Correlated Equilibrium, Public Signalling and Absorbing Games

Eilon Solan* and Rakesh V. Vohra†

October 5, 1999

Abstract

An absorbing game is a repeated game where some of the 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 n-player absorbing game admits a correlated equilibrium. In other words, for every $\epsilon > 0$ there exists a probability distribution p_ϵ over the space of pure strategy profiles such that 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.

*Department of Managerial Economics and Decision Sciences, Kellogg Graduate School of Management, Northwestern University, 2001 Sheridan Road, Evanston IL 60208. e-mail: e-solan@nwu.edu

†Department of Managerial Economics and Decision Sciences, Kellogg Graduate School of Management, Northwestern University, 2001 Sheridan Road, Evanston IL 60208. e-mail: r-vohra@nwu.edu

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).

Existence of uniform equilibrium payoffs in n-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) when they 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 (1997a,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 equilibria is established for *three-player absorbing* games, and in Solan and Vieille (1998b) where existence of uniform equilibria is established for a class of n-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. Indeed, 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.'

For sequential games (ones played in stages), one can define a *strategic form* correlated equilibrium in an analogous way to that of finite games. That is, a probability distribution 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. 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. The proof uses the ideas in Solan (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.

An equivalent formulation of correlated equilibria 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. A correlated equilibrium is a Nash equilibrium of the extended game. In this formulation, a uniform correlated equilibrium is a uniform equilibrium in a game where the signal space of each player coincides with his strategy space, and the signal to each player is a recommended strategy.

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. This signal can depend on the history of past signals as well as past plays. In contrast with the problem of existence of uniform equilibrium payoff, existence of a uniform correlated equilibrium of this kind was proved for every n-player stochastic game by Solan and Vieille (1998a).

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 study correlated equilibria in the subclass of quitting games. Many of the ideas that are used throughout the proof appear already here. In section 5 we provide sufficient conditions for existence of uniform equilibrium payoff and uniform correlated equilibrium payoff. We also study the way information can be transmitted in absorbing games. In section 6 we prove that in every absorbing game at least one of the sufficient conditions hold.

2 Examples and Main Ideas

We provide a series of examples involving quitting games 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 at least one 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, then 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).

C

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

Q

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

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

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 a positive integer t that specifies the first period in which the player 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 payoff 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 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 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 never quitting. 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 $\epsilon > 0$. 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:

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

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 payoff in the non-absorbing state. 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. 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. With probability 1/3 player i is the chosen one. W.l.o.g. assume that player 1 is the chosen one. Denote by (n_1, n_2, n_3) a pure strategy profile. As before, since player 1 is the chosen 1, n_1 is uniformly distributed in $\{1, \ldots, M\}$, where $M > 1/\epsilon^2$. Player 2 is the punisher of 1, so n_2 is uniformly distributed in $\{M+1, \ldots, 2M\}$. $n_3 = n_2+Y$, where Y is uniformly distributed in $\{1, \ldots, |1/\epsilon|\}$.

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 $\{M + |1/\epsilon|, \ldots, 2M\}$. 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.

2.3 Summary - Quitting Games

To construct correlated equilibrium for general quitting games we divide play into two phases: a quitting phase and a punishment phase. The device chooses one player to be the quitter according to a known probability distribution μ. It then chooses a quitting stage for each player, making sure that the chosen one receives the earliest stage, and his punisher receives the second earliest stage. The difficulty is to prove that there exists a μ that (i) is supported by players who have punishers and (ii) the corresponding average absorbing payoff is high for every player.

2.4 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	Q
C	$2,2,0$	$0,1,3$ *
Q	$1,3,0$ *	$1,0,1$ *
C_2	$2,2,0$	$0,4,4$ *

Q

	C	Q
C	$3,0,1$ *	$1,1,0$ *
Q	$0,1,1$ *	$0,0,0$ *
C_2	$0,4,4$ *	$0,4,4$ *

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.

As before, the device chooses a player uniformly at random. This player is informed that he should quit in the first M stages.

If the chosen player does not comply, the device will reveal his identity. Since only player 1 can transmit information (by playing either C or C_2 while the other two continue), it is his task to reveal the identity of the chosen one. However, he must not know the identity before the start of play, otherwise he might profit by deviating in the quitting phase.

Information about the identity of the chosen one will be split between the players, so that the information any one player has tells him nothing about the identity of the chosen person, while the information of player 1 with any of the other players is enough to identify the chosen one. First, let us see how players 2 and 3 can verify whether player 1 is the chosen one.

The device chooses a password, and sends it to players 2 and 3 before the start of play. Player 1 receives the password if and only if he is **not** the chosen one. In this context, a password is a sequence of actions in $\{C, C_2\}$

If the chosen person has not terminated the game by stage M, player 1, if he was not chosen, should play according to the password. This is called the revelation phase. If player 1 was not chosen, he received the password and can play according to it. If player 1 was chosen, he does not know the password. In this case, we do not care what he plays. If the password is long enough, the probability that he can mimic it will be arbitrarily small.

Notice that when player 1 plays C_2, the other players profit by quitting, to make deviations during this revelation phase non-profitable the password has to be stochastic.

Since we are interested in Nash equilibria of the extended game, we consider only unilateral deviations. Thus, if the chosen one has not deviated, the game terminates during the quitting phase and does not reach the revelation phase. If the game reaches the revelation phase, it means that the chosen one has deviated, hence once player 1 has repeated the password, we need not worry about him deviating in the sequel.

If player 1 can not repeat the password, he is revealed as the chosen one, and is punished by the others. Otherwise, since there are only three players, the remaining unchosen player can deduce the identity of the chosen one. If there are more than three players, the remaining unchosen players still do not know the identity of the chosen one. To reveal it, the device chooses a permutation π over the set of players, each permutation is chosen with equal probability. Players 2 and 3 receive π, whereas player 1 receives $\pi(i)$, where i is the chosen player. After player 1 has repeated the password, he transmits $\pi(i)$. The other players can now easily calculate the identity of the chosen player i.

Note that even though 2 and 3 know the value of i, the chosen one, player 1 does not know it, and therefore cannot take part in a punishment.

If $i = 2$, since player 3 is the punisher of 2, he can punish player 2 as was done in Example 2. If $i = 3$, his punisher is 1 who does not know who the chosen one is. However, if the chosen player is either 1 or 2, he is punished within a bounded time. If this bound is exceeded and no one has been punished, player 1 can deduce that player 3 was the chosen one. At that point player 1 knows that $i = 3$, and that he is the punisher.

2.5 Summary - General Case

In the general case, the profile that we construct is divided into three phases: a quitting phase, a revelation phase and a punishment phase. The device chooses a player to quit during the quitting phase. If the chosen player does not comply, play enters the revelation phase where some players transmit information that reveals the identity of the chosen one. Finally the chosen one is punished by his opponents.

We will see different forms of transmission of information, according to the number of players who can transmit information. In quitting games, no player can transmit information, hence the revelation phase is skipped. In other cases, where certain joint actions cause the game to terminate, the quitting and revelation phase are interleaved.

3 The Model and the Main Result

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

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

- *N is a finite set of players.*
- *A^i is a finite set of actions available for player i. Let $A = \times_{i \in N} A^i$.*
- *$r^i : A \to \mathbf{R}$ for $i \in N$. 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 $i \in N$. 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 $i \in N$ chooses, independently of his opponents, an action $a^i_n \in A^i$. The action combination $a_n = (a^i_n)_{i \in N}$ 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, so at every stage n all the moves played before that stage are known to all players.

For every finite set K, $\Delta(K)$ is the set of all probability distribution over K, and for every $\mu \in \Delta(K)$ and every $k \in K$, $\mu[k]$ is the probability of k under μ.

We assume w.l.o.g. that $0 \leq r, u \leq 1$, and denote $X^i = \Delta(A^i)$ and $X = \times_{i \in N} X^i$, the set of mixed-action combinations. For every subset $L \subseteq N$ of players, we denote $A^L = \times_{i \in L} A^i$ and $A^{-L} = \times_{i \not\in L} A^i$. Each action $a^i \in A^i$ is identified with the probability distribution in X^i that gives weight 1 to a^i. We also assume that each player has at least 2 actions; that is, $|A^i| \geq 2$ for every player $i \in N$.

Let $H_n = A^n$ be the space of all n-stage histories, and $H = \cup_{n \in \mathbf{N}} H_n$ 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 $x = (x^i)_{i \in N}$. The mixed extension of w to X is still denoted by w. A stationary profile $x \in X$ will be called **absorbing** if $w(x) > 0$ and **non-absorbing** otherwise. In particular, $x^i[a^i]$ is the per-stage probability to play a^i according to x^i, and $x[a]$ is the per-stage probability to play the action combination a 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 $\mathcal{S}^i$ denote the space of pure strategies of player i, and $\mathcal{S} = \times_{i \in N} \mathcal{S}^i$.

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

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

where θ denotes the absorption stage.

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

$$\mathbf{E}_p[\gamma_n^i(\sigma)] \geq \mathbf{E}_p[\gamma_n^i(\sigma^{-i}, f(\sigma^i))] - \epsilon \qquad \forall n \geq n_0.$$

A payoff vector $\gamma \in \mathbf{R}^N$ is a (uniform) **correlated equilibrium payoff** *if for every $\epsilon > 0$ there exists a correlated ϵ-equilibrium p_ϵ and a positive integer $n_1 \in \mathbf{N}$ such that*

$$\| \mathbf{E}_{p_\epsilon}[\gamma_n(\sigma)] - \gamma \|_\infty < \epsilon \qquad \forall n \geq n_1.$$

The payoff vector $\gamma \in \mathbf{R}^N$ is a (uniform) **equilibrium payoff** *if it is a correlated equilibrium payoff, and for every $\epsilon > 0$ the probability measure p_ϵ is a product measure $p_\epsilon = \otimes_{i \in N} p_\epsilon^i$, 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 n-player absorbing game admits a correlated equilibrium payoff.*

Since payoffs are bounded, if for every $\epsilon > 0$ 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 2ϵ-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 $i \in N$ and every two action combinations $a, b \in A$, $u^i(a) \neq u^i(b)$.

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* $S = \times_{i \in N} S^i$ *is a measurable space of signals and* $p \in \Delta(S)$ *is a probability distribution.*

Given a correlation device we define an extended game $G(\mathcal{D})$ as follows. A signal $s = (s^i)_{i \in N} \in S$ 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, a correlated equilibrium payoff of G is an equilibrium payoff of $G(\mathcal{D})$. This formulation is more general than the one we presented in section 3, but it is more convenient to work with. In our construction, the signal space S is (equivalent to) the space of pure strategy profiles $\mathcal{S}$.

The information available to each player i at stage n is an element of $S^i \times H_n$. 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 Quitting Games

In this section we study the class of quitting games which are themselves a special case of absorbing games. Such games provide a useful vehicle for conveying some of the ideas that will be used in this paper. Formally, an absorbing game is a quitting game if $A^i = \{C, Q\}$ for every player $i \in N$, $w(C, C, \dots, C) = 0$ and $w(\cdot) = 1$ otherwise.

For simplicity, we denote the absorbing payoff if all players in subset S (but none in its complement) quit, by u_S. If $S = \{j\}$ is a singleton, we denote it by u_j. The payoff to i if j alone quits is u^i_j. The non-absorbing payoff $r(C, C, \dots, C)$ is denoted simply by r. The non-absorbing payoff of all other action combinations is irrelevant for the analysis in this section.

DEFINITION 4.1 *A player* $i \in N$ *is* **punishable** *if there exists some* $j \neq i$ *with* $u^i_j \leq u^i_i$.

Player i is punishable if there is another player who, by quitting alone, can reduce the payoff of i to at least as much as i can get by quitting himself alone.

LEMMA 4.2 *If the game does not admit an equilibrium payoff then there is a probability distribution $\mu \in \Delta(N)$ that satisfies:*

1. *If $\mu_i > 0$ then i is punishable.*
2. *For every player i, $\sum_{j \in N} \mu_j u^i_j \geq u^i_i$.*

This lemma will be proven later, in the context of general absorbing games. To get a sense of its importance consider the second condition first. It says that if every player were to quit alone according to the distribution μ, they would all make more than they could by quitting unilateraly. However, it might be the case that a player may profit by not quitting (if his non-absorbing payoff is high). This is where the first condition matters.

LEMMA 4.3 *If there exists a probability distribution $\mu \in \Delta(N)$ that satisfies the conditions of Lemma 4.2, then the game admits a correlated equilibrium payoff.*

To prove the lemma, we need a result in probability. Assume a probability distribution $\lambda = (\lambda_i)$ over N is given. We wish to choose for each player i a positive integer m_i such that

1. $m_i = \min_k m_k$ for exactly one player.
2. Given m_j, the probability that $m_i = \min_k m_k$ is equal to λ_i for every $i \in N$.
3. Given m_j, the probability that $\min_k m_k = b$ is small for every $b < m_j$.

We prove that for every $\epsilon > 0$ there exists a probability distribution q over $\mathbf{N}^N$ (with finite support) such that 1) holds q-a.s., and 2) and 3) hold with probability of at least $1 - \epsilon$.

LEMMA 4.4 *For every probability distribution $\lambda = (\lambda_i)$ over N and every $\epsilon > 0$ there exists a positive integer $M \in \mathbf{N}$ and a probability distribution q over $\{1, \ldots, M\}^N$ such that every random variable $X = (X_i)_{i \in N}$ to $\{1, \ldots, M\}^N$ that has the distribution q satisfies:*

1. $\mathbf{P}_q(\exists i \neq j \text{ s.t. } X_i = X_j = \min_k X_k) = 0$.
2. *For every* $i, j \in N$, $\mathbf{P}_q\left(\mathbf{P}_q(X_j = \min_k X_k \mid X_i) = \lambda_j\right) \geq 1 - \epsilon$.
3. *For every* $i \in N$ *and every* $b < X_i$, $\mathbf{P}_q\left(\mathbf{P}_q(\min_k X_k = b \mid X_i) < \epsilon\right) \geq 1 - \epsilon$.

Proof: Let $c > 1/\epsilon$ and $d > c/\epsilon$ be fixed integers. Denote $M = c + d$. Let $Y_1, \ldots, Y_n$ be i.i.d. uniformly distributed random variables over $\{1, \ldots, c\}$. Let L be a uniformly distributed random variable over $\{1, \ldots, d\}$. Let I be a random variable over N with the distribution $\lambda = (\lambda_i)$. Define the random variable $X = (X_i)$ as follows:

$$\begin{aligned} X_I &= L \\ X_j &= L + Y_j \quad j \neq I \end{aligned}$$

Denote by q the joint distribution of $(X_i)_{i \in N}$. Observe that $\min_k X_k = L$. It is clear that requirement 1) holds.

$$\begin{aligned} \mathbf{P}(X_i = L \mid X_i = a) &= \mathbf{P}(L = a \mid X_i = a) \\ &= \frac{\mathbf{P}(X_i = a \mid L = a)\mathbf{P}(L = a)}{\mathbf{P}(X_i = a)} \\ &= \frac{\lambda_i \mathbf{P}(L = a)}{\lambda_i \mathbf{P}(L = a) + (1 - \lambda_i)\sum_{t=1}^{a-1} \mathbf{P}(L = a - t, Y_i = t)} \\ &= \frac{\lambda_i \mathbf{P}(L = a)}{\lambda_i \mathbf{P}(L = a) + (1 - \lambda_i)\sum_{t=1}^{\min\{c, a-1\}} \mathbf{P}(L = a - t, Y_i = t)}. \end{aligned}$$

For every $a \in \{c+1, c+2, \ldots, d\}$ and every $t \in \{1, \ldots, c\}$, $\mathbf{P}(L = a - t, Y_i = t) = 1/cd$ and $\mathbf{P}(L = a) = 1/d$, hence for every such a, $\mathbf{P}(X_i = L \mid X_i = a) = \lambda_i$. Since $c/d < \epsilon$, $\mathbf{P}(a \in \{c+1, c+2, \ldots, d\}) > 1 - \epsilon$, and 2) is proved for $j = i$.

By the construction of X, for every $j, k \neq i$,

$$\frac{\mathbf{P}(X_j = L \mid X_i = a)}{\mathbf{P}(X_k = L \mid X_i = a)} = \frac{\lambda_j}{\lambda_k}$$

and 2) follows for every i, j. If $X_i > 1/\epsilon$ then 3) holds, and this event occurs with probability larger than $1 - \epsilon$. ■

Proof of Lemma 4.3: Fix an $\epsilon > 0$. Let μ be a probability distribution satisfying the assumptions. For each punishable player $i \in N$ let $j(i) \in N$ be a player such that $u^i_{j(i)} \leq u^i_i$. Player $j(i)$ is the **punisher** of player i. Let $S_j = \{i \mid j = j(i)\}$ be the set of players that j punishes.

Let $\lambda_j = \sum_{i \in S_j} \mu_i$ be the probability that j is a punisher. Then $\lambda = (\lambda_j)$ is a probability distribution over N. Let q be the probability distribution over $\mathbf{N}^N$ defined by Lemma 4.4 w.r.t. λ and ϵ. The signal space of each player i is the set $\{1, \ldots, n_0\}$, where n_0 is any integer larger than $M + 1/\epsilon$ and M is defined in Lemma 4.4.

The correlation device chooses a vector $m = (m_i) \in \mathbf{N}^N$ according to q. Let $j = \text{argmin} m_i$ be the player with minimal integer. Call this player the punisher. The device now chooses a player i (the designated quitter) in S_j according to the induced probability distribution μ over S_j. Finally, the device chooses a positive integer d uniformly distributed in $\{1, \ldots, D\}$, where D is the smallest integer greater than $1/\epsilon$.

Player i receives the signal $s^i = d$. All other players k receive the signal $s^k = D + m_k$. Note that $\mathbf{P}(s^i = \min_k s^k) = \mu_i$. The signals can be interpreted as specifying the stage at which the receiver should quit.

We now define the strategy profile in the extended game. For every player j define $\sigma^j(s^j, n) = 1_{s^j = n} Q + 1_{s^j \neq n} C$; that is, player j quits with probability 1 at the stage which is recommended to him by the correlation device, and continues with probability 1 at all other periods.

If the players follow the strategy profile $\sigma = (\sigma^j)$ then their expected payoff is $\sum_{j \in N} \mu_j u_j$. An informal argument for why the players should heed the signals is that given his signal, the chosen one knows that if he does not, he will be punished by his punisher. All other players k, however, know only that they are possible punishers and with high probability they have no further information on the identity of the designated quitter.

Clearly if $\mu_i = 1$ for some i, then the identity of the designated quitter is known, and σ is an ϵ-equilibrium.

Denote by $\gamma^i_n(s^i, \sigma)$ the expected payoff of player i in the first n stages given his signal is s^i.

To verify that no player can profit more than ϵ by deviating fix a player i who receives the signal s^i. Consider a pure deviation τ^i of player i, where instead of quitting at stage s^i he quits at stage m.

First assume that $s^i \leq D$; that is, player i is the designated quitter. In that case, $\gamma^i_n(s^i, \sigma) = u^i_i$ for every $n \geq n_0$.

If $m \leq D$, we still have $\gamma^i_n(s^i, \sigma^{-i}, \tau^i) = u^i_i$. If on the other hand $m > D$, then $\mathbf{P}(m = \min_{j \neq i} s^j \mid s^i) < \epsilon$. Since payoffs are bounded by 1, for every $n \geq n_0$ $\gamma_n^i(s^i, \sigma^{-i}, \tau^i)$ is bounded by

$$\mathbf{P}(m < \min_{j \neq i} s^j \mid s^i) u_i^i + \mathbf{P}(m = \min_{j \neq i} s^j \mid s^i) + \mathbf{P}(m > \min_{j \neq i} s^j \mid s^i) u_{j(i)}^i \leq u_i^i + \epsilon.$$

Second, assume that $s^i = D + m_i$, and let $n \geq n_0$. Then $\gamma_n^i(s^i, \sigma) = \sum_{j \neq i} \mu_j u_j^i / (1 - \mu_i) \geq \sum_{j \in N} \mu_j u_j^i \geq u_i^i$. By Lemma 4.4(2), with probability higher than $1 - \epsilon$ player i has no information about the identity of the designated quitter. Thus, with probability at least $1 - \epsilon$, if $m \leq D$ then for every $n \geq n_0$ $\gamma_n^i(s_i, \sigma^{-i}, \tau^i)$ is bounded by

$$\mathbf{P}(m < \min_j s^j \mid s^i) u_i^i + \mathbf{P}(m = \min_j s^j \mid s^i) + \mathbf{P}(m > \min_j s^j \mid s^i) \sum_{j \neq i} \mu_j u_j^i / (1 - \mu_i)$$

$$\leq \sum_{j \neq i} \mu_j u_j^i / (1 - \mu_i) + \epsilon = \gamma_n^i(s^i, \sigma) + \epsilon.$$

If $m > D$, then $\gamma_n^i(s^i, \sigma^{-i}, \tau^i) = \gamma_n^i(s^i, \sigma)$.

Thus, with probability at least $1 - \epsilon$ player i cannot profit more than ϵ by deviating in any sufficiently long game. It follows that with probability at least $1 - |N|\epsilon$ no player can profit more than ϵ by any deviation. Since ϵ is arbitrary, $\sum_{j \in N} \mu_j u_j$ is a correlated equilibrium payoff. ■

4.1 Probabilistic Quitting Games

A **probabilistic quitting game** is a quitting game where, if a subset S of players quit at some stage, the game is absorbed with a positive probability $w_S > 0$, which may be strictly less than 1. In quitting games it suffices for one player to get the signal that he is the quitter. In probabilistic quitting games, even if the 'designated quitter' quits the game can continue. Once he quits, his identity is revealed to everyone. Since some players may get low payoff if the game is actually terminated by the designated quitter, a new designated quitter must be chosen. Since 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. As we will see, the fact that the game might continue even if someone quits does not pose any difficulty.

LEMMA 4.5 *If there exists a probability distribution $\mu \in \Delta(N)$ that satisfies the conditions of Lemma 4.2, then the probabilistic quitting game admits a correlated equilibrium.*

Proof: Fix $\epsilon > 0$ sufficiently small. Let $s(1), s(2), \ldots$ be an infinite sequence of independent outcomes of the correlation device described in the proof of Lemma 4.3. Thus, for each k, $s(k) \in \{1, \ldots, M + D\}^N$. The signal of player i is the sequence $(s^i(1), s^i(2), \ldots)$.

Intuitively, players play as in the proof of Lemma 4.3 with the signal $s(1)$. If the designated player does not quit, all players realize that someone deviates at stage $D + 1$. At stage t, where t is the stage in which the punisher according to $s(1)$ should quit, the identity of the punisher is revealed. From that stage on, the punisher quits at every stage with probability ϵ, and everyone else continues.

The punisher quits with probability ϵ at every stage, rather than with probability 1, so that the designated quitter would not know when the punisher will actually quit, and use this information to profit by quitting exactly at the same stage.

If the designated player indeed quits and the game does not terminate, the players forget the history, and they play as in the proof of Lemma 4.3 with the signal $s(2)$, and so on until absorption occurs.

Since there is a probability $|N|\epsilon$ that given $s(k)$ some player may get some information on the identity of the designated quitter, and this probability aggregates, ϵ should be small compared to $\min_{S \neq \emptyset} w_S$.

Let us now verify that no player i can profit by a unilateral deviation.

Clearly if $\mu_j = 1$ for some player j then σ is an ϵ-equilibrium. If the game terminates by the kth designated quitter (that is, the quitter designated by $s(k)$), the payoff to i is u^i_i if i is the kth designated quitter, and $\sum_{j \neq i} \mu_j u^i_j / (1 - \mu_i) \geq u^i_i$ if i is not the kth designated quitter. Hence the expected payoff to i by following the equilibrium strategy is at least u^i_i whatever signal he receives, which is, as we saw in the proof of Lemma 4.3, the most (up to a small error) he can get by a unilateral deviation. ∎

5 Non-Absorbing Profiles

Non-absorbing profiles generalize the 'everyone continues' profile in quitting games. This section will do three things:

1. We catalogue the different kinds of non-absorbing profiles into four groups.

2. We study the way information can be transmitted between players in absorbing games.

3. We identify sufficient conditions for the existence of correlated equilibrium in each kind of non-absorbing profile. Many of them will have the flavor of Lemma 4.3. They will consist of two parts. First the non-absorbing profile induces a distribution over payoffs with certain incentive properties. Second, 'deviations' from this non-absorbing profile can be 'punished'.

In the last section of this paper we show that every absorbing game admits either an equilibrium payoff, **or** a non-absorbing profile where one of these sufficient conditions must hold, so proving the main result.

5.1 Exits and Equilibrium

In this section we reproduce from Solan (1999, 1997) sufficient conditions for the existence of an equilibrium payoff in absorbing games. First some notation.

DEFINITION 5.1 *The real number $v^i \in \mathbf{R}$ is the* **min-max value** *of player i if for every $\epsilon > 0$ 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:*

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

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

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

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

Thus, players $N \setminus \{i\}$ 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 (1988) for N-player stochastic games.

Remark: In our construction, a deviator is punished with the min-max level and not by the max-min level. 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 n-player stochastic games. The cost is that the definition of punishment level is slightly more involved, and the proof (of Lemma 6.4) is longer.

The multi-linear extensions of r to X are 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.

Definition 5.2 *Let $\gamma \in \mathbf{R}^N$ be a payoff vector. A mixed action combination x is* **individually rational for** γ *if for every player $i \in N$ $\gamma^i \geq v^i$ and for every action $a^i \in A^i$ such that $w(x^{-i}, a^i) > 0$,*

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

x is individually rational for γ if no player i can get more than γ^i by an absorbing deviation.[1]

In absorbing games it is sometimes the case that absorption requires coordinated action on the part of a group of two or more players.

Definition 5.3 *Let $x \in X$ be a non-absorbing mixed-action combination. An* **exit** *(w.r.t. x) is a vector $a^L \in A^L$ such that (i) $\emptyset \subset L \subseteq N$, (ii) $w(x^{-L}, a^L) > 0$, and (iii) $w(x^{-L'}, a^{L'}) = 0$ for every proper subset L' of L.*

We denote by $E(x)$ the set of all exits w.r.t. x. If $L = \{i\}$ contains a single player, we 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**. For every probability distribution $\mu \in \Delta(E(x))$ we define the **expected absorbing payoff** given μ by

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

To motivate the importance of the sufficient conditions we will present, we recall from Solan (1999, Theorem 4.5) that in any absorbing game there is a profile x_0 which satisfies one of the following conditions:

[1] Usually, deviations can be followed by punishment with the min-max level, hence one gets a stronger definition of individual rationallity (see Solan (1997)). In our context players may not know the identity of the deviator, hence the deviator may deviate several times without being detected.

1. x_0 is non-absorbing and individually rational for $r(x_0)$.

2. x_0 is absorbing, individually rational for $u(x_0)$, and $u^i(x_0) = u^i(x_0^{-i}, a^i)$ for every $i \in N$ and every $a^i \in \text{supp}(x_0^i)$ such that $w(x_0^{-i}, a^i) > 0$.

3. x_0 is absorbing and there exists a probability distribution $\mu \in E(x_0)$ such that x_0 is individually rational for $u(\mu)$.

Lemma 5.4 below shows that G admits an equilibrium in the first case. By Lemma 5.5 the same holds for the second case. The third case meets all but one of the conditions of Lemma 5.6. This is the interesting case that requires special attention. Proofs of Lemmas 5.4 and 5.5 can be found in Vrieze and Thuijsman (1989) or Solan (1999). The proof of Lemma 5.6 can be found in Solan (1997). Is is also a special case of Solan (1999, Lemma 5.3). Here we only give the intuition.

LEMMA 5.4 *If $x \in X$ is non-absorbing and individually rational for $r(x)$ then $r(x)$ is an equilibrium payoff.*

The equilibrium strategies are as follows. Each player i plays the stationary strategy x^i, while checking for deviation of his opponents. Those checks include:

- Whether the realized action of each player j are compatible with x^j.

- Whether the distribution of the realized actions of each player j is approximately x^j.

The first player who fails one of these tests (or the player with minimal index in the case that more than one player fails these tests at the same stage) is punished forever by his min-max level.

With an additional condition one can extend lemma 5.4 to the case when x is absorbing.

LEMMA 5.5 *Let $x \in X$ be an absorbing mixed action combination, that is (i) individually rational for $u(x)$ and (ii) satisfies $u^i(x) = u^i(x^{-i}, a^i)$ for every $i \in N$ and every $a^i \in \text{supp}(x^i)$ such that $w(x^{-i}, a^i) > 0$. Then $u(x)$ is an equilibrium payoff.*

The equilibrium strategies are the same as above. The conditions imply that $u(x)$ is an equilibrium payoff.

Lemma 5.6 *If there exists a non-absorbing profile $x \in X$, and a probability distribution $\mu \in E(x)$ such that x is individually rational for $u(\mu)$ and $\mu[a^L] > 0$ implies that $|L| \geq 2$ then $u(\mu)$ is an equilibrium payoff.*

Let $\text{supp}(\mu) = \{a_1^{L_1}, \ldots, a_K^{L_K}\}$ and $\epsilon > 0$ be given, and let $\delta > 0$ be sufficiently small. We construct an ϵ-equilibrium profile, where the equilibrium path has a cycle of length K. At stage t, the players try to be absorbed with a small probability δ by the exit $a_k^{L_k}$, where $k = t \bmod K$. In fact each player $i \notin L_k$ plays x^i, whereas each player $i \in L_k$ plays $(1-\delta_k)x^i + \delta_k a_k^i$, where $\delta_k = (\frac{\delta\mu[a_k^{L_k}]}{w(x^{-L_k}, a_k^{L_k})})^{1/|L_k|}$. Note that $w(x^{-L_k}, a_k^{L_k})\delta_k^{|L_k|} = \delta\mu[a_k^{L_k}]$. Hence, at stage k the probability that the exit $a_k^{L_k}$ is used is $\delta\mu[a_k^{L_k}]$. In particular, if the players follow this profile and δ is sufficiently small then the game will be absorbed with high probability after $1/\delta^2$ stages, and the expected payoff is approximately $u(\mu)$.

Since players need not be indifferent between the various exits in $\text{supp}(\mu)$, suitable statistical tests are needed to make deviation non-profitable, and they can be performed effectively since $|L| \geq 2$ for every $a^L \in \text{supp}(\mu)$.

For the statistical test, the players consider at stage t only stages $j < t$ such that $j = t \bmod K$. All other stages are ignored.

- Check for each i if i's realized action is compatible with this profile; that is, if it is in $\text{supp}(x^i)$, and, if $i \in L_k$, it may be also a_k^i.
- Check for each i if the distribution of i's realized actions, when restricted to $\text{supp}(x^i)$, is approximately x^i.
- For each player $i \in L_k$ check if i plays the action a_k^i with frequency δ_k. Formally, the realized frequency p that player i plays a_k^i at stages $j < t$ such that $j = t \bmod K$, should satisfy $|\frac{p}{\delta_k t/K} - 1| < \epsilon$.

The first two tests are used in the previous sufficient conditions as well, and, if δ is sufficiently small, can be employed effectively. It was proven in Solan (1999 or 1997) or in Solan and Vieille (1998a) that the third statistical test can be employed effectively if δ is sufficiently small.

■

Notice that Lemma 5.6 does not apply to profiles that admit unilateral exits. The remainder of this paper deals with precisely this situation. We will show that absorbing profiles that admit unilateral exits can be put into one of four categories. For two of these categories, a modification of Lemma 5.6 holds. For the other two categories, an additional condition having to do with punishability is needed.

We close this subsection with another sufficient condition for existence of an equilibrium payoff, that was established by Vrieze and Thuijsman (1989) for $N = 2$ and by Solan (1997) for general N.

LEMMA 5.7 *Let x be a non-absorbing action combination. If there exists a player $i \in N$ and an action $b^i \in A^i$ such that (i) $w(x^{-i}, b^i) > 0$, (ii) x is individually rational for $u(x^{-i}, b^i)$ and (iii) $r^i(x) \leq u^i(x^{-i}, b^i)$, then $u(x^{-i}, b^i)$ is an equilibrium payoff.*

$u(x^{-i}, b^i)$ is an equilibrium payoff even if $r^i(x) \leq u^i(x^{-i}, b^i)$ does not hold, but in our setup the second inequality holds whenever this lemma is invoked.

Consider the following profile:

- All $j \neq i$ play x^j in each round.
- Player i plays x^i with probability $1 - \delta$ and b^i with probability δ in each round.

Here $\delta > 0$ is chosen sufficiently small.

If the players follow this profile the expected payoff for the players is $u(x^{-i}, b^i)$ in any sufficiently long game. To deter deviations, check for each player j, that the distribution of his realized actions is approximately x^j, and that his realized actions are compatible with this profile. The first player who fails the statistical test (or the player with minimal index, if more than one player fail the test at the same stage) is punished with his min-max value forever.

Since $r^i(x) \leq u(x^{-i}, b^i)$, player i is worse off by not letting the play terminate by b^i. Since x is individually rational for $u(x^{-i}, b^i)$, no player can profit too much by deviating.

5.2 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. We state sufficient conditions that allow for the transmission of public information.

DEFINITION 5.8 *Let $x \in X$ be a non-absorbing profile and $i \in N$ a player. Player i is a* **signaller** *w.r.t. x if for every finite message set M and every $\epsilon > 0$ 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 $\mathcal{P} = (P_m)_{m \in M}$ of H_{n_0} such that*

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

Thus the signaller 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 $\mathcal{P}$ 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.

In our context, players in $N \setminus \{i\}$ have an encrypted message, but only player i has the decryption key. Since player i is ignorant of the contents of the message, no player knows the content of the message. Once player i publicly transmits the key, every player in $N \setminus \{i\}$ can read the message.

LEMMA 5.9 *Let x be a non-absorbing profile and $i \in N$ a player. If there exist two mixed actions $y^i_1, y^i_2 \in X^i$ that satisfy*

- $w(x^{-i}, y^i_k) = 0$ *for $k = 1, 2$.*
- $\| y^i_k - x^i \|_\infty < \epsilon/2$ *for $k = 1, 2$.*
- $\| y^i_1 - y^i_2 \|_\infty > \epsilon/4$.

then i is a signaller w.r.t. x.

Proof: Intuitively, player i who wants to send a message $m \in \{1, \ldots, M\}$ will send a message of M bits, all 0 except one that corresponds to the message m. The mixed action y^i_1 is used to transmit the bit 1, and y^i_2 is used to transmit the bit 0.

Let $M = \{1, \ldots, M\}$ be a finite message space and $\epsilon > 0$ be fixed. For $k = 1, 2$, let n_k be sufficiently large such that for every sequence of i.i.d. r.v. (Y_j) with distribution y^i_k we have for every $n > n_k$,

$$\mathbf{P}(\| (Y_1 + \cdots + Y_n)/n - y^i_k \|_\infty > \epsilon/8) < \epsilon/|M|.$$

Let $n_* = \max\{n_1, n_2\}$ and $n_0 = |M|n_*$.

For every $m \in M$ define a strategy σ^i_m as follows.

$$\sigma^i_m(h) = \begin{cases} y^i_1 & h \in H_n, n_*(m-1) < n \leq n_* m \\ y^i_2 & \text{otherwise.} \end{cases}$$

Define a partition $\mathcal{P} = (P_m)_{m \in M}$ of H_{n_0} as follows. For each history $h \in H_{n_0}$ let $a^i_t(h)$ be the action taken by player i at stage t according to history h. A history h is in P_m if (i) $\| \sum_{t=n_*(m-1)+1}^{n_* m} a^i_t(h)/n_* - y^i_1 \|_\infty < \epsilon/8$ and (ii) for every $m' \neq m$, $\| \sum_{t=n_*(m'-1)+1}^{n_* m'} a^i_t(h)/n_* - y^i_2 \|_\infty < \epsilon/8$. All histories that do not satisfy any of these conditions are divided arbitrarily.

By construction,

$$\mathbf{P}_{x^{-i},\sigma^i_m}\left(\| \sum_{t=n_*(m-1)+1}^{n_* m} a^i_t(h)/n_* - y^i_1 \|_\infty < \epsilon/8\right) \geq 1 - \epsilon/|M|$$

and for every $m' \neq m$

$$\mathbf{P}_{x^{-i},\sigma^i_m}\left(\| \sum_{t=n_*(m-1)+1}^{n_* m} a^i_t(h)/n_* - y^i_2 \|_\infty < \epsilon/8\right) \geq 1 - \epsilon/|M|.$$

Thus, $\mathbf{P}_{x^{-i},\sigma^i_m}(P_m) \geq 1 - \epsilon$, as required. ∎

Corollary 5.10 *Let x be a non-absorbing profile and $i \in N$ a player. If either (i) $|\mathrm{supp}(x^i)| \geq 2$ or (ii) there exists an action $a^i \notin \mathrm{supp}(x^i)$ such that $w(x^{-i}, a^i) = 0$, then player i is a signaller w.r.t. x.*

Proof: If (i) holds, there exists $a^i \in \mathrm{supp}(x^i)$ such that $x^i[a^i] \leq 1/2$. Define $y^i_1 = x^i$ and $y^i_2 = (1 - \epsilon/2)x^i + (\epsilon/2)a^i$ (interpret this to mean play x^i with probability $1 - \epsilon/2$ and a^i with probability $\epsilon/2$). If (ii) holds, let $y^i_1 = x^i$ and $y^i_2 = (1 - \epsilon/2)x^i + (\epsilon/2)a^i$. ∎

It is clear that the condition in Corollary 5.10 is also a necessary condition. Indeed, otherwise, there is only one history that is non-absorbing when $N \setminus \{i\}$ play x^{-i}.

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 some 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 always be clear from the context what is the stationary profile x to be used.

During the signalling period, players who are not signallers may deviate in two ways. They can either alter the frequency in which they play actions in $\mathrm{supp}(x^i)$, or they can play actions outside $\mathrm{supp}(x^i)$. The second type of deviations 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 deviations can be deterred. The first type of deviations does not change the message that is sent, since $\mathcal{P}$ depends only on the actions of the signallers.

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

Definition 5.11 *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 there exist $a^{i_1} \in A^{i_1}$ and $a^{i_2} \in A^{i_2}$ 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 will see later, with the help of the signaller he can transmit information.

5.3 Classification of Non-Absorbing Profiles

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

Definition 5.12 *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.*

We assign no appelation 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.

For example, consider the following two-player absorbing games where each player has 2 actions, and only the absorbing structure is given (an asterisked cell means that the probability of absorption is positive, and a non-asterisked cell 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 $\mathrm{supp}(x)$ and $\mathrm{supp}(y)$ 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 action, then x admits at least two signallers. In particular, if x is isolated, semi-isolated or weak, $E(x)$ includes only unilateral exits.

In section 5.4 we deal with isolated actions, in section 5.5 with semi-isolated profiles, and in section 5.6 with all other non-absorbing profiles.

5.4 Isolated Actions

In this section we consider isolated actions. We show that the punishability condition stated for probabilistic quitting games can be used in this setup. To make the analogy complete we define what it means to be punishable in this context.

Fix an isolated action a. Since the function u is generic, each player i has a unique action $b^i(a)$ that maximizes the expression $u^i(a^{-i}, d^i)$ over $d^i \neq a^i$. If the other players play a^{-i} then $b^i(a)$ is i's best absorbing response. Let $g^i(a) = u^i(a^{-i}, b^i(a))$ be the best absorbing response payoff: the maximum that player i can get by a unilateral deviation from a. Set $p^i(a) = \min_{j \neq i} u^i(a^{-j}, b^j(a))$ to be the **punishment level** of player i, and let $j_i(a)$ be a player that minimizes this expression. Player $j_i(a)$ is the **punisher** of i at a. Player $i \in N$ is **punishable** at a if $p^i(a) \leq g^i(a)$.

Thus, in an isolated action a player may punish an opponent only by his best absorbing response.

The proof of Lemma 4.5 yields:

LEMMA 5.13 *Let a be an isolated action. If there exists a probability distribution $\mu \in \Delta(N)$ that satisfies the following two conditions:*

1. *If $\mu_i > 0$ then i is punishable at a.*
2. *For every player i, $\sum_{j \in N} \mu_j u^i(a^{-j}, b^j(a)) \geq g^i(a)$.*

then $\sum_{i \in N} \mu_i u(a^{-i}, b^i(a))$ is a correlated equilibrium payoff.

To build equilibria around action combinations that are not isolated we need players to signal to each other. How this is done is described next.

5.5 Semi-isolated Profiles

In this section we study semi-isolated profiles. We define the punishability notion for these profiles, and we give a condition on existence of correlated equilibrium payoff using the signalling strategy we described earlier.

5.5.1 On Punishments

First we extend the notion of punishment level and punisher to semi-isolated actions. As discussed in example 3, in a semi-isolated profile, the identity of the designated quitter will be revealed to everyone but the signaller i_0. In particular, player i_0 can be punished with his min-max level, whereas any player $j \neq i_0$ can be punished either by i_0 or jointly by $N \setminus \{j, i_0\}$. It is easier to restrict ourselves to punishments by single players.

Let x be a semi-isolated profile with unique signaller i_0. Then $|\text{supp}(x^j)| = 1$ for every $j \neq i$. Since the game is generic, there is a unique action $b^{i_0}(x)$ that maximizes $u^{i_0}(x^{-i_0}, d^{i_0})$ over $d^{i_0} \in A^{i_0}$. For each player $i \in N$ define $g^i(x) = \max_{d^i \in A^i} u^i(x^{-i}, d^i)$ to be the maximal absorbing level of player i given that the other players follow x. Notice that $g^{i_0}(x)$ is independent of x^{i_0}, and depends only on x^{-i_0}.

Any player $i \in N$ can be punished by any player $j \neq i$. Define

$$p^i_{i_0}(x) = u^i(x^{-i_0}, b^{i_0}(x))$$

and for every $j \neq i, i_0$

$$p^i_j(x) = \min_{d^j \neq x^j} u^i(x^{-j}, d^j). \tag{3}$$

The definition reflects the idea that any player $j \neq i$ knows the identity of i, and therefore can choose the action that punishes i the most, whereas player i_0 does not know the identity of the punished one, so he must use only one action to punish.

Finally, define the punishment level of each player $i \neq i_0$ by

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

and the punishment level of i_0 by

$$p^{i_0}(x) = \min\{r^{i_0}(x), \min_{j \neq i_0} p^{i_0}_j(x)\}.$$

That is, players $N \setminus \{i_0\}$ can either punish player i_0 by some absorbing action, or by never absorbing, whichever yields i_0 a lower payoff. Player i is **punishable** at x if $p^i(x) \leq g^i(x)$.

Let $j_i(x)$ be the **punisher** of player i at x; that is, it is the player j that minimizes $p^i_j(x)$ over all $j \neq i$. Denote by $b^{j_i(x)}_i(x)$ the action of player $j_i(x)$ that minimizes the payoff of i.

5.5.2 A Sufficient Condition

LEMMA 5.14 *Let x be a semi-isolated profile with signaller i_0 and $\mu \in \Delta(E(x))$. Assume that the following conditions hold.*

1. *If $i \neq i_0$ and $\sum_{a^i \in E(x)} \mu[a^i] > 0$ then i is punishable at x.*
2. *x is individually rational for $u(\mu)$.*
3. *If $a^i, b^i \in E(x)$ are two unilateral exits of i and $\mu[a^i] > 0$ then $u^i(x^{-i}, a^i) \geq u^i(x^{-i}, b^i)$.*

Then $u(\mu)$ is a correlated equilibrium payoff.

Proof: Note that since x is semi-isolated, μ is supported by unilateral exits. Let $\mu_i = \sum_{a^i \in E(x)} \mu[a^i]$ be the overall weight of unilateral exits of i in μ. Then $(\mu_i)_{i \in N}$ is a probability distribution. If $\mu_i > 0$, define $y^i = \sum_{a^i \in E(x)} \frac{\mu[a^i]}{\mu_i} a^i$ to be the probability distribution induced by μ over the unilateral exits of player i.

We define the following mechanism:

1. The correlation device chooses a player i according to the probability distribution $(\mu_i)_{i \in N}$.
2. The device sends i the signal "you have been chosen".
3. The device chooses a **verification key** $v \in \{1, \ldots, K\}$, where $K \geq 1/\epsilon$, with the uniform distribution.
4. The device chooses an **encrypting key** $k \in \{1, \ldots, K\}$ with the uniform distribution.
5. If $i \neq i_0$, the device sends to i_0 both v and k. If $i = i_0$ it does not.
6. The device sends to each player $j \neq i, i_0$ both v and the sum $k + i \bmod K$.
7. In the first $1/\epsilon^2$ stages,[2] all unchosen players j play x^j, and player i plays the mixed action $(1-\epsilon)x^i + \epsilon y^i$.

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

If the players follow this mechanism, the expected payoff if absorption occurs is $u(\mu)$, and absorption occurs with probability greater than $1-\epsilon$. Moreover, for every unchosen player $j \neq i$ and every $j' \neq j$, the probability that $j' = i$ conditional on the information revealed to player j by the device is $\mu_i/(1-\mu_j)$. Therefore by condition 2 no player $j \in N \setminus \{i, i_0\}$ has a profitable deviation.

If $i \neq i_0$, player i_0 might profit by deviating from x^{i_0}. To make such a deviation non-profitable, we add a standard statistical test: players in $N\setminus\{i_0\}$ check whether the distribution of the realized actions of i_0 is approximately x^{i_0}. If a deviation is detected, they punish him with his min-max value.

By condition 3, player i cannot profit by altering the probability in which he plays absorbing actions.

It is in the interest of the chosen player to never play an absorbing action if $r^i(x) > g^i(x)$. If absorption has not occurred in the first $1/\epsilon^2$ stages, the identity of the chosen player should be revealed so that he can be punished. To accommodate this we add the following instructions to the mechanism. If the chosen one did not play an absorbing action until stage $1/\epsilon^2$, do the following.

8. Player i_0 publicly sends v. Denote by v' the actual message sent.
9. If $v \neq v'$, players $N \setminus \{i_0\}$ punish i_0 with his min-max value.
10. If $v' = v$, player i_0 publicly sends k. Now every player $j \neq i_0$ knows the identity of i.
11. For $1/\epsilon^2$ stages all players j that are *not* the punisher of i play x^j, and player i_0 plays x^{i_0}. If the punisher $j_i(x)$ of i is **not** i_0, then this player plays $(1-\epsilon)x^{j_i(x)} + \epsilon b_i^{j_i(x)}(x)$.
12. If absorption has not occurred yet within $1/\epsilon^2$ stages (which happens through the luck of the draw or if $j_i(x) = i_0$), from then on player i_0 plays $(1-\epsilon)x^j + \epsilon b^{i_0}(x)$, and all the other players play x^{-i_0}.

In step 8 player i_0 reveals whether he is the chosen one or not. If he is not the chosen one, the probability that he can duplicate v is smaller than ϵ.

The only thing left to specify is, what happens if the chosen one plays a unilateral exit and the game is not absorbed. As in the proof of Lemma 4.5, the device actually chooses an i.i.d. sequence of chosen players, verification keys and encrypting keys, one for each attempt to use an exit. If play is not terminated by the first chosen player, it is the turn of the second chosen player to use one of his unilateral exits, and so on until the game terminates.

Let us now verify that the mechanism induces a correlated $R\epsilon$-equilibrium for some fixed $R \in \mathbf{R}$. As in the proof of Lemma 4.5, since the sequence the device chooses is i.i.d. and by condition 2, it is sufficient to check that no player can profit by deviating in the first round of this mechanism.

Let us first check that the chosen player i cannot profit by deviating. The expected payoff of i is $g^i(x)$ if absorption occurs while he uses a unilateral exit, and by condition 3 at least $g^i(x)$ if the play continues. If he does not use one of his unilateral exits with positive probability under μ, his identity is revealed in step 8 or 10, and he is punished by $p^i(x) \leq g^i(x)$. By the definition of the signalling mechanism, player i cannot profit too much during the signalling process.

Let us now verify that each unchosen player j cannot profit by a deviation. Since the probability that any player $k \neq j$ is actually the chosen one given the information of i is exactly $\mu_k/(1-\mu_k)$, the expected payoff of j is at least $u^j(\mu) \geq g^j(x)$ if the play is absorbed by the designated quitter, and at least $g^j(x)$ if it continues. The most he can get by deviating is $g^j(x)$. The only opportunity for profit is if j guesses correctly the stage in which i uses a unilateral exit, but this chance is small. As before, deviations during the signalling process cannot yield high profit. ■

5.6 Other Non-Absorbing Profiles

In this section we deal with weak profiles and non-absorbing profiles that admit at least two signallers. In these two 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.

LEMMA 5.15 *Let x be a non-absorbing profile and i_1, i_2 be two distinct signallers w.r.t. x. If there exists a probability distribution $\mu \in \Delta(E(x))$ such that for every $i \in N$*

1. *x is individually rational for $u(\mu)$.*
2. *If $a^i, b^i \in E(x)$ are two unilateral exits of i and $\mu[a^i] > 0$ then $u^i(x^{-i}, a^i) \geq u^i(x^{-i}, b^i)$.*

Then $u(\mu)$ is a correlated equilibrium payoff.

The conditions of the lemma are similar to those in Lemma 5.6, except that we have at least two signallers, and unilateral exits may have positive weight in μ. If, for every such unilateral exit we had $u^i(x^{-i}, a^i) = u^i(\mu)$ then, as was proved in Solan (1999, Lemma 5.3), $u(\mu)$ would be an equilibrium payoff. If players are indifferent to their unilateral exits, no statistical tests are needed to ensure that players use their unilateral exits as they should. In general we cannot guarantee that this condition holds, and we can only guarantee the weaker form presented here.

As seen in the proof of Lemma 5.6, 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 other players.

Proof: Recall that the profile described in the proof of Lemma 5.6 was played in rounds. Each round consists of $|\text{supp}(\mu)|$ stages, and in each stage players use a different exit in $\text{supp}(\mu)$. In Lemma 5.6 we had only joint exits, whereas here we also have unilateral exits. The mechanism we construct is similar to the one presented in Lemma 5.6, but at the end of **each** round we add a revelation period, where the signallers reveal whether any player was supposed to use one of his unilateral exits during that round. If that player has not used his unilateral exit, he is punished with his min-max value.

For each player i let $\mu_i = \sum_{a^i \in E(x)} \mu[a^i]$ be the overall probability that player i should use a unilateral exit according to μ. If μ contains joint exits, then $\sum_{i \in N} \mu_i < 1$. Let $\eta \in (0, 1)$ to be chosen later. Let (Y_t) be a sequence of i.i.d. r.v. with values in $\{0\} \cup N$ and distribution $\mathbf{P}(Y_t = i) = \eta\mu_i$ for each $i \in N$ and $\mathbf{P}(Y_t = 0) = 1 - \sum_{i \in N} \mathbf{P}(Y_t = i)$.

The interpretation of Y_t is that if $Y_t = i$ for some $i \in N$ then player i should use a unilateral exit during the t^{th} round.

Let $i_1, i_2 \in N$ be two distinct signallers.

Before start of play, each player i receives all rounds t such that $Y_t = i$. In addition, the device sends, for each round t the following:

- For each $l = 1, 2$, if $Y_t \neq i_l$, player i_l receives a uniformly distributed **verification key** $v_t^l \in \{1, \ldots, K\}$ and a uniformly distributed **encrypting key** $k_t^l \in \{1, \ldots, K\}$, where $K > 1/\epsilon$.
- For each $l = 1, 2$, each player $i \neq i_l$ receives v_t^l and $k_t^l + Y_t \mod K$.

We now recall the profile constructed in the proof of Lemma 5.6. Let $\delta \in (0, \epsilon)$ be sufficiently small. Let $\mathrm{supp}(\mu) = \{a_1^{L_1}, \ldots, a_k^{L_k}\}$, and for each k set $\delta_k = (\delta\mu[a_k^{L_k}]/w(x^{-L_k}, a_k^{L_k}))^{1/|L_k|}$.

We define a strategy σ^i in rounds. We first define only the first $|\mathrm{supp}(\mu)|$ stages of the round, which form a **quitting period**. Let t be the current round.

- If $Y_t = i$, player i chooses at the beginning of the round a unilateral exit $a^i \in E(x)$, each action a^i is chosen with probability $\mu[a^i]/\mu_i$.

At the kth stage of the round, player i plays as follows:

- If $i \notin L_k$, he plays x^i.
- If $i \in L_k$ and $|L_k| \geq 2$, he plays $(1 - \delta_k)x^i + \delta_k a_k^i$.
- If $L_k = \{i\}$ and $Y_t \neq i$, he plays x^i.
- If $L_k = \{i\}$, $Y_t = i$ and a_k^i is the unilateral exit he chose at the beginning of the round, he plays a_k^i. If a_k^i is not the unilateral action he chose, he plays x^i.

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

Since it might be in the interest of some player i to alter the frequency with which he plays different actions in $\mathrm{supp}(x^i)$, or in which he perturbs to a_k^i in stages that correspond to a joint exit, we employ the same statistical tests that are used in the proof of Lemma 5.6.

Since it might be in the interest of a player supposed to use a unilateral exit not to use it, we append a revelation period to the end of the first $|\mathrm{supp}(\mu)|$ stages **every** round t.

Revelation Period:

- Player i_1 publicly sends v_t^1.
- Player i_2 publicly sends v_t^2.
- Player i_1 publicly sends k_t^1.
- Player i_2 publicly sends k_t^2.

At the end of the revelation period all players know (with high probability) whether some player was supposed to use his unilateral exit and did not, and who it was.

Let us now verify that no player i can profit too much by deviating. Since x is individually rational for $u(\mu)$, no player can profit too much by playing an action that is not compatible with this profile, or alter the probability in which he plays actions in $\text{supp}(x^i)$.

Moreover, since the value of Y_t is revealed at the end of the round, if $Y_t = i$, player i cannot profit by not using one of his unilateral exits, and by condition 2 he is indifferent between them.

By the construction of the signalling mechanism, no player can profit too much by deviating during the signalling process.

If i is a signaller, say i_1, he can signal an incorrect signal at some round. Clearly he cannot profit by sending an incorrect verification key, but maybe he can profit by altering k and having another player punished. Such a deviation cannot happen in a round t where $Y_t = i_1$, since in that round player i_1 does not have the verification key. Since i_2 does not deviate, his report contradicts the report of i_1. Whatever the value of Y_t, the report of i_2 coincides with the realized play, whereas the report of i_1 does not, hence i_1 is declared a deviator, and is punished by his min-max value. Hence no player can profit too much by any deviation. ■

Recall that a weak non-absorbing profile x admits exactly one signaller, and at least one weak signaller. Moreover, in this case $E(x)$ contains only unilateral exits.

Lemma 5.16 *Let x be a weak non-absorbing profile. If there exists a probability distribution $\mu \in \Delta(E(x))$ such that for every $i \in N$*

1. *x is individually rational for $u(\mu)$.*
2. *If $a^i, b^i \in E(x)$ are two unilateral exits of i and $\mu[a^i] > 0$ then $u^i(x^{-i}, a^i) \geq u^i(x^{-i}, b^i)$.*

Then $u(\mu)$ is a correlated equilibrium payoff.

The proof of this Lemma is similar to the proof of Lemma 5.14. However, since we have only one signaller, the identity of the chosen player must be revealed by the signaller and the weak signaller together.

Proof: Denote by i_0 the unique signaller w.r.t. x, and by i_1 one of the weak signallers. Let $a^{i_0} \in A^{i_0}$ and $a^{i_1} \notin \text{supp}(x^{i_1})$ such that $w(x^{-i_0}, a^{i_0}) = w(x^{-i_0,i_1}, a^{i_0}, a^{i_1}) = 0$.

For each player i let $\mu_i = \sum_{a^i \in E(x)} \mu_i[a^i]$. Since μ is supported by unilateral exits, $\sum_{i \in N} \mu_i = 1$. If $\mu_i > 0$, define $y^i = \sum_{a^i \in E(x)} \frac{\mu[a^i]}{\mu_i} a^i$ to be the probability distribution induced by μ over the unilateral exits of player i.

The steps describing the correlation device from Lemma 5.14 are reproduced below with ammendments.

1. The correlation device chooses a player i according to the probability distribution $(\mu_i)_i$.

2. The device sends i the signal "you have been chosen".

3. The device chooses a **verification key** $v \in \{1, \ldots, K\}$ and an **encrypting key** $k \in \{1, \ldots, K\}$, where $K \geq 1/\epsilon$, both with the uniform distribution.

4. If $i \neq i_0$, the device sends to i_0 both v and k. If $i = i_0$ it does not.

5. The device sends to each player $j \neq i, i_0$ both v and the sum $k + i \bmod K$.

6. The device chooses $|N|$ different numbers $t_1 < t_2 < \cdots < t_N$ in the range $\{1, \ldots, T\}$ with the uniform distribution,[3] where T is sufficiently large so that $\mathbf{P}(t_N < T - 1/\epsilon) > 1 - \epsilon$. To each member of $\{i_0, i_1\}$ who was not chosen, the device sends these numbers.

7. The players play as follows for $1/\epsilon^2$ stages. Every unchosen player $j \neq i$ plays x^j, and the chosen one plays $(1 - \epsilon)x^i + \epsilon y^i$.

Next, the revelation phase is modified:

[3]That is, every increasing sequence of N numbers in this range has the same probability to be chosen.

8. Player i_0 publicly sends v. Denote by v' the actual message he sent.

9. If $v \neq v'$, players $N \setminus \{i_0\}$ punish i_0 with his min-max value.

10. If $v = v'$ player i_0 publicly sends k. Now every player $j \neq i_0$ knows the identity of i. In particular, i_1 knows it. Let t be the stage of the game where this step is over.

11. In the next T stages, the players play as follows.

    - Each player $j \neq i_0, i_1$ plays x^j.
    - If i_0 was not chosen, he plays a^{i_0} at every stage $t + t_j$. At all other stages he plays x^{i_0}.
    - If i_1 was not chosen, he plays a^{i_1} at stage $t + t_i$, 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.

At the end of the revelation phase all players know the identity of the chosen player i. Indeed, if $i = i_0$, his identity is revealed in step 9. If $i \neq i_0$, i's identity is revealed to all but i_0 in step 10. i_0 can then infer the identity of i at the end of step 11. Indeed, if at stage $t + t_k$ player i_1 plays a^{i_1}, then i_0 concludes that $i = k$, whereas if i_1 never plays a^{i_1}, i_0 concludes that $i = i_1$. If i_1 plays a^{i_1} at a stage $t \notin \{t_1, \ldots, t_N\}$, his identity is revealed (and the game might be absorbed). If he never plays it, his identity is also revealed.

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

6 An Auxiliary Game

In this section we 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 stationary profile that satisfies one of the sufficient conditions identified in the previous section.

6.1 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, 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 function $\tilde{r} : X \to \mathbf{R}^N$ and every discount factor $\lambda \in (0,1)$ we define an auxiliary discounted game $G_\lambda(\tilde{r})$, where the payoff associated with every strategy profile σ is:

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

where 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 $\tilde{r}(x_n)$.

If for every $i \in N$ the function

$$x^{-i} \mapsto \text{argmax}_{x^i \in X^i} \tilde{\gamma}^i_\lambda(x^{-i}, x^i) \tag{5}$$

has non-empty, convex values and is upper-hemi-continuous, by Kakutani's Fixed Point Theorem the game $G_\lambda(\tilde{r})$ admits a stationary equilibrium.

It is easy to see that if $\tilde{\gamma}^i_\lambda$ is continuous and for every $x^{-i} \in X^{-i}$ and 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, (5) holds.

LEMMA 6.1 *If for every $i \in N$ the function $\tilde{r}^i$ is continuous, and for every fixed $x^{-i} \in X^{-i}$ the function $\tilde{r}^i(x^{-i}, \cdot)$ is quasi-concave on X^i, then $G_\lambda(\tilde{r})$ admits a stationary equilibrium.*

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

$$\tilde{\gamma}_\lambda(x) = \frac{\lambda \tilde{r}(x) + (1-\lambda) w(x) u(x)}{\lambda + (1-\lambda) w(x)}. \tag{6}$$

In particular, $\tilde{\gamma}_\lambda$ is continuous.

Let $x^{-i} \in X^{-i}$, $x^i, y^i \in X^i$, $\beta \in [0,1]$ and $c \in \mathbf{R}$. Denote $x = (x^{-i}, x^i)$, $y = (y^{-i}, y^i)$ and $z = \beta x + (1-\beta)y$. We assume that $\tilde{\gamma}_\lambda(x), \tilde{\gamma}_\lambda(y) \geq c$, and prove that $\tilde{\gamma}_\lambda(z) \geq c$. By the previous remark, that would suffice to prove the lemma. By assumption, $\lambda\tilde{r}(x) \geq c(\lambda + (1-\lambda)w(x)) - (1-\lambda)w(x)u(x)$ and $\lambda\tilde{r}(y) \geq c(\lambda + (1-\lambda)w(y)) - (1-\lambda)w(y)u(y)$. By the linearity of w and wu, and the quasi-concavity of $\tilde{r}$, $\lambda\tilde{r}(z) \geq c(\lambda + (1-\lambda)w(z)) - (1-\lambda)w(z)u(z)$. By (6), $\tilde{\gamma}_\lambda(z) \geq c$. ∎

A similar proof proves that the ratio of two multi-linear functions with a positive denominator is quasi-concave.

So that the game $G_\lambda(\tilde{r})$ admits a stationary equilibria, we will make sure that $\tilde{r}^i$ is continuous and quasi-concave on X^i for every fixed $x^{-i} \in X^{-i}$.

Let $B \subseteq A$ be the set of all non-absorbing action combinations, and let $X' = \{x \in X \mid \text{supp}(x) \subseteq B\}$ be the collection of all the non-absorbing stationary profiles. Define a function $\tilde{r} : X' \to \mathbf{R}$ as follows:

$$\tilde{r}^i(x) = \begin{cases} p^i(x) & x \text{ is isolated} \\ p^i(x) & x \text{ is semi-isolated} \\ \min\{r^i(x), v^i\} & \text{otherwise} \end{cases} \tag{7}$$

Before proving that $\tilde{r}$ can be extended to a continuous quasi-concave function on X, we need the following result.

Recall that $B_1, \ldots, B_K$ are the maximal semi-isolated sets.

LEMMA 6.2 *Let $B_k = \times_{i \in N} B^i_k$ be a maximal semi-isolated set and let i_0 be the unique player such that $|B^{i_0}_k| > 1$. Then the function $p^{i_0} : \Delta(B^{i_0}_k) \to \mathbf{R}$ can be extended to a continuous quasi-concave function on X^i. The extended function is still denoted by p^{i_0}.*

Proof: Since the minimum of a finite number of continuous quasi-concave functions is continuous and quasi-concave, it is sufficient to prove that for every $i \neq i_0$ and every $a^i \in A^i$ such that $w(x^{-i}, a^i) > 0$, the function $u^{i_0}(x^{-i,i_0}, a^i, \cdot) : \Delta(B^{i_0}_k) \to \mathbf{R}$ can be extended to a continuous and quasi-concave function from X^{i_0}.

Since B_k is a maximal semi-isolated set and $i \neq i_0$,

$$w(x^{-i,i_0}, a^i, x^i) > 0 \qquad \forall x^{i_0} \in \Delta(B^{i_0}_k). \tag{8}$$

Every $x^{i_0} \in X^{i_0}$ such that $x^{i_0}[B_k^{i_0}] < 1$ can be uniquely decomposed to a sum $x^{i_0} = x^{i_0}[B_k^{i_0}]x_1^{i_0} + (1 - x^{i_0}[B_k^{i_0}])x_2^{i_0}$, where $\operatorname{supp}(x_1^{i_0}) \subseteq B_k^{i_0}$ and $\operatorname{supp}(x_2^{i_0}) \cap B_k^{i_0} = \emptyset$. If $x^{i_0}[B_k^{i_0}] = 1$, define $x_1^{i_0} = x^{i_0}$ and $x_2^{i_0}$ arbitrarily.

Define

$$\tilde{u}^i(x) = \frac{w(x^{-i_0}, x^{i_0})u^i(x^{-i_0}, x^{i_0}) + (1 - x^{i_0}[B_k^{i_0}])}{w(x^{-i_0}, x^{i_0}) + 1 - x^{i_0}[B_k^{i_0}]}.$$

By (8), $\tilde{u}^i$ is well defined, and clearly it agrees with u^i on $\Delta(B_k^{i_0})$.

Since wu and $x^{i_0} \mapsto x^{i_0}[B_k^{i_0}]$ are linear in x^{i_0}, one can verify as in the proof of Lemma 6.1 that $\tilde{u}^i$ is quasi-concave, and clearly it is continuous. ■

LEMMA 6.3 *For every $i \in N$ there exists a continuous, quasi-concave function $\tilde{r} : X \to \mathbf{R}$ that agrees with (7) on X'.*

Proof: Fix $i \in N$, and let $\epsilon > 0$ be sufficiently small. Let B_0 be the collection of all the non-absorbing action combinations that are neither isolated nor semi-isolated. Define

$$\tilde{r}^i(x^{-i}, x^i) = \begin{cases} p^i(x^{-i}, a^i) & \exists a^i \in A^i \text{ s.t. } (x^{-i}, a^i) \text{ is isolated} \\ p^i(x^{-i}, a^i) & \exists a^i \in A^i \text{ s.t. } (x^{-i}, a^i) \text{ is semi-isolated with signaller } i_0 \neq \\ p^i(x^{-i}, x^i) & \exists a^i \in A^i \text{ s.t. } (x^{-i}, a^i) \text{ is semi-isolated with signaller } i_0 = \\ \min\{r^i(x^{-i}, x^i), v^i\} & \exists a^i \in A^i \text{ s.t. } \operatorname{supp}(x^{-i}, a^i) \subseteq B_0 \end{cases} \quad (9)$$

Let B_{K+1} be the collection of all isolated actions. Then $B = B_0 \cup (\cup_{k=1}^k B_k) \cup B_{K+1}$ is the set of all non-absorbing action combinations.

Let B^{-i} be the projection of B on X^{-i}: $b^{-i} \in B^{-i}$ if and only if there exists $a^i \in A^i$ such that (b^{-i}, a^i) is non absorbing.

For every $x^{-i} \in X^{-i}$, $x^{-i}[B^{-i}]$ is the probability that x^{-i} gives to the set B^{-i}.

Note that $\tilde{r}$ was already defined for every x such that $x^{-i}[B^{-i}] = 1$. Define $B_k^\epsilon = \{y^{-i} \in X^{-i} \mid y^{-i}[B_k] \geq 1 - \epsilon\}$. Since all the sets $(B_k)_{k=0}^{K+1}$ are compact and disjoint, one can choose ϵ sufficiently small so that the sets $(B_k^\epsilon)_{k=0}^{K+1}$ are compact and disjoint. For every $y^{-i} \in B_k^\epsilon \setminus B_k$ there is a unique decomposition $y^{-i} = y^{-i}[B_k]y_1^{-i} + (1 - y^{-i}[B_k])y_2^{-i}$ where $\operatorname{supp}(y_1^{-i}) \subseteq B_k$ and $\operatorname{supp}(y_2^{-i}) \cap B_k = \emptyset$.

Define for every $y^{-i} \in B_k^\epsilon \setminus B_k$ and every $x^i \in X^i$

$$\tilde{r}^i(y^{-i}, x^i) = (1 - y^{-i}[B_k]/\epsilon)\tilde{r}(y_1^{-i}, x^i)$$

and for every $y^{-i} \notin \cup_{k=0}^{K+1} B_k^\epsilon$ and every $x^i \in X^i$

$$\tilde{r}^i(y^{-i}, x^i) = 0.$$

One can easily check that the function $\tilde{r}^i$ is continuous. Moreover, for every $x^{-i} \in \cup_{k=0}^{K+1} B_k$ the function $\tilde{r}^i(x^{-i}, \cdot) : X^i \to \mathbf{R}$ is quasi-concave by definition, for every $x^{-i} \notin \cup_{k=0}^{K+1} B_k^\epsilon$ it is identically 0, hence quasi-concave, and for every other x^{-i} it is a multiplication by constant of a quasi-concave function, hence quasi-concave. ∎

By Lemmas 6.1 and 6.3, 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 $i \in N$, the support $\mathrm{supp}(x_\lambda^i)$ is independent of λ. In the sequel we will assume using the same reasoning that other limits we take exist.

LEMMA 6.4 *$\tilde{\gamma}_0^i \geq v^i$ for every player $i \in N$.*

The proof of this lemma is postponed to Section 7.

6.2 Asymptotic Analysis

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 $\lim_{\lambda\to 0} \tilde{\gamma}_\lambda(y_\lambda) = u(y)$.

In this section we study the asymptotic properties of the sequence $(x_\lambda)_{\lambda\to 0}$ of λ-discounted equilibria of $G_\lambda(\tilde{r})$.

If x_0 is absorbing then it is easy to prove, as in Vrieze and Thuijsman (1989) or Solan (1999) that the conditions of Lemma 5.5 are satisfied.

If x_0 is non-absorbing, we have 8 cases, according to whether x_0 is an isolated action, semi-isolated profile, weak profile or neither, and according to whether $\alpha_0 = 1$ or $\alpha_0 < 1$. In each of these cases, we invoke one of the sufficient conditions stated before, as summarized by the following table:

	$\alpha_0 = 1$	$\alpha_0 < 1$
x_0 is isolated	Lemma 5.4 , Lemma 5.7	Lemma 4.5
x_0 is semi-isolated	Lemma 5.4 , Lemma 5.7	Lemma 5.14
x_0 is weak	Lemma 5.5	Lemma 5.16
Otherwise	Lemma 5.5	Lemma 5.15

The definition of the function $\tilde{r}$ was involved because it must ensure two different things: for $\alpha_0 = 1$ the existence of a uniform equilibrium, and for $\alpha_0 < 1$ the existence of a good convex combination that is supported by punishable players.

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

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

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

■

The following corollary follows easily from the definition of individual rationality and Lemmas 6.4 and 6.5.

Corollary 6.6 *x_0 is individually rational for $\tilde{\gamma}_0$.*

Lemma 6.7 *If x_0 is absorbing then $u(x_0)$ is an equilibrium payoff.*

Proof: We prove that x_0 satisfies the conditions of Lemma 5.5. By Corollary 6.6, x_0 is individually rational for $\tilde{\gamma}_0$.

Since x_0 is absorbing, $\tilde{\gamma}_0 = u(x_0)$. By Lemma 6.5, for every player $i \in N$,

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

hence $u^i(x^{-i}, a^i) = u^i(x_0)$ whenever $a^i \in \text{supp}(x_0^i)$ with $w(x_0^{-i}, a^i) > 0$, and the second condition holds. ■

If $\alpha_0 = 1$ then $\tilde{\gamma}_0 = \lim \tilde{\gamma}_\lambda(x_\lambda) = \tilde{r}(x_0)$. In this case, as shown below, G admits an equilibrium payoff.

LEMMA 6.8 *If $\alpha_0 = 1$ then G admits an equilibrium payoff. In particular no correlation amongst the players is needed.*

Proof: Since $\alpha_0 = 1$, $\tilde{\gamma}_0 = \tilde{r}(x_0)$ and x_0 is non-absorbing. By Corollary 6.6, x_0 is individually rational for $\tilde{\gamma}_0$. 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. Since the support of x_0 does not include either isolated actions nor semi-isolated profiles $\tilde{r}(x_0) = \min\{r(x_0), v\}$ by the definition of $\tilde{r}$. Now

$$v \leq \tilde{\gamma}_0 = \tilde{r}(x_0) = \min\{r(x_0), v\} \leq r(x_0).$$

By Lemmas 6.4 and 6.5 and Corollary 6.6, x_0 satisfies the conditions of Lemma 5.4, hence $r(x_0)$ is an equilibrium payoff.

Assume now that $x_0 = a$ is an isolated action. If $r^i(a) \geq g^i(a)$ for every player i then by Lemma 5.4 $r(a)$ is a uniform equilibrium payoff. Suppose then there is $i_1 \in N$ such that $r^{i_1}(a) < g^{i_1}(a)$.

By the definition of $\tilde{r}$ at isolated actions and by Lemma 6.5, $p^i(a) = \tilde{r}^i(a) = \tilde{\gamma}_0^i \geq g^i(a)$ for every $i \in N$. But this implies that for every $j \neq i$, $u^i(a^{-j}, b^j(a)) \geq g^i(a)$. In particular, for every player $i \in N$, $u^i(a^{-i_1}, b^{i_1}(a)) \geq g^i(a)$. It follows that the conditions of Lemma 5.7 hold w.r.t. x_0, i_1 and $b^{i_1}(a)$.

Assume now that x_0 is a semi-isolated profile with signaller i_0. By the definition of $\tilde{r}^i(x_0)$ and by Lemma 6.5 it follows that $p^i(x_0) = \tilde{r}^i(x_0) = \tilde{\gamma}_0^i \geq g^i(x_0)$ for every $i \in N$. In particular, for $i = i_0$, $r^{i_0}(x_0) \geq p^{i_0}(x_0) \geq g^{i_0}(x_0)$.

If for every $i \neq i_0$, $r^i(x_0) \geq g^i(x_0)$, then by Lemma 5.4 $r(x_0)$ is a uniform equilibrium payoff. Otherwise, there exists a player $i_1 \neq i_0$ with $r^{i_1}(x_0) < g^{i_1}(x_0)$. As above, the conditions of Lemma 5.7 are satisfied for x_0, i_1 and d^{i_1}, where $d^{i_1} \neq x^{i_1}$ maximizes $u^{i_1}(x^{-i_1}, d^{i_1})$. ■

Assume now that x_0 is non-absorbing, but $\alpha_0 < 1$. In particular, x_λ is absorbing for every λ sufficiently small. For every exit $a^L \in E(x_0)$ define

$$x_\lambda[a^L] = \prod_{i \in L} x^i_\lambda[a^i] \prod_{i \notin L} x^i_\lambda[x^i_0].$$

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

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

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 $E(x_0)$.

It is easy to verify that (Solan 1999, Lemma 6.6)

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

It follows that

$$\tilde{\gamma}_0 = \alpha_0 \tilde{r}(x_0) + (1 - \alpha_0) u(\mu_0). \tag{10}$$

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 $\tilde{\gamma}^i_0$.

Lemma 6.9 *If $a^i \in E(x_0)$ and $\mu_0[a^i] > 0$ then $u^i(x_0^{-i}, a^i) = \tilde{\gamma}^i_0$.*

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 6.5, $u^i(x_0^{-i}, a^i) \leq \gamma^i_0$. Since the function $\tilde{r}$ is **not** multi-linear this lemma is not an immediate consequence of Lemma 6.5 and (10).

Lemma 6.10 *If x_0 is non-absorbing and neither isolated nor semi-isolated, and $\alpha < 1$ then $u(\mu_0)$ is a correlated equilibrium payoff.*

Proof: If x_0 is neither isolated nor semi-isolated, it is sufficient to prove that the conditions of either Lemma 5.15 or Lemma 5.16 hold.

We first prove that $u(\mu)$ is individually rational for x_0. By Corollary 6.6, it is sufficient to show that $u^i(\mu_0) \geq \tilde{\gamma}_0^i$ for every $i \in N$.

Since x_0 is neither isolated nor semi-isolated than $\tilde{r}^i(x_0) \leq v^i \leq \tilde{\gamma}_0$, and in particular (10) implies that $u^i(\mu_0) \geq \tilde{\gamma}_0^i \geq v^i$, as desired.

By Lemma 6.9, if x_0 admits two signallers then the conditions of Lemma 5.15 hold, and otherwise the conditions of Lemma 5.16 hold. ■

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 6.11 *If $x_0 = a$ is an isolated action and $\alpha < 1$ then $u(\mu_0)$ is a correlated equilibrium payoff.*

Proof: We prove that the conditions of Lemma 5.13 hold. Define $\mu_i = \sum_{b^i \in E(x_0)} \mu_0[b^i]$ to be the overall weight of unilateral exits of player i under μ_0. Since $E(x_0)$ contains only unilateral exits, $\sum_{i \in N} \mu_i = 1$.

We have to prove the following: (i) If $\mu[b^i] > 0$ then $b^i = b^i[a]$ (hence $u(\mu) = \sum_{i \in N} \mu_i u(a^{-i}, b^i(a))$). (ii) If $\mu_i > 0$ then i is punishable at x_0, and (iii) $u^j(\mu_0) \geq g^j(a)$.

Since the game is generic, it follows by Lemma 6.9 that if $\mu[b_i] > 0$ then $b^i = b^i(a)$ and $\tilde{\gamma}_0^i = g^i(a)$. In particular, (i) holds.

By (10), for every player $i \in N$, $\lim_{\lambda \to 0} \tilde{\gamma}_\lambda^i(x_\lambda^{-i}, a^i)$ is a convex combination of $\tilde{r}^i(a) = p^i(a)$ and $u^i(a^{-j}, b^j(a)) \geq p^i(a)$, for $j \neq i$, and in particular is at least $p^i(a)$.

To prove (ii), assume that $\mu_i > 0$. Then

$$p^i(a) \leq \lim_{\lambda \to 0} \tilde{\gamma}_\lambda^i(x_\lambda^{-i}, a^i) \leq \lim_{\lambda \to 0} \tilde{\gamma}_\lambda^i(x_\lambda) = \tilde{\gamma}_0^i = g^i(a),$$

and i is punishable.

(iii) follows by (10) and since $\tilde{r}^i(a) = p^i(a) \leq u^i(\mu)$. ■

A similar argument establishes:

LEMMA 6.12 *If x_0 is a semi-isolated profile and $\alpha < 1$ then $u(\mu)$ is a correlated equilibrium payoff.*

Proof: We prove that the conditions of Lemma 5.14 hold.

Let i_0 be the unique signaller at x_0. The proof that the second condition holds is similar to the proof provided in Lemma 6.11 for $i \neq i_0$, and to the proof provided in Lemma 6.10 for $i = i_0$.

For $i = i_0$, $\tilde{r}^{i_0}(x_0) \leq v^{i_0} \leq \tilde{\gamma}_0^{i_0}$. Since $\alpha < 1$ and by Lemma 6.4, (10) implies that $u^{i_0}(\mu_0) \geq \tilde{\gamma}_0^{i_0}$. For $i \neq i_0$, $u^i(\mu_0) \geq p^i(x_0) = \tilde{r}^i(x_0)$ since μ_0 is supported by unilateral exits. (10) implies again that $u^{i_0}(\mu_0) \geq \tilde{\gamma}_0^i$. The second condition follows now from Corollary 6.6. The third condition follows from Lemma 6.9. ■

7 Proof of Lemma 6.4

This section is devoted to the proof of Lemma 6.4. From now on we fix a player $i \in N$ and $\epsilon > 0$.

We need to show that for every λ sufficiently close to 0, player i has a mixed action $x^i \in X^i$ such that $\tilde{\gamma}_\lambda^i(x_\lambda^{-i}, x^i) \geq v^i - \epsilon$. In the sequel we use the fact that x_λ converge to a limit x_0.

One way of proving the lemma would be to prove 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 $\lim v_\lambda^i(\tilde{r}) \geq v^i$. 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 $\lim v_\lambda^i(\min\{r, v\}) = v^i$. Since p^i is incomparable to v^i, we cannot invoke Solan's result to prove Lemma 6.4. However, we will use some of his arguments.

We begin by proving that several quantities are at least the min-max value of player i.

LEMMA 7.1 *For every isolated action a,* $\max\{p^i(a), g^i(a)\} \geq v^i$.

Proof: Consider the following profile of players $N \setminus \{i\}$:

1. Each player $k \in N \setminus \{i, j_i(a)\}$ plays a^k.
2. Player $j_i(a)$, the punisher of i at a, plays $(1-\epsilon)a^{j_i(a)} + \epsilon b^{j_i(a)}(a)$.

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

Lemma 7.2 *For every semi-isolated profile x with signaller i_0, if $i \neq i_0$ then* $\max\{p^i(x), g^i(x)\} \geq v^i$.

Proof: Consider the following profile for players $N \setminus \{i\}$:

1. Players $k \neq i, j_i(x)$ play x^k.
2. Player $j_i(x)$ plays $(1-\epsilon)a^{j_i(x)} + \epsilon b_i^{j_i(x)}(x)$, where $a^{j_i(x)}$ is the sole action in player $j_i(x)$'s support.

The best player i can do against this profile is $\max\{p^i(x), g^i(x)\}$, and the lemma follows. ■

A **correlated** profile of players $N \setminus \{i\}$ is a function $\sigma^{-i} : H \to \Delta(\times_{j \neq i} A^i)$.

Definition 7.3 *Let $x^{-i} \in X^{-i}$ be a mixed action combination and $\eta > 0$. A correlated profile σ^{-i} is (x^{-i}, η)-**perturbed** if for every finite history h, $\| \sigma^{-i}(h) - x^{-i} \|_\infty < \eta$. The profile x^{-i} is called the **base** of σ^{-i}.*

In words, a perturbed correlated profile is one that is close to a stationary profile. Denote by $\mathcal{S}_\eta(x^{-i})$ the class of all (x^{-i}, η)-perturbed correlated profiles of players $N \setminus \{i\}$.

Lemma 7.4 *Let B 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* $\mathrm{supp}(x) \subseteq B$.[4] *Denote by a^{-i_0} the unique action combination of $N \setminus \{i_0\}$ in B. Then*

$$\max_{x^{i_0} \in \mathrm{supp}(B^{i_0})} p^{i_0}(a^{-i_0}, x^{i_0}) \geq v^{i_0}.$$

Proof: Let $c = \max_{x^{i_0} \in \mathrm{supp}(B^{i_0})} p^{i_0}(a^{-i_0}, x^{i_0})$. First we show that the max-min value of player i_0 in G is at most c. This implies by Neyman (1988) that the min-max value of player i_0 in G, when the minimizers are allowed to use correlated profiles, is also at most c. We then show that the minimizers have a minimizing correlated profile where the amount of correlation is "small". Finally, we approximate this correlated profile with a non-correlated profile in G, and show that the best player i can do against this non-correlated profile is at most $c + \epsilon$. Since ϵ is arbitrary, we get $c \geq v^{i_0}$, as required.

Instead of studying the original game G, we study a modification $\bar{G}$ of G.

Consider the zero-sum absorbing game $\bar{G}$, where player i_0 is the maximizer and players $N \setminus \{i_0\}$ are the minimizers, that is defined as follows.

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

- The action set of player i_0 is $Y^{i_0} = \{y^{i_0} \mid w(a^{-i_0}, y^{i_0}) = 0\}$.
- The action set of each player $i \neq i_0$ is A^i.
- Transition $\bar{w}$ and non-absorbing payoff $\bar{r}$ are the same as in G.
- The absorbing payoff is $\bar{u}(a^{-i_0,i}, x^{i_0}, d^i) = u^{i_0}(a^{-i_0,i}, x^{i_0}, d^i)$ for every $i \neq i_0$ and $d^i \neq a^i$, and $\bar{u} = 2$ elsewhere.
- Each player $i \neq i_0$ is restricted to play strategies that give probability at least $1 - \eta$ to a^i at every period, where η is sufficiently small.
- Player i_0 is restricted to play pure strategies (that is, he cannot perform lotteries between mixed actions).

In particular, any profile (correlated or non-correlated) of players $N \setminus \{i_0\}$ in $\bar{G}$ is in $S_\eta(C^{-i_0})$.

The game $\bar{G}$ is the same as G, but we restricted the actions of i_0, and changed the non-absorbing payoff only at absorbing action combinations.

Extend $\bar{w}$, $\bar{r}$ and $\bar{u}$ to mixed actions of players $N \setminus \{i_0\}$ as was done for w, r and u in G. In particular, $\bar{w}$, $\bar{r}$ and $\bar{w}\bar{u}$ are multi-linear, and $\bar{u}$ is quasi-concave and quasi-convex. Similar argument to the one used in Lemma 6.1 show that for every fixed discount factor, the function $\bar{\gamma}_\lambda$ that assigns to each stationary profile in $\bar{G}$ the corresponding λ-discounted payoff is quasi-concave in x^{i_0} and quasi-convex in x^{-i_0}. In particular, for every discounted factor, the discounted min-max value and the discounted max-min value are equal, provided that players $N \setminus \{i_0\}$ can correlate their actions.

For every fixed $x^{i_0} \in Y^{i_0}$, players $N \setminus \{i_0\}$ have a profile in S_η that lowers player i_0's payoff to c: If $r^{i_0}(a^{-i_0}, x_0) \leq c$, players $N \setminus \{i_0\}$ play a^{-i_0}, and otherwise player $j_{i_0}(x)$ plays $(1-\eta)a^{j_{i_0}(x)} + \eta b_{i_0}^{j_{i_0}(x)}(x)$ and each other player $i \neq i_0, j_{i_0}(x)$ players a^i.

It follows that for every $\epsilon > 0$ and every fixed strategy of player i_0 in $\bar{G}$, players $N \setminus \{i_0\}$ have a reply that lowers i_0 expected average payoff to $c + \epsilon$ in any sufficiently long game.

Since $\bar{r}$ and $\bar{u}$ are semi-algebraic, it follows that the result of Neyman (1988) holds for this game. In particular, the (uniform) min-max and the max-min values of player i_0 in $\bar{G}$ exist. Moreover, they are equal provided that players $N \setminus \{i_0\}$ can use correlated profiles.

Thus, c is at least the max-min value of player i_0 in $\bar{G}$, and therefore there exists $n_0 \in \mathbf{N}$ and a correlated profile $\sigma^{-i_0} \in S_\eta(a^{-i_0})$ of players $N \setminus \{i_0\}$ such that for every strategy σ^i of player i,

$$\bar{\gamma}_n(\sigma^{-i_0}, \sigma^{i_0}) \geq c^i - \epsilon \qquad \forall n \geq n_0,$$

where $\bar{\gamma}^i_n$ is the expected average payoff of player i_0 in $\bar{G}$ during the first n stages.

Let E be the collection of all action combinations where at least two players in $N \setminus \{i_0\}$ play an action different then that prescribed by a^{-i_0}.

We first note that we can assume w.l.o.g. that the probability under σ^{-i_0} that an action combination in E is played is 0. Indeed, since action combinations in E are absorbing, and the corresponding absorbing payoff is 2, whereas both $\bar{r}$ and $\bar{u}$ are bounded above by 1, by redefining σ^{-i} to give probability 0 to action combinations in E, and normalizing the remaining probability distribution we get a new profile that satisfies our requirements.

We now define a *non-correlated* profile $\tilde{\sigma}^{-i_0}$ that approximates σ^{-i_0}. Recall that for every finite history h, $\sigma^{-i_0}(h)$ gives positive probability only to action combination a^{-i_0} and to the action combinations (a^{-i,i_0}, d^i) where only player i plays an action $d^i \neq a^i$.

For every finite history h, every player i and every $d^i \neq a^i$, define $\tilde{\sigma}^i(h)[d^i] = \sigma^{-i_0}(h)[a^{-i,i_0}, d^i]$; that is, player i plays d^i with the same probability that (a^{-i,i_0}, d^i) should have occurred according to σ^{-i_0}.

Clearly $\tilde{\sigma}^{-i} \in \mathcal{S}_\eta$. Moreover, the probability of action combinations in E under $\tilde{\sigma}^{-i_0}$ is small. Let $h \in H$ be a finite history. Then $\tilde{\sigma}^{-i_0}(h)[E]$ is at most 2^{N-1} times $\sum_{i \neq i_0} \sum_{d^i \neq a^i} \sigma^{-i_0}(h)[a^{-i,i_0}, d^i]$.

Since $\sum_{d^i \neq a^i} \sigma^{-i_0}(h)[a^{-i,i_0}, d^i] \leq \eta$, the probability of ever playing an action combination in E under $\tilde{\sigma}^{-i_0}$ is at most $\eta 2^{N-1}$.

It follows that by playing $\tilde{\sigma}^{-i_0}$ players $N \setminus \{i\}$ bound the payoff of i_0 in $\bar{G}$ from above by $c - 2\epsilon$, provided η is sufficiently small.

Since the probability of a ever playing an action combination in E under $\tilde{\sigma}^{-i_0}$ is low, and since $g^{i_0}(x) < v^{i_0}$, by playing $\tilde{\sigma}^{-i_0}$ in the original game G players $N \setminus \{i\}$ bound the payoff of i_0 from above by $c + \epsilon$. ■

Proof of Lemma 6.4: Fix a player $i \in N$. 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.

Recall that (x_λ) is a sequence of λ-discounted equilibria in $G_\lambda(\tilde{r})$ that converge to x_0.

Assume that there exists $a^i \in A^i$ such that $a = (x_0^{-i}, a^i)$ is an isolated profile. Clearly $\lim \tilde{\gamma}^i_\lambda(a^{-i}, b^i(a)) = g^i(a)$. By the definition of $\tilde{r}$, $\lim \tilde{\gamma}^i_\lambda(a) \geq p^i(a)$. It follows that player i can guarantee $\max\{p^i(a), g^i(a)\}$, which, by Lemma 7.1 is at least v^i. Thus $\tilde{\gamma}^i_0 \geq v^i$, as desired.

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 7.2, show that player i can guarantee v^i in $G_\lambda(\tilde{r})$.

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 $g^i(x) \geq v^i$ then $\tilde{\gamma}^i_0 \geq \lim \tilde{\gamma}^i_\lambda(x_\lambda^{-i}, b^i(x)) = g^i(x) \geq v^i$. Assume then that $g^i(x) < v^i$.

Let $Y^i = \{y^i \in X^i \mid w(x_0^{-i}, y^i) = 0\}$. By (10), for every $y^i \in Y^i$,

$$\tilde{\gamma}^i_0 = \lim_{\lambda\to 0} \tilde{\gamma}^i_\lambda(x_\lambda^{-i}, y^i) \geq \min\{\tilde{r}^i(x_0^{-i}, y^i), p^i(x_0^{-i}, y^i)\} = \min\{r^i(x_0^{-i}, y), p^i(x_0^{-i}, y^i)\}.$$

In particular, $\tilde{\gamma}^i_0 \geq \max_{y^i \in \mathrm{supp}(C^i)} \min\{r^i(x_0^{-i}, y), p^i(x_0^{-i}, y^i\}$. By Lemma 7.4 the latter is at least v^i.

Last, assume that there is no action $a^i \in A^i$ such that one of the first three case hold.

If there exists an action $a^i \in A^i$ such that $u^i(x_0^{-i}, a^i) \geq v^i$ then

$$\tilde{\gamma}^i_0 \geq \lim_{\lambda\to 0} \tilde{\gamma}^i_\lambda(x_\lambda^{-i}, a^i) = u^i(x_0^{-i}, a^i) \geq v^i$$

as desired. Otherwise, for each $j \neq i$ let $D^j = \mathrm{supp}(x_0^j)$. The functions $\tilde{r}$ and r are continuous over X. Moreover, $\tilde{r} = \min\{r, v\}$ on $Y = \times_{j\in N}\Delta(D^j)$. Hence, if η is sufficiently small, $\tilde{r}^j(x) \geq \min\{r^j(x), v^j\} - \epsilon$ for every $j \in N$ and every $x \in X$ such that $d(x, Y) \leq \eta$. It follows from Solan (1999, Eq. (30)) that for every λ sufficiently small there exists $x^i \in \mathrm{supp}(D^i)$ such that $\tilde{\gamma}^i_\lambda(x_\lambda^{-i}, x^i) \geq v^i - 2\epsilon$, as desired. ■

References

[1] Aumann R.J. (1974) Subjectivity and Correlation in Randomized Strategies, **J. Math. Econ.**, **1**, 67-96

[2] Aumann R.J. (1987) Correlated Equilibrium as an Expression of Bayesian Rationality, **Econometrica**, **55**, 1-18

[3] Aumann R.J. 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 D. P. and Vohra R. V. (1998) Calibrated Learning and Correlated Equilibrium, **Games Econ. Behavior**, **21**, 40-55

[6] Mertens J.F. and Neyman A. (1981) Stochastic Games, **Int. J. Game Th.**, **10**, 53-66

[7] Neyman A. (1988) *mimeo*

[8] Solan E. (1999) Three-Person Absorbing Games, **Math. Oper. Res.**, **24**(3)

[9] Solan E. (1997) Absorbing Team Games, **Games Econ. Behavior**, *to appear*

[10] Solan E. and Vieille N. (1998a), Correlated Equilibrium in Stochastic Games, D.P. # 1226, The Center for Mathematical Studies in Economics and Management Science, Northwestern University **and** Cahiers du CEREMADE #9848

[11] Solan E. and Vieille N. (1998b), Quitting Games, D.P. # 1227, The Center for Mathematical Studies in Economics and Management Science, Northwestern University **and** Cahiers du CEREMADE #9846

[12] Vieille N. (1997a) Equilibrium in 2-Person Stochastic Games I: A Reduction, CEREMADE D.P. #9745

[13] Vieille N. (1997b) Equilibrium in 2-Person Stochastic Games II: The Case of Recursive Games, CEREMADE D.P. #9747

[14] Vrieze O.J. and Thuijsman F. (1989) On Equilibria in Repeated Games with Absorbing States, **Int. J. Game Th.**, **18**, 293-310