification of non-absorbing profiles

Here we divide non-absorbing stationary profiles into four groups, according to the way information can be transmitted.

---

[2] Lotteries made at each stage are independent of the outcome of previous lotteries.

[3] In Example 3 we used a different mechanism for signalling: Player 1 had an action $a^1 \notin \text{supp}(x^1)$ such that $w(x^{-1}, a^1) = 0$, and he played that action at most once during some pre-specified time interval to transmit information. Since we do not know how to replicate this construction if (i) in Definition 4.4 is satisfied, we chose the present construction.

Definition 4.6. *A non-absorbing profile x is* **isolated** *if it admits no signallers. It is* **semi-isolated** *if it admits exactly one signaller, but no weak signallers. It is* **weak** *if it admits exactly one signaller and at least one weak signaller.*

No appellation is assigned to non-absorbing profiles that admit at least two signallers. We refer to isolated profiles also as isolated actions, to emphasize that they are pure action combinations. If x is semi-isolated, and if player i is the unique signaller it admits, we say that i is the signaller at x.

For example, consider the following two-player absorbing games where each player has 2 actions, and only the absorbing structure is given (an asterisked entry means that the probability of absorption is positive, and a non-asterisked entry means that the probability of absorption is 0):

Game 1

	L	R
T		*
B	*	*

Game 2

	L	R
T		
B	*	*

Game 3

	L	R
T		
B	*	

In game 1, (T, L) is an isolated profile. In game 2, any convex combination of (T, L) and (T, R) is semi-isolated. In game 3, (T, L) and (B, R) are weak, as is any convex combination of (T, L) and (T, R) which gives positive probability to (T, L), and any convex combination of (T, R) and (B, R) which gives positive probability to (B, R). The profile (T, R) admits two signallers.

It is easy to see that the support of any isolated action is disjoint from the support of any semi-isolated or weak profile, and that the support of any semi-isolated profile is disjoint from the support of any weak profile.

If x and y are semi-isolated, then either $\text{supp}(x)$ and $\text{supp}(y)$ are disjoint, or they have the same signaller, and any convex combination $\beta x + (1 - \beta)y$ is also semi-isolated. In particular, there are disjoint sets $B_1, \ldots, B_K$ that form the maximal supports of semi-isolated profiles: the support of any semi-isolated profile is contained in some B_k, and for each k there is some semi-isolated profile whose support is B_k. We call each set B_k a **maximal semi-isolated set**. In game 2, $K = 1$ and $B_1 = \{(T, L), (T, R)\}$.

If x is non-absorbing and $E(x)$ contains a joint exit, then x admits at least two signallers. If x is isolated, semi-isolated or weak, $E(x)$ includes only unilateral exits.

4.4. The punishment level

In this section we define the punishment level of player i at a mixed action profile x. Roughly speaking, this is the lowest payoff players $I \setminus \{i\}$ can inflict on player i when everyone is supposed to follow mainly x.

For every non absorbing profile x, denote

$$g^i(x) = \max_{a^i \in A^i \mid w(x^{-i}, a^i) > 0} u^i(x^{-i}, a^i). \tag{3}$$

By convention, the maximum over an empty set is $-\infty$. This is the best absorbing payoff player i can get if players $I \setminus \{i\}$ play x^{-i}. Let $b^i(x)$ be an action that maximizes the expression in (3). It is arbitrary if $g^i(x) = -\infty$.

Note that $b^i(x)$ is independent of x^i. Moreover, since the game is generic, if x^{-i} is a pure action then $b^i(x)$ is uniquely determined. Thus, if x is semi-isolated with signaller i then $b^i(x)$ is uniquely determined.

Define for every isolated action or semi-isolated profile x the punishment level (by absorption) player j can inflict on player i by

$$p_j^i(x) = \begin{cases} u^i(a^{-j}, b^j(a)) & x = a \text{ is an isolated action} \\ u^i(x^{-j}, b^j(x)) & x \text{ is semi-isolated with signaller } j \\ \min_{d^j \neq x^j} u^i(x^{-j}, d^j) & x \text{ is semi-isolated with signaller not } j \end{cases}$$

If x is semi-isolated with signaller j, and there is no action $d^j \in A^j$ such that $w(x^{-j}, d^j) > 0$, $p_j^i(x) = +\infty$.

In our construction, on the equilibrium path, if a player uses a unilateral exit, he uses an exit that maximizes his absorbing payoff. In particular, if x is isolated, or semi-isolated with signaller j, the only unilateral exit player j may use is $b^j(x)$. The definition of $p_j^i(x)$ captures the idea that if x is isolated, or semi-isolated with signaller j, then player j does not know the identity of the deviator, hence only the action $b^j(x)$ can be used for punishment. If x is semi-isolated with signaller not j, then our mechanism will reveal the identity of the deviator to j, hence j can choose the action that punishes the deviator the most.

Define the **punishment level** (by absorption) of player i at x by

$$p^i(x) = \min_{j \neq i} p_j^i(x). \tag{4}$$

This definition captures the idea that one can choose (through an appropriate definition of a correlation device) the player who punishes the deviator the most.

Player i is **punishable** at x if $p^i(x) \leq g^i(x)$. In this case, let $j_i(x)$ be the **punisher** of player i at x; that is, a player j that attains the minimum in the right hand side of (4). Observe that since there are at least three players, and since each player has at least two actions, $p^i(x)$ is always finite (for isolated or semi-isolated x).

The next Lemma claims that for every maximal semi-isolated set B_k and every $i \in I$, the function $p^i : \Delta(B_k) \to [0, 1]$ is quasi-concave.

Lemma 4.7. *Let B_k be a maximal semi-isolated set, and let i be the signaller at B_k. Then the function $p^i : \Delta(B_k) \to [0, 1]$ is quasi-concave.*

Proof: Since the minimum of quasi-concave functions is quasi-concave, it is sufficient to prove that for every $j \neq i$ and every $d^j \notin \text{supp}(x^j)$, the function $f : \Delta(B_k) \to [0, 1]$ defined by $f(x) = u^i(x^{-j}, d^j)$ is quasi-concave. Since the ratio of two linear functions is quasi-concave, the result follows. ∎

Corollary 4.8. *Let B_k be a maximal semi-isolated set, and let i be the signaller at B_k. There exists a concave function $\hat{p}^i : X^i \to [0, 1]$ that satisfies:* (i) $\hat{p}^i(x) = v^i$ *when $x \in \Delta(B_k)$ and $p^i(x) \geq v^i$, and* (ii) $\hat{p}^i(x) < v^i$ *otherwise.*

Proof: Let $C = \{x \in \Delta(B_k) \mid p^i(x) \geq v^i\}$. By Lemma 4.7 the function p^i is quasi-concave, hence C is convex. Since payoffs are non negative, the function $\hat{p}^i(x) = v^i \times (1 - d(x, C))$, where $d(x, C)$ is the Euclidean distance between x and C, satisfies the requirements. ■

Our next goal is to combine the punishment level and the daily payoff function to a single continuous concave function.

Lemma 4.9. *For every $i \in I$ there exists a continuous function $\tilde{r}^i : X \to [0, 1]$ that is concave in x^i for every fixed $x^{-i} \in X^{-i}$, and that satisfies:*

$$\tilde{r}^i(x) = \begin{cases} p^i(x) & x \text{ is isolated} \\ p^i(x) & x \text{ is semi-isolated with signaller not } i \\ \hat{p}^i(x) & x \text{ is semi-isolated with signaller } i \\ \min\{r^i(x), v^i\} & x \text{ is weak or admits two signallers} \end{cases} \tag{5}$$

Proof: Fix a player $i \in I$. Let $B_0 \subset A$ be the set of all isolated actions, and $B_{K+1} \subset A$ be the set of all non absorbing action profiles that are neither isolated nor contained in any maximal semi-isolated set. For every $k = 0, 1, \ldots, K+1$, let B_k^{-i} be the projection of B_k on A^{-i}:

$$B_k^{-i} = \{a^{-i} \in A^{-i} \mid (\{a^{-i}\} \times A^i) \cap B_k \neq \emptyset\}.$$

Define $B'_k = B_k^{-i} \times A^i$. Observe that the sets B_k^{-i}, $k = 0, 1, \ldots, K+1$ are disjoint, and therefore so are B'_k, $k = 0, 1, \ldots, K+1$.

We first define the function $\tilde{r}^i$ only for mixed action profiles x such that $\text{supp}(x) \subseteq B'_k$, for some k. We then extend $\tilde{r}^i$ to all X.

Let x be a mixed action profile such that $\text{supp}(x) \subseteq B'_k$, for some $k = 0, 1, \ldots, K+1$. Define

$$\tilde{r}^i(x) = \begin{cases} p^i(x^{-i}, y^i) & k = 0, (x^{-i}, y^i) \in B_k \\ p^i(x^{-i}, y^i) & 1 \leq k \leq K, (x^{-i}, y^i) \in B_k, \\ & \quad \text{and } i \text{ is not the signaller at } B_k \\ \hat{p}^i(x) & 1 \leq k \leq K, x \in B'_k, i \text{ is the signaller at } B_k \\ \min\{r^i(x), v^i\} & k = K+1, x \in B'_k. \end{cases} \tag{6}$$

Observe that (6) agrees with (5) for every mixed action profile x such that $\text{supp}(x) \subseteq \bigcup_{k=0}^{K+1} B_k$, and that for every k and every fixed $x^{-i} \in \Delta(B_k^{-i})$, the function $\tilde{r}^i(x^{-i}, x^i)$ is concave in x^i.

We now extend $\tilde{r}^i$ to X. For every $x \in X$ and every $k = 0, 1, \ldots, K+1$, let $\pi_k : X \to \Delta(B'_k)$ be the projection function:

$$\pi_k(x)[a] = \frac{x[a]}{x[B'_k]}, \quad \forall a \in B'_k.$$

The projection is defined arbitrarily if $x[B'_k] = 0$. Note that since $B'_k = B_k^{-i} \times A^i$, $x[B'_k] = x^{-i}[B_k^{-i}]$. Note also that $\tilde{r}_k^i(\pi_k(x))$ is already defined for every x such that $x[B'_k] > 0$.

Fix $\delta \in (0, 1/2)$. Since the sets B'_k, $k = 0, 1, \ldots, K+1$, are disjoint, and since $\delta < 1/2$, if $x^{-i}[B_k^{-i}] > 0$ and $x^{-i}[B_l^{-i}] > 0$ then $k = l$. Define for every $x \in X$

$$\tilde{r}^i(x) = \sum_{k=0}^{K+1} 1_{x^{-i}[B_k^{-i}] \geq 1-\delta} \frac{x^{-i}[B_k^{-i}] - (1-\delta)}{\delta} \tilde{r}^i(\pi_k(x)).$$

Observe that at most one term in this summation is non zero.

The extended function $\tilde{r}^i$ is a sum of finitely many continuous functions, hence continuous, and it clearly agrees with (5) on $\bigcup_{k=1}^{K+1} \Delta(B'_k)$. For every fixed $x^{-i} \in X^{-i}$, $1_{x^{-i}[B_k^{-i}] \geq 1-\delta} \frac{x^{-i}[B_k^{-i}] - (1-\delta)}{\delta}$ is independent of x^i, hence $\tilde{r}^i(x^{-i}, x^i)$ is concave in x^i. ■

4.5. A classification result

For every $x \in X$ and every probability distribution $\mu \in \Delta(E(x))$ we define the **expected absorbing payoff** given by μ to be

$$u(\mu) = \sum_{a^L \in E(x)} \mu[a^L] w(x^{-L}, a^L) u(x^{-L}, a^L) \Big/ \sum_{a^L \in E(x)} \mu[a^L] w(x^{-L}, a^L).$$

Recall that if $E(x)$ contains joint exits then x admits two signallers.

Proposition 4.10. *For every absorbing game there is a mixed action profile $x \in X$ and a probability distribution $\mu \in \Delta(E(x))$ that satisfy one of the following conditions.*

1. *x is absorbing, x is individually rational for $u(x)$, and $u^i(x) = u^i(x^{-i}, a^i)$ for every player i and every action $a^i \in \text{supp}(x^i)$ such that $w(x^{-i}, a^i) > 0$.*
2. *x is non absorbing, and x is individually rational for $r(x)$.*
3. *x is non absorbing,* supp(μ) *contains a single exit, which is unilateral, and x is individually rational for $u(\mu)$.*
4. (a) *x is non absorbing,* (b) *x is individually rational for $u(\mu)$,* (c) *for every player i, if $a^i \in \text{supp}(\mu)$, then $u^i(x^{-i}, a^i) = g^i(x) \geq v^i$, and one of the following conditions holds:*

   d) (i) *x is isolated, and* (ii) *for every player $i \in I$, $\mu[E(x) \cap A^i] > 0$ implies that i is a punishable player at x.*
   d′) (i) *x is semi-isolated with signaller i_0, and* (ii) *for every player $i \neq i_0$, $\mu[E(x) \cap A^i] > 0$ imply that i is a punishable player at x.*
   d″) *x is either weak, or admits at least two signallers.*

Since the proof of this Proposition is involved, it is deferred to Section 9.

It is well known that if condition 1 (*resp.* 2, 3) holds, then $u(x)$ (*resp.* $r(x)$, $u(\mu)$) is an equilibrium payoff. Thus, given Proposition 4.10, to prove Theorem 3.3 it suffices to show that if 4 holds, the game admits a correlated equilibrium payoff. Moreover, we will see that in this case, $u(\mu)$ is a correlated equilibrium payoff.

In the next section we sketch the construction of equilibrium payoffs in the first three cases. In the following three sections we show how to construct a correlated equilibrium payoff if the three cases (4.d), (4.d′), (4.d″) that appear in condition 4 Proposition 4.10 hold.

5. Cases 1, 2 and 3

If either one of the first three cases of Proposition 4.10 hold, an equilibrium payoff exists. We will construct for each of the cases an ε-equilibrium profile; namely, a correlated ε-equilibrium with a trivial correlation device that sends no messages. The construction is known and standard, and the interested reader is referred to Vrieze and Thuijsman (1989), Solan (1999) or Vieille (2000b) for more details.

Assume that the conditions of Case 1 are satisfied. The players play the stationary profile x, and monitor their opponents for deviations. If the players follow the stationary profile x the expected payoff is $u(x)$. There are deviations of two types: (i) player i may play an action not in $\text{supp}(x^i)$, and (ii) player i may alter the frequency in which he plays actions in $\text{supp}(x^i)$. Deviations of the first type are detected immediately, and can be punished at the min-max level. Since x is individually rational for $u(x)$, such deviations are not profitable. Deviations of the second type cannot be detected immediately, but since $u^i(x) = u^i(x^{-i}, a^i)$ whenever $w(x^{-i}, a^i) > 0$ and $a^i \in \text{supp}(x^i)$, those deviations are not profitable as well. Player i may nevertheless profit if there exists $y^i \in X^i$ such that (a) $\text{supp}(y^i) \subset \text{supp}(x^i)$, (b) $w(x^{-i}, y^i) = 0$, and (c) $r^i(x^{-i}, y^i) > u^i(x)$. Indeed, instead of playing the mixed action x^i at each stage, he plays the mixed action y^i. To deter this type of deviations, players should verify at each stage n that the distribution of the realized actions of each player i up to stage n is approximately x^i. The first player to fail this test, is punished at his min-max level.

Assume that the conditions of Case 2 are satisfied. The players play as in Case 1 the stationary profile x, and monitor their opponents for deviations. If the players follow the stationary profile x the expected payoff is $r(x)$. The two types of deviations mentioned for Case 1 apply here too, and they can be deterred as above.

Assume that the conditions of Case 3 are satisfied. Let a^i be the unique unilateral exit in $\text{supp}(\mu)$. The players play the stationary profile $(x^{-i}, (1-\eta)x^i + \eta a^i)$, where $\eta > 0$ is sufficiently small, while monitoring their opponents for deviations. If the players follow this profile the game will be eventually absorbed, and the expected average payoff is $u(\mu)$. Deviations are deterred as in the previous two cases.[4]

6. Case 4.d: Isolated actions

In this section we consider case 4.d of Proposition 4.10. Thus, we assume that $x = a$ is isolated.

[4] Actually, $u(\mu)$ is an equilibrium payoff even when the unique exit in $\text{supp}(\mu)$ is a joint exit. As this case is covered by case 4.d″, we do not solve it here.

Recall that $b^i(a)$ is the unique action of player i that maximizes the expression $u^i(a^{-i}, d^i)$ over $d^i \neq a^i$, that player i is punishable at a if $g^i(a) = u^i(a^{-i}, b^i(a)) \geq \min_{j \neq i} u^i(a^{-j}, b^j(a)) = p^i(a)$, and that player $j_i(a)$ is the punisher of i at a.

The next lemma follows from Solan and Vohra (2001, section 4.2). Since the game is generic, this Lemma resolves case 4.d.

Lemma 6.1. *If there is a probability distribution $\nu \in \Delta(I)$ that satisfies* (i) $\nu[i] > 0$ *implies that i is punishable at a, and* (ii) $\sum_{i \in I} \nu[i] u^j(a^{-i}, b^i(a)) \geq g^j(a)$ *for every $j \in I$, then the game admits a correlated equilibrium payoff.*

Sketch of Proof: Fix $\varepsilon > 0$. Assume first that $w(a^{-i}, b^i(a)) = 1$ for every player i.

Define the following mechanism, where $M \in \mathbf{N}$ is sufficiently large.

1. A quitter i is chosen according to ν.
2. Player i receives a positive integer d, uniformly distributed in $\{1, 2, \ldots, M\}$.
3. The punisher of i at a, player $j = j_i(a)$ receives the positive integer $M + d'$, where d' is uniformly distributed in $\{1, 2, \ldots, M\}$.
4. Each other player $i' \neq i, j$ receives the positive integer $M + d' + 1$.

Define now the following strategy σ^i for each player i:

- If you received the signal c (which is a positive integer), play a^i in all stages but stage c, in which you play $b^i(a)$.

It is easy to check that if the players follow the strategy profile $\sigma = (\sigma^i)$ then the expected payoff is $\sum_{i \in I} \nu[i] u(a^{-i}, b^i(a))$.

We now verify that if M is sufficiently large, no player can gain too much by deviating.

First, if M is sufficiently large, the probability a player correctly guesses d (if he is not i) or d' (if he is i) is low. Since player i is punishable, he cannot profit to much by deviating.

Second, if M is sufficiently large, then, with high probability, no player $j \neq i$ knows whether he is the punisher or not. Therefore, if a player $j \neq i$ plays some action $b^j \neq a^j$ before stage d, js expected payoff is

$$u^j(a^{-j}, b^j) \leq u^j(a^{-j}, b^j(a)) \leq g^j(a) \leq \sum_{i \in I} \nu[i] u^j(a^{-i}, b^i(a)).$$

In particular, no player $j \neq i$ can gain too much by deviating.

If $w^i(a^{-i}, b^i(a))$ is strictly less than 1, even if the 'designated quitter' plays the action $b^i(a)$ at stage d the game can continue. Once he plays $b^i(a)$, his identity is revealed to everyone. Since some players may get a low payoff if the game is actually terminated by the designated quitter, a new designated quitter must be chosen. As signals are sent only before start of play, this player needs to know in advance that, if the game is not terminated by the first quitter, he should do the job.

Thus, in this case the correlation device chooses an infinite sequence of quitters and punishers, which are chosen independently according to the procedure explained above, so that every player receives an infinite sequence of positive integers.

The players play in rounds; at round k, if the game was not already absorbed, the players play as explained above using the kth signal from the infinite sequence.

If some player i plays the action $b^i(a)$ in one of the first M stages of the round, everyone treat him as if he was the designated quitter, and continue to the next round. Note that if i is not the kth designated quitter, only the kth designated quitter knows of i's deviation, but he has no way to transmit this information to the other players.[5]

If no player i played the action $b^i(a)$ in the first M stages of the round, the identity of the punisher j is revealed at the punishment stage. From that stage on, the punisher plays $(1-\eta)a^j + \eta b^j(a)$, where $\eta > 0$ is sufficiently small, while all other players play a^{-j}. The punisher punishes with small probability at every stage, to mask the punishment stage.

One can verify that if η is small compared to $\min_{i \in N} w(a^{-i}, b^i(a)) > 0$, this mechanism is a correlated ε-equilibrium. ■

7. Case 4.d′: Semi-isolated profiles

In this section we consider case 4.d′ of Proposition 4.10, and prove that $u(\mu)$ is a correlated equilibrium payoff.

Since x is semi-isolated, μ is supported by unilateral exits. Let $\varepsilon > 0$ be sufficiently small, and let i_0 be the signaller at x.

We define the following mechanism, that is performed in rounds, and depends on the parameters $\eta \in (0,1)$, $K_1, K_2 \in \mathbf{N}$.

Coordination phase

1. The correlation device chooses for every $t \in \mathbf{N}$ an element $Y_t \in \{0\} \cup E(x)$, where $\mathbf{P}(Y_t = 0) = 1 - \eta$ and $\mathbf{P}(Y_t = a^i) = \eta\mu[a^i]$. Define for every $t \in \mathbf{N}$
$$i_t = \begin{cases} 0 & Y_t = 0 \\ i & Y_t \in A^i \end{cases}$$
2. For every $t \in \mathbf{N}$ and every $a^i \in E(x)$, the device chooses an integer $k_t(a^i) \in \{1, 2, \ldots, K_2\}$ according to the uniform distribution.
3. Each player $i \in I$ receives, for every t such that $i_t = i$, both Y_t and $k_t(Y_t)$.
4. Each player $i \in I$ such that $i \neq i_t$ receives $\{k_t(a^j), j \neq i, t \in \mathbf{N}\}$.
5. For every $t \in \mathbf{N}$ the device chooses a verification key $v_t \in \{1, 2, \ldots, K_2\}$, and an encryption key $e_t \in \{1, 2, \ldots, |I|\}$ according to the uniform distribution.
6. The signaller i_0 receives the sequence $\{e_t, t \in \mathbf{N}\}$, and, for every t such that $i_t \neq i_0$, he receives v_t as well.
7. Each player $i \neq i_0$ receives $\{v_t, t \in \mathbf{N}\}$, and $\{e_t + i_t \bmod |I|, t \in \mathbf{N}\}$.
8. All choices of the device are done independently.

If $i_t \neq 0$ then player i_t has to use the exit Y_t at the tth round. If $i_t = 0$, no player will use any unilateral exit.

[5] At the cost of a more complicated correlation device one can ensure that this type of deviation is not profitable. For more details, see the construction in section 7.

The mechanism proceeds in rounds. Each round consists of two phases, a quitting phase, which lasts for K_2 stages, and a revelation phase. We now explain the structure of round t.

Quitting Phase

9. For K_2 stages each player $i \neq i_t$ plays x^i.
10. If $i_t \neq 0$, player i_t chooses a stage $t_0 \in \{k_t(Y_t), k_t(Y_t + 1), \ldots, k_t(Y_t + K_1)\}$ according to the uniform distribution. He plays Y_t at stage t_0 of the round, and x^{i_t} at all other stages.

If this mechanism is followed then the expected payoff if $Y_t \neq 0$ is $u(\mu)$. If η is sufficiently small, the expected payoff of every player $j \neq i$ along the round, given his information, is approximately $u(\mu)$.

There are four types of deviations possible from this procedure. (i) Player i_t may play an action $a^{i_t} \neq Y_t$. (ii) Player i_t may not play the action Y_t at all. (iii) Player i_t may play (at least) twice the action Y_t. (iv) Player $i \neq i_t$ may play some action a^i.

Let us see which of those deviations can be detected by the players. Since player i_t does not know $k_t(a^{i_t})$ for $a^{i_t} \neq Y_t$, while players $i \neq i_t$ do know it, the chances that player i_t can correctly guess $k_t(a^{i_t})$ are small, provided K_2 is much large than K_1. Deviation (i) can therefore be detected with high probability. For the same reason, deviation (iv) can be detected with high probability. Deviation (iii) can be detected once player i_t plays Y_t for the second time. Since x is individually rational for $u(\mu)$, these three types of deviations are not profitable, provided a deviator is punished by his min-max level upon deviation.

To deter deviation (ii), the identity of i_t should be revealed, so that he can be punished. If no player used any unilateral exits in the first K_2 stages of round t, a revelation phase takes place.

Revelation + Punishment phase

11. The signaller i_0 publicly transmits v_t and e_t.

By transmitting v_t, the signaller i_0 proves that he is not i_t; if K_2 is sufficiently large than the chance that he can correctly guess v_t is low. After the revelation phase, all players *but player* i_0 know the identity of i_t (unless $i_0 = i_t$, in which case i_0 is also aware of that). If i_0 is the deviator, he can be punished at his min-max level. If $i_t \neq i_0$ and the punisher of i_t is $j \neq i_0$, then player j has to punish i_t.

12. If the punisher j of i_t is not i_0, for $1/\eta^2$ stages,[6] player j plays $(1-\eta)x^j + \eta d^j$, where $d^j \in A^j$ is the action that minimizes $u^i(x^{-j}, a^j)$ among all actions a^j such that $w(x^{-j}, a^j) > 0$ and η is sufficiently small. In those stages, player i_0 plays x^{i_0}, and every player $j' \neq j, i_0$, plays $x^{j'}$.

This ends the description of round t.

[6] Whenever we refer to a non-integer number s of stages, it should be understood as the smallest integer larger than s.

Absorption need not occur during the last $1/\varepsilon^2$ stages of the round for three reasons: (i) Player i_t was supposed to be punished, but by the luck of the draw, player j did not punish him. This event occurs with low probability, provided η is sufficiently small. (ii) $i_t = 0$, and no player was supposed to be punished. (iii) Player i_0 is the punisher of player i_t.

In the first two cases the players continue to the next round. However, if i_0 is the punisher of i_t, then i_0 should play his punishing action. Recall that the punishing action of i_0 is $b^{i_0}(x)$, so that, if all other players never use a unilateral exit, i_0 will eventually play $b^{i_0}(x)$, thereby punishing i_t, without knowing who i_t is. Thus, if i_0 is the punisher of i_t, in all subsequent rounds each player $j' \neq i_0$ stops following the above procedure, and plays the mixed action $x^{j'}$.

The only complication that may arise is if i_0 is the punisher of i_t, but $\mu[b^{i_0}(x)] = 0$. Observe that if the players follow the above mechanism then absorption eventually occurs, and the expected payoff is $u(\mu)$. Thus, if for $1/\eta^2$ subsequent rounds no player has used any exit in $E(x)$, player i_0 understands that he is the punisher, and plays at every subsequent stage the mixed action $(1-\eta)x^{i_0} + \eta b^{i_0}(x)$.[7]

It is straightforward to verify that no player can profit too much by deviating, provided η is chosen sufficiently small and K_1 and K_2 sufficiently large.

8. Case 4.d″: Other non-absorbing profiles

In this section we deal with weak profiles and non-absorbing profiles that admit at least two signallers. In these cases the identity of the chosen one can be revealed to every player, so that he can be punished with his min-max level, rather than by single punishments. We will prove that $u(\mu)$ is a correlated equilibrium payoff.

8.1. x admits at least two signallers

In this section we assume that x admits at least two signallers. In particular, $E(x)$ may contain joint exits.

It is well known (see, e.g. Vieille (2000b) or Solan (1999)) that joint exits can be controlled by the players. To control unilateral exits the device chooses whether any player should use a unilateral exit, and if so who it is. The signallers will then reveal the identity of the chosen player. Since there are at least two signallers, the identity is revealed to everyone, and if the chosen player does not use a unilateral exit, he can be punished. If one of the signallers misreports, the report of the other signaller is still consistent with the realized play. So such a deviation can be detected by the players.

Our construction here is similar to the one presented in Case 4.d′. We describe here only the relevant changes.

Let i_1 and i_2 be two distinct signallers. The coordination phase is similar to that presented in Case 4.d′, with the following exception. The device chooses a verification key and an encryption key independently for the two signallers at

[7] It can be shown that if $\mu[b^{i_0}(x)] = 0$ then there is a player $j \neq i_0$ who can punish i_t, so that this case essentially need not arise.

every stage. The verification and encryption keys are handled in an analogous way to that in Case 4.d′.

Let $\delta \in (0, \varepsilon)$ be sufficiently small. For every $a^L \in \text{supp}(\mu)$ define $\delta(a^L) = (\delta\mu[a^L]/w(x^{-L}, a^L))^{1/|L|}$.

Define a strategy σ^i in rounds. Recall that K_1 and K_2 are two sufficiently large integers. The first K_2 stages of the round are devoted to unilateral exits in $E(x) \cap \text{supp}(\mu)$, as done in Case 4.d′. Let E' be the set of all joint exits in $E(x) \cap \text{supp}(\mu)$. Each one of the following $|E'|$ stages corresponds to one joint exit a^L in E'. At the stage that corresponds to the joint exit a^L each player $i \notin L$ plays x^i, and each player $i \in L$ plays $(1 - \delta(a^L))x^i + \delta(a^L)a^i$.

In the revelation phase both signallers execute step 11 as described in Case 4.d′.

If the players follow $\sigma = (\sigma^i)$ then the game will eventually be absorbed. Moreover, provided that δ is sufficiently small, there exists $\eta \in (0, 1)$ such that the probability that the game is absorbed through the exit $a_k^{L_k}$ is approximately $\mu[a_k^{L_k}]$, thereby the expected payoff for the players is approximately $u(\mu)$.

There are several ways players may deviate from this procedure. (i) A player could play an action that has probability 0 under this procedure. Such a deviation is detected immediately, and can be punished at the min-max level. By condition 4.b of Proposition 4.10 such a deviation is not profitable. (ii) Player i may play an action $a^i \in \text{supp}(\mu)$ when he is not supposed to, or not play it when he is supposed to. If i deviates in this way, and the game is not terminated, his deviation is detected after the revelation phase, and can be punished at the min-max level. As in (i), it is not profitable. (iii) Player i may alter the frequency with which he plays different actions in $\text{supp}(x^i)$, or with which he perturbs to a_k^i in stages that correspond to a joint exit. To deter this kind of deviations, we add standard statistical tests (see, e.g., Solan (1999) or Vieille (2000b)). (iv) A signaller, say i_1, can signal an incorrect signal at some round. Since he does not know $v_{a^j}^t$ for $a^j \neq Y_t$, if he sends an incorrect verification key, this key does not correspond to the key the other players possess, and his deviation can be identified. If he sends an incorrect encryption key, trying to frame an innocent player, the report of the other signaller coincides with the realized play. Thus this deviation is detectable as well.

8.2. x is weak

In this section we assume that x is weak; that is, x admits one signaller i_0 and at least one weak signaller i_1. Since x is weak, $E(x)$ contains only unilateral exits.

We will see that the identity of the designated quitter can be revealed to everyone. The construction is similar to the construction presented in section 7. The signaller i_0 can reveal the identity of the designated quitter to everyone. However, i_0 will be ignorant of the identity of the designated quitter. We then append a phase in which the weak signaller reveals the identity of the designated quitter to i_0. Afterwards, the designated quitter is punished by his min-max value. Here we will explain how the weak signaller i_1, with the help of the signaller, reveals the identity of the designated quitter to the signaller.

Fix $\varepsilon > 0$. Let $a^{i_0} \in A^{i_0}$ and $a^{i_1} \notin \text{supp}(x^{i_1})$ be two actions that satisfy

$$w(x^{-i_0}, a^{i_0}) = w(x^{-i_0,i_1}, a^{i_0}, a^{i_1}) = 0.$$

The device chooses $|I|$ different numbers $t_1 < t_2 < \cdots < t_{|I|}$ in the range $\{1, \ldots, T\}$ with the uniform distribution,[8] where T is sufficiently large so that $\mathbf{P}(t_{|I|} < T - 1/\varepsilon) > 1 - \varepsilon$. To each member of $\{i_0, i_1\}$ who is not the designated quitter, the device sends these numbers.

Next, the revelation phase is modified. After player i_0 sends the verification key and the encryption key, player i_0 either reveals that he is the designated quitter, or reveals the identity of the designated quitter i to every player $j \neq i_0$. Player i_1 now has to reveal the identity of the designated quitter to i_0, assuming i_0 is not the deviator.

For simplicity, number the following T stages by $\{1, 2, \ldots, T\}$. In those stages the players play as follows.

- Each player $j \neq i_0, i_1$ plays x^j.
- If i_0 is not the designated quitter, he plays a^{i_0} at every stage t_k, $k = 1, 2, \ldots, |I|$. At all other stages he plays x^{i_0}.
- If i_1 is not the designated quitter, he plays a^{i_1} at stage t_i, and x^{i_1} at all other stages.

Since $w(x^{-i_0}, a^{i_0}) = w(x^{-i_0,i_1}, a^{i_0}, a^{i_1}) = 0$, if the players follow the revelation phase the game is not absorbed.

If i_1 is not the designated quitter, he knows $t_1, \ldots, t_{|I|}$ and therefore reveals the identity of the designated quitter to i_0. If, on the other hand, i_1 is the designated quitter, he does not know $t_1, \ldots, t_{|I|}$. If he ever plays the action a^{i_1}, with high probability it will be in a stage different than $t_1, \ldots, t_{|I|}$, and his identity as the designated quitter be revealed. If he never plays the action a^{i_1}, he is declared the deviator.

It is easy to verify that no player can profit too much by any type of deviation.

9. Proof of Proposition 4.10

The goal of this section is to prove Proposition 4.10. Our approach is similar in spirit to that of Solan (1999). We first introduce an auxiliary game that is 'close' in some sense to the original absorbing game. By studying the asymptotic behavior of a sequence of discounted equilibria of the auxiliary game, we establish the existence of a mixed action x and a probability distribution over $E(x)$ that satisfy one of the sufficient conditions listed in Proposition 4.10.

9.1. Definition of an auxiliary game

In Solan (1999) an auxiliary game is defined by changing the non-absorbing payoff of the original game. For every discount factor $\lambda \in (0, 1)$ the auxiliary game is shown to admit a stationary λ-discounted equilibrium x_λ. Moreover,

[8] That is, every increasing sequence of $|I|$ numbers in this range has the same probability to be chosen.

the limit of the λ-discounted min-max values of the auxiliary game is equal to the min-max value of the original game. It is then proved that if there is no uniform ε-equilibrium where the players play the limit stationary strategy $x_0 = \lim_{\lambda \to 0} x_\lambda$ and statistically check for deviations of their opponents, then there exists a probability distribution μ over the exits $E(x_0)$ such that x_0 is individually rational for $u(\mu)$. We cannot apply this result directly to our case since we require the μ to satisfy an additional punishability condition. Nevertheless it is still possible to execute something similar.

For every discount factor $\lambda \in (0,1)$ we define an auxiliary discounted game $G_\lambda(\tilde{r})$. The payoff to player i in $G_\lambda(\tilde{r})$ associated with a strategy profile σ is:

$$\tilde{\gamma}^i_\lambda(\sigma) = \mathbf{E}_\sigma\left(\lambda \sum_{n=1}^{\infty} (1-\lambda)^{n-1} (1_{n \le \theta} \tilde{r}^i(x_n) + 1_{n > \theta} u^i(x_\theta))\right)$$

where $\tilde{r}^i$ is given by Lemma 4.9, x_n is the mixed-action prescribed by σ at stage n, and θ is the stage of absorption. That is, the absorbing game with non-absorbing payoff $\tilde{r}$, but at stage n if the game is not yet absorbed, instead of getting the payoff $r(a_n)$ the players get the payoff $\tilde{r}(x_n)$.

Lemma 9.1. *The game $G_\lambda(\tilde{r})$ admits a stationary equilibrium.*

Proof: By Lemma 4.9, for every player $i \in I$ the function $\tilde{r}^i$ is continuous and concave in x^i for every fixed $x^{-i} \in X^{-i}$.

It is well known (see, e.g., Vrieze and Thuijsman (1989) or Solan (1999)) that for every player i, every discount factor $\lambda \in (0,1)$ and every stationary profile x

$$\tilde{\gamma}^i_\lambda(x) = \frac{\lambda \tilde{r}^i(x) + (1-\lambda) w(x) u^i(x)}{\lambda + (1-\lambda) w(x)}. \tag{7}$$

Since the denominator is strictly positive, $\tilde{\gamma}^i_\lambda$ is continuous.

We now show that for every player $i \in I$ and every fixed $x^{-i} \in X^{-i}$, the function $\tilde{\gamma}^i_\lambda(x^{-i}, x^i) : X^i \to [0,1]$ is quasi-concave; that is, for every $c \in \mathbf{R}$, the set $\{x^i \in X^i \mid \tilde{\gamma}^i_\lambda(x^{-i}, x^i) \ge c\}$ is convex. Let $x^{-i} \in X^{-i}$, $x^i, y^i \in X^i$, $\beta \in [0,1]$ and $c \in \mathbf{R}$ be fixed. Denote $x = (x^{-i}, x^i)$, $y = (x^{-i}, y^i)$ and $z = \beta x + (1-\beta) y$. We assume that $\tilde{\gamma}^i_\lambda(x), \tilde{\gamma}^i_\lambda(y) \ge c$, and prove that $\tilde{\gamma}^i_\lambda(z) \ge c$. By assumption, $\lambda \tilde{r}^i(x) \ge c(\lambda + (1-\lambda) w(x)) - (1-\lambda) w(x) u^i(x)$ and $\lambda \tilde{r}^i(y) \ge c(\lambda + (1-\lambda) w(y)) - (1-\lambda) w(y) u^i(y)$. By the multi-linearity of w and wu, and the concavity of $\tilde{r}^i(x^{-i}, \cdot)$, $\lambda \tilde{r}^i(z) \ge c(\lambda + (1-\lambda) w(z)) - (1-\lambda) w(z) u^i(z)$. By (7), $\tilde{\gamma}^i_\lambda(z) \ge c$.

By Theorem 4.4.1 in Mertens, Sorin and Zamir (1994), the game $G_\lambda(\tilde{r})$ admits a stationary equilibrium, as desired. ■

By Lemma 9.1 for every discount factor λ the game $G_\lambda(\tilde{r})$ admits a stationary equilibrium x_λ. $\tilde{\gamma}_\lambda(x_\lambda)$ is the corresponding discounted equilibrium payoff. By taking a subsequence, we assume w.l.o.g. that the limits $x_0 = \lim_{\lambda \to 0} x_\lambda$ and $\tilde{\gamma}_0 = \lim_{\lambda \to 0} \tilde{\gamma}_\lambda(x_\lambda)$ exist, and that for every $i \in I$, the support, $\text{supp}(x^i_\lambda)$, is independent of λ. In the sequel we will assume using the same reasoning that other limits we take exist.

Recall that for every discount factor $\lambda \in (0, 1)$ and every profile x

$$\tilde{\gamma}_\lambda(x) = \alpha_\lambda(x)\tilde{r}(x) + (1 - \alpha_\lambda(x))u(x),$$

where $\alpha_\lambda(x) = \lambda/(\lambda + (1 - \lambda)w(x))$. We define $\alpha_0 = \lim_{\lambda\to 0} \alpha_\lambda(x_\lambda)$.

Note that if y is an absorbing profile and (y_λ) are stationary profiles such that $y_\lambda \to y$ then $\lim_{\lambda\to 0} \alpha_\lambda(y_\lambda) = 0$ and $\lim_{\lambda\to 0} \tilde{\gamma}_\lambda(y_\lambda) = u(y)$.

For every exit $a^L \in E(x_0)$ define

$$x_\lambda[a^L] = \prod_{i\in L} x_\lambda^i[a^i] \prod_{i\notin L} x_\lambda^i[x_0^i].$$

This is the per-stage probability that the game is absorbed through a^L if the players play x_λ. x_λ induces a probability distribution over $E(x_0)$ as follows:

$$\mu_\lambda[a^L] = w(x_\lambda^{-L}, a^L)x_\lambda[a^L] \Bigg/ \sum_{b^L\in E(x_0)} w(x_\lambda^{-L}, b^L)x_\lambda[b^L].$$

This is the conditional probability that the game is absorbed by the exit a^L when the players follow x_λ, given that an exit in $E(x_0)$ is used.

We define for every $a^L \in E(x_0)$

$$\mu_0[a^L] = \lim_{\lambda\to 0} \mu_\lambda[a^L].$$

Then μ_0 is a probability distribution over $E(x_0)$.

One can verify that (Solan 1999, Lemma 6.6)

$$\lim_{\lambda\to 0} u^i(x_\lambda) = \sum_{a^L\in E(x_0)} \mu_0[a^L]u^i(x_0^{-L}, a^L) = u^i(\mu_0).$$

It follows that

$$\tilde{\gamma}_0 = \alpha_0\tilde{r}(x_0) + (1 - \alpha_0)u(\mu_0). \tag{8}$$

We first prove that if player i has some action a^i that is absorbing against x_0^{-i}, then his absorbing payoff by using a^i cannot exceed $\tilde{\gamma}_0^i$.

Lemma 9.2. *If $a^i \in A^i$ satisfies $w(x_0^{-i}, a^i) > 0$ then $u^i(x_0^{-i}, a^i) \le \tilde{\gamma}_0^i$.*

Proof: Since $w(x_0^{-i}, a^i) > 0$ it follows that $\lim_{\lambda\to 0} \alpha_\lambda(x_\lambda^{-i}, a^i) = 0$. Therefore

$$\tilde{\gamma}_0^i = \lim_{\lambda\to 0} \tilde{\gamma}_\lambda^i(x_\lambda) \ge \lim_{\lambda\to 0} \tilde{\gamma}_\lambda^i(x_\lambda^{-i}, a^i) = u^i(x_0^{-i}, a^i). \qquad \blacksquare$$

Lemma 9.3. *Let $i \in I$. If x_0 is either isolated or semi-isolated with signaller not i, then $\tilde{\gamma}_0^i \ge p^i(x_0)$. In particular, if $\alpha_0 < 0$ then $u^i(\mu_0) \ge \tilde{\gamma}_0^i$.*

Proof: We prove the result when x_0 is isolated. The proof when x_0 is semi-isolated with signaller not i is similar.

By the definition of $\tilde{r}^i$, $\tilde{r}^i(x_0) = p^i(x_0)$. If $\alpha_0 = 1$, the result follows by (8).

If $\alpha_0 < 1$ then μ_0 is supported by unilateral exits. By the definition of the punishment level, $u^i(x_0^{-j}, b^j(x_0)) \geq p^i(x_0)$ for every $j \neq i$. Recall that $\tilde{\gamma}_0^i = \lim_{\lambda\to 0} \gamma_\lambda^i(x_\lambda) \geq \lim_{\lambda\to 0} \gamma_\lambda^i(x_\lambda^{-i}, x_0^i)$. By (8), the right hand side is a convex combination of $\tilde{r}^i(x_0)$ and $u^i(x_0^{-j}, b^j(x_0))$, $j \neq i$. The first claim follows. The second claim now follows by (8). ■

In the rest of the section we study the asymptotic properties of the sequence $(x_\lambda)_{\lambda\to 0}$.

9.2. The limit of the discounted equilibrium payoffs

In the present section we compare various quantities to the min-max value.

Lemma 9.4. *For every isolated action a and every player* $i \in I$, $\max\{p^i(a), g^i(a)\} \geq v^i$.

Proof: Consider the following profile of players $I\backslash\{i\}$:

1. Each player $k \in I\backslash\{i, j_i(a)\}$ plays a^k.
2. Player $j_i(a)$, the punisher of i at a, plays $(1-\eta)a^{j_i(a)} + \eta b^{j_i(a)}(a)$, where $\eta \in (0,1)$.

The best that player i can do against that profile is (up to η) $\max\{p^i(a), g^i(a)\}$. Thus, players $I\backslash\{i\}$ can bound the payoff of i from above by $\max\{p^i(a), g^i(a)\}$, and therefore his min-max value cannot exceed that number. ■

A similar argument proves the following.

Lemma 9.5. *For every semi-isolated profile x with signaller* i_0*, and every player* $i \neq i_0$, $\max\{p^i(x), g^i(x)\} \geq v^i$.

Lemma 9.6. *Let* $B = \times_{i\in I} B^i$ *be the support of a maximal semi-isolated profile with signaller* i_0. *Assume that* $g^{i_0}(x) < v^{i_0}$ *for any semi-isolated x such that* $\text{supp}(x) \subseteq B$.[9] *Denote by* a^{-i_0} *the unique action combination of* $I\backslash\{i_0\}$ *in B. Then*

$$\max_{y^{i_0} \in \Delta(B^{i_0})} \min\{r^{i_0}(a^{-i_0}, y^{i_0}), p^{i_0}(a^{-i_0}, y^{i_0})\} \geq v^{i_0}.$$

Before proving the Lemma, we define the max-min value, and recall a result due to Neyman (2002). The real number v^i is the **max-min value** of player i if for every $\varepsilon > 0$ there exists a positive integer $n_0 \in \mathbf{N}$ and a strategy σ^i of player i such that $\gamma_n^i(\sigma^{-i}, \sigma^i) \geq v^i - \varepsilon$ for every profile σ^{-i} and every $n \geq n_0$, and for every strategy τ^i of player i there is a profile σ^{-i} of players $I\backslash\{i\}$ such that $\gamma_n^i(\sigma^{-i}, \tau^i) \geq v^i - \varepsilon$ for every $n \geq n_0$.

Neyman (2002) proves that in a two player zero-sum stochastic games with finitely many states and actions, if each player is restricted to use strategies

[9] Recall that $g^{i_0}(x)$ depends only on x^{-i_0}, and is independent of x^{i_0}.

such that the mixed action chosen at every stage is in some fixed convex, compact and semi-algebraic subset of the set of mixed actions, then the min-max value and the max-min value coincide.

Proof: Let $c = \max_{y^{i_0} \in \Delta(B^{i_0})} \min\{r^{i_0}(a^{-i_0}, y^{i_0}), p^{i_0}(a^{-i_0}, y^{i_0})\}$.

Fix $\delta, \eta > 0$ sufficiently small. Let G' be a game similar to G, but (i) in G' every player $i \neq i_0$ can only use strategies that at every stage choose the action a^i with probability at least $1 - \eta$, and (ii) for every action combination $b^{-i_0} \in A^{-i_0}$, if there are at least two distinct players j in $I \backslash \{i_0\}$ such that $b^j \neq a^j$ then $w(b^{-i_0}, b^{i_0}) = 1$ and $u^{i_0}(b^{-i_0}, b^{i_0}) = 2$ for every $b^{i_0} \in A^{i_0}$.[10] That is, if at least two players in $I \backslash \{i_0\}$ play an action that differs from that indicated by a^{-i_0}, the game is absorbed, and the absorbing payoff of player i_0 is high.

Let $\hat{u}^{i_0}$ be the max-min value of player i_0 in G'.

We will show that $c \geq \hat{u}^{i_0} \geq v^{i_0}$, thereby proving the result.

We first show that $\hat{u}^{i_0} \geq v^{i_0}$.

By collapsing players $I \backslash \{i\}$ into a single player and using Neyman (2002), the max-min value of player i_0 in G', is equal to the min-max value of player i_0 in G', provided players $I \backslash \{i_0\}$ can correlate their actions. In particular, there exists a *correlated* profile τ^{-i_0} such that (i) at every stage, each τ chooses a with probability at least $(1 - \eta)^{|I|-1} > 1 - |I|\eta$, and (ii) for every strategy σ^{i_0} of player i_0, the expected payoff of player i_0 under $(\sigma^{i_0}, \tau^{-i_0})$ is at most $\hat{u}^{i_0} + \eta$ in every sufficiently long game.

Since if at least two players j in $I \backslash \{i_0\}$ play an action different than a^j the game is absorbed, and the absorbing payoff is 2, which is strictly more than $\hat{u}^{i_0} + \eta$, there exists a profile τ^{-i_0} that satisfies (i) and (ii), and such that the overall probability that at some stage more than one player $j \neq i_0$ plays an action different than a^j is 0.

We now define a *non-correlated* profile $\tilde{\tau}^{-i_0}$ that approximates τ^{-i_0}. For every finite history h, every player $i \neq i_0$ and every $d^i \neq a^i$, define $\tilde{\tau}^i(h)[d^i] = \tau^{-i_0}(h)[a^{-i,i_0}, d^i]$; that is, player i plays d^i with the same probability that (a^{-i,i_0}, d^i) should have occurred according to τ^{-i_0}. We set $\tilde{\tau}^i(h)[a^i] = 1 - \sum_{d^i \neq a^i} \tilde{\tau}^i(h)[d^i] \geq 1 - |I|\eta$.

Observe that under $\tilde{\tau}^{-i_0}$ at every stage each action $d^j \neq a^j$ is played with probability at most $|I|\eta$. In particular, under $\tilde{\tau}^{-i_0}$ the overall probability that at some stage at least two players $j \neq i_0$ play an action other than a^j is at most $\eta|I|2^{|I|-1}$.

It follows that by playing $\tilde{\tau}^{-i_0}$ in G' (though this profile is not permissible in G'), players $I \backslash \{i_0\}$ bound the payoff of i_0 from above by $\hat{u}^{i_0} + K\eta$, where $K > 0$ is some constant.

Since in G payoffs do not exceed the payoffs in G', by playing $\tilde{\tau}^{-i_0}$ in G players $I \backslash \{i_0\}$ bound the payoff of i_0 from above by $\hat{u}^{i_0} + K\eta$. Since η is arbitrary, $\hat{u}^{i_0} \geq v^{i_0}$.

We now prove that $c \geq \hat{u}^{i_0}$.

Let σ^{i_0} be any strategy of player i_0 in G', and define the process $(\mathbf{x}_n^{i_0})_{n \in \mathbf{N}}$ as the mixed action played by player i_0 at stage n. To simplify the proof, we assume that σ^{i_0} never chooses actions that are not in B^{i_0}: for every $a^i \notin B^{i_0}$ and every mixed action x^{-i_0} such that $x^i[a^i] \geq 1 - \eta$ for every $i \neq i_0$,

[10] The absorbing payoff of players $I \backslash \{i_0\}$ is irrelevant.

$w(x^{-i_0}, x^{i_0}) > 0$ and, if η and δ are sufficiently small, $u^{i_0}(x^{-i_0}, x^{i_0}) < g^{i_0}(x) + \delta < v^{i_0} \leq \hat{u}^{i_0}$, so that such actions can only reduce the expected payoff of player i_0.

We will now define a strategy profile τ^{-i_0} such that the expected payoff to player i_0 under $(\sigma^{i_0}, \tau^{-i_0})$ is below c. By the simplifying assumption, the value of $\mathbf{x}_n^{i_0}$ is in $\Delta(B^{i_0})$ a.s. By the definition of c, either $r^{i_0}(a^{-i_0}, \mathbf{x}_n^{i_0}) \leq c$ or $p^{i_0}(a^{-i_0}, \mathbf{x}_n^{i_0}) \leq c$ (or both).

If $r^{i_0}(a^{-i_0}, \mathbf{x}_n^{i_0}) \leq c$, each player $i \neq i_0$ plays a^i.

If $p^{i_0}(a^{-i_0}, \mathbf{x}_n^{i_0}) \leq c$, denote by j the punisher of i_0 at $\mathbf{x}_n^{i_0}$, and let $d^j \neq a^j$ satisfy $u^j(a^{-i_0,j}, \mathbf{x}_n^{i_0}, d^j) = p^{i_0}(a^{-i_0}, \mathbf{x}_n^{i_0})$. Player j plays $(1-\eta)a^j + \eta d^j$, while each player $i \neq i_0, j$ plays a^i.

Thus, under $(\sigma^{i_0}, \tau^{-i_0})$, at every stage, either the game is absorbed with probability 0, and the non absorbing payoff is at most c, or the game is absorbed with probability bounded away from 0, and the expected absorbing payoff is at most c.

It follows that the expected payoff of player i_0 under $(\sigma^{i_0}, \tau^{-i_0})$ is at most $c + \delta$ in every sufficiently long game, as desired. ■

Lemma 9.7. $\tilde{\gamma}_0^i \geq v^i$ *for every player* $i \in I$.

One way of proving the lemma would be to show that for every $\lambda \in (0,1)$ the min-max value of player i in $G_\lambda(\tilde{r})$, $v_\lambda^i(\tilde{r})$, exists and $\lim_{\lambda\to 0} v_\lambda^i(\tilde{r}) \geq v^i$. This would yield a stronger result than needed. This approach is taken in Solan (1999), where $\tilde{r}^i$ was defined as $\min\{r^i, v^i\}$, and it was proven that $\lim_{\lambda\to 0} v_\lambda^i(\min\{r, v\}) = v^i$. Since p^i is incomparable to v^i, we cannot invoke Solan's result to prove the Lemma.

Proof: Fix a player $i \in I$. In the sequel we use the fact that x_λ converge to a limit x_0, and that $\tilde{\gamma}_0^i = \lim_{\lambda\to 0} \tilde{\gamma}_\lambda^i(x_\lambda)$.

We have four cases, that correspond to isolated actions, semi-isolated actions with signaller i, semi-isolated actions with a signaller that is not i, and a case that deals with all other possibilities.

Assume that there exists $a^i \in A^i$ such that $a = (x_0^{-i}, a^i)$ is an isolated profile. By Lemma 9.2, $\tilde{\gamma}_0^i \geq g^i(a)$. Lemma 9.3 implies that $\tilde{\gamma}_0^i \geq p^i(a)$. The result follows by Lemma 9.4.

Assume that there exists $a^i \in A^i$ such that (x_0^{-i}, a^i) is a semi-isolated profile with signaller which is *not i*. Similar arguments, using Lemma 9.5, show that $\tilde{\gamma}_0^i \geq v^i$.

Assume that there exists $a^i \in A^i$ such that $x = (x_0^{-i}, a^i)$ is a semi-isolated profile with signaller i. If $g^i(x) \geq v^i$ the result follows by Lemma 9.2. Assume then that $g^i(x) < v^i$.

By Lemma 9.6 there is $y^i \in \Delta(B^i)$ such that $p^i(a^{-i}, y^i) \geq v^i$, so that $\tilde{r}^i(a^{-i}, y^i) = v^i$ and $u^{i_0}(a^{-j,i}, y^i, d^j) \geq v^i$ for every $j \neq i$ and every $d^j \neq a^j$. In particular, by (8),

$$\tilde{\gamma}_0^i = \lim_{\lambda\to 0} \tilde{\gamma}_\lambda^i(x_\lambda) \geq \lim_{\lambda\to 0} \tilde{\gamma}_\lambda^i(x_\lambda^{-i}, y^i) \geq v^i.$$

Last, assume that there is no action $a^i \in A^i$ such that one of the first three cases hold. If there exists an action $a^i \in A^i$ such that $w(x_0^{-i}, a^i) > 0$ and $u^i(x_0^{-i}, a^i) \geq v^i$ the result follows by Lemma 9.2.

Otherwise, by the definition of the min-max value, the set $D^i = \{a^i \in A^i \mid w(x_0^{-i}, a^i) = 0$ is not empty. For each $j \neq i$ set $D^j = \mathrm{supp}(x_0^j)$.

Fix $\eta, \delta > 0$ sufficiently small. The functions $\tilde{r}$ and r are continuous over X. Moreover, $\tilde{r} = \min\{r, v\}$ on $D = \times_{j \in I} \Delta(D^j)$. Hence, if η is sufficiently small, $\tilde{r}^j(x) \geq \min\{r^j(x), v^j\} - \delta$ for every $j \in I$ and every $x \in X$ such that $d(x, D) \leq \eta$. It follows from Solan (1999, Eq. (30)) that for every λ sufficiently small there exists $x^i \in \mathrm{supp}(D^i)$ such that $\tilde{\gamma}_\lambda^i(x_\lambda^{-i}, x^i) \geq v^i - 2\delta$. Therefore $\tilde{\gamma}_0^i \geq \lim_{\lambda \to 0} \tilde{\gamma}_\lambda^i(x_\lambda^{-i}, x^i) \geq v^i - 2\delta$. Since δ is arbitrary, the result follows. ■

9.3. Asymptotic analysis

In the present section we prove that x_0 satisfies one of the conditions of Proposition 4.10.

The following corollary follows easily from the definition of individual rationality and Lemmas 9.7 and 9.2.

Corollary 9.8. *x_0 is individually rational for $\tilde{\gamma}_0$.*

Lemma 9.9. *If x_0 is absorbing then condition 1 in Proposition 4.10 holds.*

Proof: Since x_0 is absorbing, $\tilde{\gamma}_0 = u(x_0)$. By Corollary 9.8, x_0 is individually rational for $\tilde{\gamma}_0 = u(x_0)$.

By Lemma 9.2, for every player $i \in I$,

$$u^i(x_0) = \sum_{a^i \in A^i} x_0^i[a^i] w(x_0^{-i}, a^i) u^i(x^{-i}, a^i) / w(x_0) \leq u^i(x_0),$$

hence $u^i(x^{-i}, a^i) = u^i(x_0)$ whenever $a^i \in \mathrm{supp}(x_0^i)$ with $w(x_0^{-i}, a^i) > 0$. ■

Lemma 9.10. *If $\alpha_0 = 1$ then either condition 2 or condition 3 of Proposition 4.10 hold.*

Proof: Since $\alpha_0 = 1$, $\tilde{\gamma}_0 = \tilde{r}(x_0)$ and x_0 is non-absorbing. By Corollary 9.8, x_0 is individually rational for $\tilde{\gamma}_0 = \tilde{r}(x_0)$. We have three cases:

1. $x_0 = a$ is an isolated action.
2. x_0 is a semi-isolated profile.
3. None of the first two cases hold.

Consider the last case first. We show that condition 2 of Proposition 4.10 holds.

By the definition of $\tilde{r}^i$, $\tilde{\gamma}_0^i = \tilde{r}^i(x_0) = \min\{r^i(x_0), v^i\} \leq r^i(x_0)$ for every $i \in I$. By Corollary 9.8 x_0 is individually rational for $r(x_0)$, as desired.

Assume now that $x_0 = a$ is an isolated action. We show that condition 3 of Proposition 4.10 holds, with $\mathrm{supp}(\mu) = b^j(a)$, for any $j \in I$.

By the definition of $\tilde{r}^i$ at isolated actions, by Corollary 9.8 and Lemma 9.2, $p^i(a) = \tilde{r}^i(a) = \tilde{\gamma}_0^i \geq g^i(a)$ for every $i \in I$. But this implies that for every $j \neq i$, $u^i(a^{-j}, b^j(a)) \geq p^i(a) \geq g^i(a)$. By Lemma 9.4, $u^i(a^{-j}, b^j(a)) \geq v^i$ for every $i, j \in I$. The claim follows.

Assume now that x_0 is a semi-isolated profile with signaller i_0. We show that condition 3 of Proposition 4.10 holds, with $\text{supp}(\mu) = b^j(a)$, for any $j \neq i_0$.

By Lemma 9.2 and the definition of $\tilde{r}^i(x_0)$, $g^i(x_0) \leq \tilde{\gamma}_0^i = \tilde{r}^i(x_0) = p^i(x_0) \leq u^i(x_0^{-j}, b^j(x_0))$ for every $i \neq i_0$ and every $j \neq i$. By Lemma 9.5 $p^i(x_0) \geq v^i$ for every $i \neq i_0$. For player i_0 we have, by Lemma 9.7, $v^{i_0} \leq \tilde{\gamma}_0^{i_0} = \tilde{r}^{i_0}(x_0)$, so that by the definition of $\tilde{r}^{i_0}(x_0)$, $\tilde{\gamma}_0^{i_0} = \tilde{r}^{i_0}(x_0) = v^{i_0}$. Since $\tilde{r}^{i_0}(x_0) = \hat{p}^{i_0}(x_0)$, it follows that $p^{i_0}(x_0) \geq v^{i_0}$. This implies, by Lemma 9.2, that $g^{i_0}(x_0) \leq \tilde{\gamma}_0^{i_0} = v^{i_0} \leq p^{i_0}(x_0) \leq u^{i_0}(x_0^{-j}, b^j(x_0))$ for each $j \neq i_0$. As in the case of isolated actions, the claim follows. ■

Assume now that x_0 is non-absorbing, but $\alpha_0 < 1$. We prove that condition 4 of Proposition 4.10 holds. Since $\alpha_0 < 1$, x_λ is absorbing for every λ sufficiently small.

If player i has a unilateral exit a^i that receives a positive probability under μ_0, then his absorbing payoff by using it is $\tilde{\gamma}_0^i$.

Lemma 9.11. *If $a^i \in E(x_0)$ and $\mu_0[a^i] > 0$ then $u^i(x_0^{-i}, a^i) = \tilde{\gamma}_0^i \geq v^i$.*

The lemma is proved in Solan (1999, proof of Theorem 4.5, Step 8). Note that if $a^i \in E(x_0)$ then $w(x_0^{-i}, a^i) > 0$, and that by Lemma 9.2, $u^i(x_0^{-i}, a^i) \leq \tilde{\gamma}_0^i$. Since the function $\tilde{r}$ is *not* multi-linear this lemma is not an immediate consequence of Lemma 9.2 and (8).

Lemmas 9.11 and 9.2 imply that condition 4.c in Proposition 4.10 holds.

Lemma 9.12. *If x_0 is non-absorbing and neither isolated nor semi-isolated, and $\alpha < 1$ then condition 4.b in Proposition 4.10 holds.*

Proof: Since x_0 is neither isolated nor semi-isolated, it is either weak or admits two signallers.

By Corollary 9.8 it is sufficient to show that $u^i(\mu_0) \geq \tilde{\gamma}_0^i$ for every $i \in I$.

Since x_0 is neither isolated nor semi-isolated, $\tilde{r}^i(x_0) \leq v^i \leq \tilde{\gamma}_0^i$, and in particular (8) implies that $u^i(\mu_0) \geq \tilde{\gamma}_0^i \geq v^i$, as desired. ■

We now confine our attention to the case when $x_0 = a$ is an isolated action, or x_0 is a semi-isolated action. Recall that in these cases $E(x_0)$ includes only unilateral exits.

Lemma 9.13. *If $x_0 = a$ is an isolated action and $\alpha < 1$ then conditions 4.b and 4.d.ii in Proposition 4.10 hold.*

Proof: Condition 4.b holds by (8), Corollary 9.8 and Lemma 9.3.

Since $E(x_0)$ contains only unilateral exits, and since the game is generic, Lemma 9.11 implies that if b^i is a unilateral exit of player i w.r.t. a and $\mu_0[b^i] > 0$ then $b^i = b^i(a)$. By Lemmas 9.11 and 9.3, if $\mu_0[b^i(a)] > 0$ then i is punishable. ■

Lemma 9.14. *If x_0 is a semi-isolated profile and $\alpha < 1$ then conditions 4.b and 4.d$'$.ii in Proposition 4.10 hold.*

Proof: Let i_0 be the unique signaller at x_0. The proof that condition 4.d$''$

holds, as well as the proof that $u^i(\mu) \geq \tilde{\gamma}_0^i$ for $i \neq i_0$, is similar to the proof provided in Lemma 9.13.

To see that $u^{i_0}(\mu) \geq \tilde{\gamma}_0^{i_0}$, use (8), Lemma 9.7 and the fact that $\tilde{r}^{i_0}(x_0) \leq v^{i_0}$. ■

References

[1] Aumann RJ (1974) Subjectivity and Correlation in Randomized Strategies. J. Math. Econ., 1:67–96

[2] Aumann RJ (1987) Correlated Equilibrium as an Expression of Bayesian Rationality. Econometrica, 55:1–18

[3] Aumann RJ and Maschler M (1995) Repeated Games with Incomplete Information, The MIT Press

[4] Flesch J, Thuijsman F and Vrieze K (1997) Cyclic Markov Equilibria in Stochastic Games. Int. J. Game Th., 26:303–314

[5] Foster DP and Vohra RV (1998) Calibrated Learning and Correlated Equilibrium. Games Econ. Behavior, 21:40–55

[6] Mertens JF and Neyman A (1981) Stochastic Games. Int. J. Game Th., 10:53–66

[7] Mertens JF, Sorin S and Zamir S (1994) Repeated Games, CORE Discussion Paper 9420-9422

[8] Neyman A (2002) Existence of the min-max of stochastic games. Discussion paper 295, the Center for Rationality and Interactive Decision Theory, the Hebrew University, Jerusalem.

[9] Nowak A (1994) Stationary Equilbria for Nonzerosum Average Payoff Ergodic Stochastic Games with General State Space, *Advances in Dynamic Games and Applications*, Annals of the International Society of Dynamic Games, Vol. 1, Birkhäuser, Basar T and Huarie A (eds.), 231–246

[10] Solan E (1999) Three-Person Absorbing Games. Math. Oper. Res., 24:669–698

[11] Solan E and Vieille N (2001a) Correlated Equilibrium in Stochastic Games. Games Econ. Behavior, 38:362–399

[12] Solan E and Vieille N (2001b) Quitting Games. Math. Oper. Res., 26:265–285

[13] Solan E and Vohra RV (2001) Correlated Equilibrium in Quitting Games. Math. Oper. Res., 26:601–610

[14] Vieille N (2000a) Equilibrium in 2-Person Stochastic Games I: A Reduction. Israel J. Math., 119:55–91

[15] Vieille N (2000b) Equilibrium in 2-Person Stochastic Games II: The Case of Recursive Games. Israel J. Math., 119:93–126

[16] Vrieze OJ and Thuijsman F (1989) On Equilibria in Repeated Games with Absorbing States. Int. J. Game Th., 18:293–310