On the MaxMin Value of Stochastic Games with Imperfect Monitoring

Dinah Rosenberg*, Eilon Solan† and Nicolas Vieille‡§

December 24, 2001

Abstract

We study zero-sum stochastic games in which players do not observe the actions of the opponent. Rather, they observe a stochastic signal that may depend on the state, and on the pair of actions chosen by the players. We assume each player observes the state and his own action.

We propose a candidate for the max-min value, which does not depend on the information structure of player 2. We prove that player 2 can defend the proposed max-min value, and that in absorbing games player 1 can guarantee it. Analogous results hold for the min-max value. This paper thereby unites several results due to Coulomb.

*Laboratoire d'Analyse Géométrie et Applications, Institut Galilée, Université Paris Nord, avenue Jean-Baptiste Clément, 93430 Villetaneuse, France. e-mail: dinah@zeus.math.univ-paris13.fr

†MEDS Department, Kellogg School of Management, Northwestern University, *and* the School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel. e-mail: eilons@post.tau.ac.il

‡Ecole Polytechnique and Département Finance et Economie, HEC, 1, rue de la Libération, 78 351 Jouy-en-Josas, France. e-mail: vieille@hec.fr

§We deeply thank J.-F. Mertens, A. Neyman, S. Sorin and E. Lehrer. The discussions we had, and the suggestions they provided, were extremely useful, and simplified the proof. We acknowledge the financial support of the Arc-en-Ciel/Keshet program for 2001/2002. The research of the second author was supported by the Israel Science Foundation (grant No. 03620191).

1 Introduction

The classical literature on repeated games and stochastic games considers models with perfect monitoring – past play is observed by the players.

In the last two decades models with imperfect monitoring were explored, and several applications of these models were studied (see, e.g., Radner (1981), Rubinstein and Yaari (1983)). Lehrer (1989, 1990, 1992a, 1992b) has characterized various notions of undiscounted equilibria in infinitely repeated games with imperfect monitoring.

In the present paper we are interested in two-player zero sum stochastic games with imperfect monitoring. These games are played as follows. At every stage, the game is in one of finitely many states. Each player chooses an action, independently of his opponent. The current state, together with the pair of actions, determine a daily payoff that player 2 pays player 1, a probability distribution according to which a new state is chosen, and a probability distribution over pairs of signals, one for each player. Each player is then informed of his private signal, and of the new state. However, no player is informed of his opponent's signal, and of the daily payoff.

Coulomb (1992, 1999, 2001) was the first to study stochastic games with imperfect monitoring. He studied the class of absorbing games, and proved that the max-min value (and similarly the min-max value) exists. That is, there is a quantity $v \in \mathbf{R}$ such that for every $\varepsilon > 0$, (i) player 1 has a strategy that guarantees that his expected average payoff in all sufficiently long games exceeds $v - \varepsilon$, whatever player 2 plays, and (ii) for every strategy of player 1, player 2 has a reply such that the expected average payoff in all sufficiently long games does not exceed $v + \varepsilon$. Coulomb also provides a formula for the calculation of the max-min value and the min-max value in absorbing games.

One of Coulomb's main findings is that the max-min value does not depend on the signaling structure of player 2. Similarly, the min-max value does not depend on the signaling structure of player 1.

In the present paper we propose a candidate for the max-min and the min-max values in general stochastic games. As in the case of absorbing games, the candidate for the max-min value is independent of the information structure of player 2, and the candidate for the min-max value is independent of the information structure of player 1. These values are limits of max-min and min-max values of certain auxiliary discounted games.

We prove that player 2 can defend the proposed max-min value, and that in absorbing games player 1 can guarantee it. The proof that player 1 can guarantee the max-min value in general stochastic games is substantially more difficult, and is dealt with in a companion paper.

The approach that we take is quite different from that of Coulomb. We first define an equivalence relation over mixed actions of player 2, that has similarities with the one used in Lehrer's and Coulomb's works. The definition takes into account the fact that we use discounted games, hence events that occur rarely (relative to the discount factor) do not affect the payoff. Using this equivalence relation we define a new daily payoff function. We then define an auxiliary discounted max-min value as a fixed point of a functional equation that is based on the auxiliary daily payoff function. Finally, we prove that the max-min value is the limit of these auxiliary discounted max-min values.

To prove the last claim we would like to use the method developed by Mertens and Neyman (1981) for perfect monitoring stochastic games. Unfortunately, the new payoff function is convex but not necessarily multi-linear, which is a crucial requirement in Mertens and Neyman's proof.

Nevertheless, the method of Mertens and Neyman goes through for player 2; that is, it shows that player 2 can defend the proposed max-min value.

The method of studying asymptotic properties of auxiliary discounted games by defining a new payoff function appears already in Solan (1999) and in Solan and Vohra (1999), in the study of equilibria in n-player absorbing games. The present work is the first time where this method is successfully applied to general stochastic games.

The paper is organized as follows. Section 2 contains the model and the statement of the main results. In Section 3 we introduce a number of tools, define the auxiliary discounted games, and study some of their basic properties. Section 4 contains a reminder of the analysis of Mertens and Neyman. In Section 5 we deal with player 2's side of the analysis in the general case and in Section 6 with player 1's side in the case of absorbing games. We end the paper by showing how our tools can be used to show that if player 1 receives no signals, he can guarantee the max-min value by *stationary* strategies, as was already shown by Coulomb (1992).

Thus, our paper offers a unified treatment of Coulomb (1992,1999,2001). It also lays down tools that are used in the analysis of the general case.

2 The model

For every finite set K, $\Delta(K)$ is the space of probability distributions over K. We identify each element $k \in K$ with the probability distribution in $\Delta(K)$ that gives weight 1 to k.

We consider the standard model of finite two-person zero-sum stochastic games with signals. Such a game is described by: (i) a finite set S of states, (ii) finite action sets A and B for the two players, (iii) a daily reward function $r : S \times A \times B \to \mathbf{R}$, (iv) finite sets M^1 and M^2 of signals for the two players and (v) a transition function $\psi : S \times A \times B \to \Delta(M^1 \times M^2 \times S)$.

The game is played in stages. The initial state s_1 is known to both players. At each stage $n \in \mathbf{N}$, (a) the players independently choose actions a_n and b_n; (b) player 1 gains $r(s_n, a_n, b_n)$, and player 2 looses the same amount; (c) a triple (m^1_n, m^2_n, s_{n+1}) is drawn according to $\psi(s_n, a_n, b_n)$; (d) players 1 and 2 are told respectively m^1_n and m^2_n, but they are *not* informed of a_n, b_n, or $r(s_n, a_n, b_n)$; and (e) the game proceeds to stage $n + 1$.

We assume throughout that each player always knows the current state, and the action he is playing. In terms of ψ, this amounts to assuming that the signal received by player i contains the identity of the new state and the name of the action he has just chosen: if $m^1 \in M^1$ has positive probability under both $\psi(s, a, b)$ and $\psi(s', a', b')$ then $(s, a) = (s', a')$. An analogous property holds with reversed roles for player 2. We also assume perfect recall, so each player remembers the sequence of signals he has received so far.

We assume w.l.o.g. that payoffs are non-negative and bounded by 1.

We denote by $H_n = S \times (A \times B \times M^1 \times M^2 \times S)^{n-1}$ the set of histories up to stage n,[1] by $H^1_n = S \times (M^1)^{n-1}$ and $H^2_n = S \times (M^2)^{n-1}$ the set of histories to players 1 and 2 respectively. We equip these spaces with the discrete topology. We also let $H_\infty = (S \times A \times B \times M^1 \times M^2)^{\mathbf{N}}$ denote the set of infinite plays, $\mathcal{H}^i_n$ denote the cylinder algebra over H_∞ induced by H^i_n, and we set $\mathcal{H}_\infty = \sigma(\mathcal{H}^1_n \cup \mathcal{H}^2_n, n \geq 1)$, the σ-algebra generated by all those cylinder algebras. We let $\mathbf{s}_n, \mathbf{a}_n$

[1] Since the signal of each player contains the current state and his action, some information in this representation is redundant.

and $\mathbf{b}_n$ denote respectively the state at stage n, and the action played at stage n: these are random variables, and are respectively $\mathcal{H}_n^i$, $\mathcal{H}_{n+1}^1$ and $\mathcal{H}_{n+1}^2$ measurable.

Whenever convenient, we use the convention that boldfaced letters denote random variables, and non boldfaced letters denote the value of the random variable.

A (behavioral) strategy of player 1 (resp. player 2) is a sequence $\sigma = (\sigma_n)_{n\geq 1}$ (resp. $\tau = (\tau_n)_{n\geq 1}$) of functions $\sigma_n : H_n^1 \to \Delta(A)$ (resp. $\tau_n : H_n^2 \to \Delta(B)$). Such a strategy is *stationary* if the mixed move used at stage n depends only on $\mathbf{s}_n$ (which is known to both players). Every stationary strategy of player 1 (resp. player 2) can be identified with a vector $x \in (\Delta(A))^S$ (resp. $y \in (\Delta(B))^S$), with the interpretation that $x(s)$ is the lottery used by player 1 whenever the play visits state s.

Given a pair (σ, τ) of strategies and an initial state s, we denote by $\mathbf{P}_{s,\sigma,\tau}$ the probability distribution induced over $(H_\infty, \mathcal{H}_\infty)$ by (σ, τ) and s, and by $\mathbf{E}_{s,\sigma,\tau}$ the corresponding expectation operator. The expected average payoff up to stage n is

$$\gamma_n(s, \sigma, \tau) = \mathbf{E}_{s,\sigma,\tau}\left[\frac{1}{n}\sum_{k=1}^{n} r(\mathbf{s}_k, \mathbf{a}_k, \mathbf{b}_k)\right].$$

Definition 1 *$v(s)$ is the* max-min value *of the game with initial state s if:*

- *Player 1 can* guarantee *$v(s)$: for every $\varepsilon > 0$, there exists a strategy σ of player 1 and $N \in \mathbf{N}$, such that:*
$$\forall \tau, \forall n \geq N,\ \gamma_n(s, \sigma, \tau) \geq v(s) - \varepsilon.$$
- *Player 2 can* defend *$v(s)$: for every $\varepsilon > 0$ and every strategy σ of player 1 there exists a strategy τ of player 2 and $N \in \mathbf{N}$, such that:*
$$\forall n \geq N, \gamma_n(s, \sigma, \tau) \leq v(s) + \varepsilon.$$

The definition of the min-max value is obtained by exchanging the roles of the two players.

Coulomb (1999, 2001) proved that the max-min value exists in every absorbing game. Moreover, Coulomb provides a formula for the max-min value, which is independent of the signaling structure of player 2. By exchanging the roles of the two players, Coulomb derives an analogous result for the min-max value.

3 Indistinguishable moves

3.1 Definition and continuity properties

We start by defining an equivalence relation between mixed actions of player 2; two mixed actions y and y' are equivalent if the probability that player 1 can not distinguish y from y' is high. Variants of this relation have played a central role in earlier analysis of games with imperfect monitoring, such as in the work of Aumann and Mashler (1995), Lehrer (1989,1990,1992) and Coulomb (1999,2001).

Given $\varepsilon, \lambda > 0$,[2] $s \in S$ and $x \in \Delta(A)$, we define a binary relation $\sim_{\lambda,\varepsilon,s,x}$ over $\Delta(B)$ as follows:

$$y \sim_{\lambda,\varepsilon,s,x} y' \text{ if and only if } \psi(s, a, y) = \psi(s, a, y') \text{ whenever } x[a] \geq \lambda/\varepsilon.$$

[2] λ always stands for a discount factor. Here and in the sequel we omit the condition $\lambda \leq 1$.

Thus, y and y' are equivalent for x at s if every action of player 1 that can be used to distinguish between y and y' is played under x with low probability. This notion plays a fundamental role in our work. We therefore study it in some detail.

First, we provide a geometric rewriting of the definition. For the time being, we let λ, ε, x and s be given, and denote by $\overline{A} = \left\{a \in A, x[a] \geq \frac{\lambda}{\varepsilon}\right\}$ the set of actions that are relevant for the relation. We also write $y \sim z$ for the more cumbersome $y \sim_{\lambda,\varepsilon,s,x} y'$. We denote by $\Psi : \mathbf{R}^B \to \mathbf{R}^{\overline{A}}$ the linear map $\Psi(Y) = (\psi(s,a,Y))_{a \in \overline{A}}$, where $\psi(s,a,Y) = \sum_{b \in B} Y(b)\psi(s,a,b)$, and by N the kernel of Ψ: $N = \{Y \in \mathbf{R}^B \mid \Psi(Y) = \vec{0}\}$. Plainly,

$$y \sim z \Leftrightarrow z - y \in N.$$

Thus, denoting $F(y) := \{z \in \Delta(B), y \sim z\}$, one has

$$F(y) = (y + N) \cap \Delta(B).$$

This formulation allows for a straightforward verification of some simple properties. In particular, $F(y)$ is compact. Also, for every $y_1, y_2 \in \Delta(B)$ and every $\alpha \in [0,1]$, by convexity of both N and $\Delta(K)$,

$$\begin{aligned} \alpha F(y_1) + (1-\alpha)F(y_2) &= \alpha((y_1 + N) \cap \Delta(K)) + (1-\alpha)((y_2 + N) \cap \Delta(K)) \\ &\subseteq F(\alpha y_1 + (1 - \alpha y_2)). \end{aligned} \tag{1}$$

Lemma 2 *$\sim$ is an equivalence relation.*

Proof. We prove only transitivity. Assume $y_1 \sim y_2$ and $y_2 \sim y_3$. This means $y_2 - y_1 \in N$ and $y_3 - y_2 \in N$. This implies $y_3 - y_1 \in N$ hence $y_1 \sim y_3$. ■

Remark: An alternative definition of the relation $\sim_{\lambda,\varepsilon,s,x}$ is the following: $y \sim_{\lambda,\varepsilon,s,x} y'$ if and only if $\|\psi(s,x,y) - \psi(s,x,y')\| < \lambda/\varepsilon$. However, in this definition the relation is not transitive, and we were unable to prove Lemma 4 below.

Let $k > 0$. Recall that a set-valued map F between two metric sets (E_1, d_1) and (E_2, d_2) is k-*Lipschitz* if $F(y) \subseteq B(F(y'), kd_1(y,y'))$ for each $y, y' \in E_1$ (Definition 4, p.44 in Aubin and Cellina (1984)).

By Aubin and Frankowska (1990, Theorem 2.2.6) it follows that the set-valued map F is k-Lipschitz, for some $k > 0$.

Note that the set-valued map F depends on λ, ε, s and x only through the set $\overline{A}$. Since A and S are finite, only finitely many set-valued maps arise as λ, ε, s and x vary. In particular, we may choose a constant $K > 0$ such that all these maps are K-Lipschitz.

We define a function $\widetilde{r}$ that is to be thought of as the worst payoff consistent with a given distribution of signals to player 1: given $\varepsilon, \lambda > 0$, $s \in S$ and $(x,y) \in \Delta(A) \times \Delta(B)$, we set

$$\widetilde{r}^{\varepsilon}_{\lambda}(s,x,y) = \inf_{z \sim_{\lambda,\varepsilon,s,x} y} r(s,x,z). \tag{2}$$

We now establish useful properties of $\widetilde{r}$.

Lemma 3 *For every $\varepsilon, \lambda > 0$, every $x \in \Delta(A)$ and every $s \in S$, the function $\widetilde{r}_\lambda^\varepsilon(s, x, \cdot) : \Delta(B) \to \mathrm{R}$ is convex and $(K+1)$-Lipshitz.*

Proof. We start with convexity. Let $y = \alpha y_1 + (1-\alpha)y_2$, where $y_1, y_2 \in \Delta(B)$ and $\alpha \in [0,1]$. We use the previous notations. By (1) and since r is multi-linear,

$$\begin{aligned}\widetilde{r}_\lambda^\varepsilon(s,x,y) &= \inf_{z \in F(y)} r(s,x,z) \leq \inf_{z \in \alpha F(y_1) + (1-\alpha)F(y_2)} r(s,x,z) \\ &= \alpha \inf_{z \in F(y_1)} r(s,x,z) + (1-\alpha) \inf_{z \in F(y_2)} r(s,x,z),\end{aligned}$$

Since r is non-negative and multi-linear, the function $y \mapsto r(s,x,y)$ is 1-Lipshitz. By Aubin and Cellina (1984, Theorem 1.7), $y \mapsto \inf_{z \in F(y)} r(s,x,z)$ is $(K+1)$-Lipshitz. ■

Denote by $\Delta(B)/N$ the quotient space of $\Delta(B)$ by the equivalence relation $\sim_{\lambda,\varepsilon,s,x}$, endowed with the quotient topology. For $y \in \Delta(B)$, we denote by $cl(y)$ its projection over $\Delta(B)/N$. Plainly,

$$d(cl(y), cl(z)) := \|\Psi(y) - \Psi(z)\| = \|\Psi(y-z)\|$$

defines a metric over $\Delta(B)/N$, whose topology coincides with the quotient topology. By definition, $\widetilde{r}_\lambda^\varepsilon(s,x,y)$ depends on y only through $cl(y)$, hence there is a map $\widehat{r} : \Delta(B)/N \to \mathbf{R}$, such that $\widetilde{r}_\lambda^\varepsilon(s,x,\cdot) = \widehat{r} \circ \pi$, where $\pi : \Delta(B) \to \Delta(B)/N$ is the projection map. Since $\widetilde{r}_\lambda^\varepsilon(s,x,\cdot)$ is continuous, so is $\widehat{r}$. Therefore, we have established the following.

Lemma 4 *For every $\delta > 0$, there is $\eta > 0$ such that for every $s \in S$, every $x \in \Delta(A)$, and every $y, z \in \Delta(B)$, the following is satisfied: if $\|\psi(s,a,y) - \psi(s,a,z)\| < \eta$ for every $a \in A$ that satisfies $x[a] \geq \lambda/\varepsilon$, then $|\widetilde{r}_\lambda^\varepsilon(s,x,y) - \widetilde{r}_\lambda^\varepsilon(s,x,z)| < \delta$.*

3.2 The candidate for the max-min value

We here posit a quantity which is our candidate for the max-min value of the game. We denote by $q : S \times A \times B \to \Delta(S)$ the transition function induced by ψ :

$$q(s'|s,a,b) = \psi(s,a,b)[\{s'\} \times M^1 \times M^2].$$

the multi-linear extension of q to $S \times \Delta(A) \times \Delta(B)$ is still denoted by q. For every $s \in S$, and every $(x,y) \in \Delta(A) \times \Delta(B)$, denote by $\mathbf{E}[\cdot|s,x,y]$ the expectation with respect to $q(\cdot|s,x,y)$.

Given $\lambda, \varepsilon > 0$ and $w : S \to \mathbf{R}$, define $T_{\lambda,\varepsilon} w : S \to \mathbf{R}$ by

$$T_{\lambda,\varepsilon} w(s) := \sup_{x \in \Delta(A)} \inf_{y \in \Delta(B)} \left\{ \lambda \widetilde{r}_\lambda^\varepsilon(s,x,y) + (1-\lambda)\mathbf{E}[w|s,x,y] \right\}.$$

Thus, $T_{\lambda,\varepsilon}$ is an operator that acts on real-valued functions defined over S.

Lemma 5 *For each $\lambda, \varepsilon > 0$, the operator $T_{\lambda,\varepsilon}$ has a unique fixed point, v_λ^ε.*

Proof. Plainly, $T_{\lambda,\varepsilon}$ is non-decreasing. Moreover, $T_{\lambda,\varepsilon}(w + c\mathbf{1}) = T_{\lambda,\varepsilon}w + (1-\lambda)c\mathbf{1}$, for each $c \in \mathbf{R}$ and $w : S \to \mathbf{R}$, where $\mathbf{1} : S \to \mathbf{R}$ is the function defined by $\mathbf{1}(s) = 1$ for every $s \in S$. By Blackwell's criterion, $T_{\lambda,\varepsilon}$ is strictly contracting, hence has a unique fixed point. ■

Our candidate for the max-min value when the initial state is s is

$$v(s) = \lim_{\varepsilon\to 0}\lim_{\lambda\to 0} v_\lambda^\varepsilon(s).$$

The existence of the limit, together with additional properties of interest, is established in the next section. Observe that $v_\lambda^\varepsilon(s)$ depends only on the structure of signals to player 1, hence so does $v(s)$.

The main result of this paper is the following.

Theorem 6 *Player 2 can defend v. In addition, if there is one non-absorbing state, player 1 can guarantee v.*

3.3 Algebraic properties

We collect in Proposition 8 below the semi-algebraic properties that are useful for our purposes.

Lemma 7 *For every state $s \in S$, the function $\phi_s : (\varepsilon, \lambda, x, y) \mapsto \widetilde{r}_\lambda^\varepsilon(s, x, y)$ is semi-algebraic.*

Proof. Fix $s \in S$. The set

$$E = \left\{(\varepsilon, \lambda, x, y, y', r) \in (0,1)^2 \times \Delta(A) \times (\Delta(B))^2 \times \mathbf{R} : y \sim_{\lambda,\varepsilon,s,x} y', r = r(s, x, y')\right\}$$

is defined by finitely many polynomial inequalities. In particular, it is semi-algebraic. Therefore the graph of ϕ_s, which is equal to

$$\left\{(\varepsilon, \lambda, x, y, r) \in (0,1)^2 \times \Delta(A) \times \Delta(B) \times \mathbf{R} : r = \inf\{r' \in \mathbf{R}, (\varepsilon, \lambda, x, y, y', r') \in E\}\right\}$$

is semi-algebraic as well. ■

Using Lemma 5 one can now deduce the following.

Proposition 8 *For every state $s \in S$, (i) the function $(\lambda, \varepsilon) \mapsto v_\lambda^\varepsilon(s)$ is semi-algebraic, and (ii) the set*

$$\left\{(\varepsilon, \lambda, x) \in (0,1)^2 \times \Delta(A),\ \inf_y \left\{\lambda \widetilde{r}_\lambda^\varepsilon(s, x, y) + (1-\lambda)\mathbf{E}\left[v_\lambda^\varepsilon | s, x, y\right]\right\} \geq v_\lambda^\varepsilon(s) - \varepsilon\lambda\right\}$$

is semi-algebraic.

In particular, for every fixed $\varepsilon > 0$, $\lim_{\lambda\to 0} v_\lambda^\varepsilon(s)$ exists. Since the function $\varepsilon \mapsto \lim_{\lambda\to 0} v_\lambda^\varepsilon(s)$ is monotonic, the limit $v(s) = \lim_{\varepsilon\to 0}\lim_{\lambda\to 0} v_\lambda^\varepsilon(s)$ exists.

For every state $s \in S$ define

$$G = \{(\lambda, \varepsilon, z) \in (0,1)^2 \times \mathbf{R}^S \mid \lambda \leq \varepsilon^2, z = v_\lambda^\varepsilon\}. \tag{3}$$

G_s is a semi-algebraic set, whose closure contains $(0,0,v)$. Indeed, for every $\eta > 0$ there is $\varepsilon_0 > 0$ sufficiently small such that $\| \lim_{\lambda \to 0} v^\varepsilon_\lambda - v\| < \eta$ for every $\varepsilon \in (0, \varepsilon_0)$. Hence for every $\varepsilon \in (0, \varepsilon_0)$ there is $\lambda_0(\varepsilon) \in (0,1)$ such that $\|v^\varepsilon_\lambda - v\| < 2\eta$ for every $\lambda \in (0, \lambda_0(\varepsilon))$.[3]

By the Curve Selection Theorem (see, e.g. Bochnak et al (1998, Theorem 2.5.5)) there is a continuous semi-algebraic function $f : (0,1) \to G$ such that $\lim_{r\to 0} f(r) = (0,0,v)$.

Write $f(r) = (\lambda(r), \varepsilon(r), v^{\varepsilon(r)}_{\lambda(r)})$. The functions $r \mapsto \lambda(r)$, $r \mapsto \varepsilon(r)$ and $r \mapsto v^{\varepsilon(r)}_{\lambda(r)}(s)$ (for $s \in S$) are semi-algebraic, hence monotone in a neighborhood of zero. Since $\lambda > 0$ for each $(\lambda, \varepsilon, v) \in G$, and since $\lim_{r\to 0} \lambda(r) = 0$, the function $\lambda(r)$ is invertible in a neighborhood of zero. Hence, there is a semi-algebraic function $\lambda \mapsto \varepsilon(\lambda)$ such that, in a neighborhood of 0, $(\lambda, \varepsilon(\lambda), v^{\varepsilon(\lambda)}_\lambda) \in G$ and $\lim_{\lambda\to 0} v^{\varepsilon(\lambda)}_\lambda = v$.

We denote by d the degree in λ of the function $\lambda \mapsto \varepsilon(\lambda)$. That is, $\lim_{\lambda\to 0} \lambda^d/\varepsilon(\lambda) \in (0, \infty)$. By the definition of G, $d \in (0, 1/2]$.

4 Reminder on zero-sum games

The purpose of this section is to provide a slight modification of a result due to Mertens and Neyman (henceforth MN) (1981) , that we will use. We let $\lambda \longmapsto w_\lambda$ be a $\mathbf{R}^S$-valued semi-algebraic function, and $w = \lim_{\lambda\to 0} w_\lambda$.

Let $\varepsilon > 0$, $Z \geq 0$ and two functions $\lambda : (0, +\infty) \to (0, 1)$ and $L : (0, +\infty) \to \mathbf{N}$ be given. Set $\delta = \varepsilon/48$. Assume that the following conditions are satisfied for every $z \geq Z$, every $|\eta| \leq 4$ and every $s \in S$:

C1 $|w_\lambda(s) - w(s)| \leq 4\delta$;

C2 $4L(z) \leq \delta z$;

C3 $|\lambda(z + \eta L(z)) - \lambda(z)| \leq \delta\lambda(z)$

C4 $\left|w_{\lambda(z+\eta L(z))}(s) - w_{\lambda(z)}(s)\right| \leq 4\delta L(z)\lambda(z)$

C5 $\int_Z^\infty \lambda(z)dz \leq 4\delta$.

MN (1981) note that **C1-C5** hold for Z large enough, in each of the next two cases:

Case 1 $\lambda(z) = z^{-\beta}$ and $L(z) = \lceil \lambda(z)^{-\alpha} \rceil$,[4] where $\alpha \in (0,1)$ and $\beta > 1$ satisfy $\alpha\beta < 1$;

Case 2 $L(z) = 1$ and $\lambda(z) = 1/z(\ln z)^2$.

Let $(\widehat{r}_k)_{k\in\mathbf{N}}$ be a $[0,1]$-valued process defined on the set of plays. Define recursively processes

[3]The condition $\lambda \leq \varepsilon^2$ in (3) can be replaced by $\lambda \leq \varepsilon^c$ for any $c > 1$.

[4]For every $c \in \mathbf{R}$, $\lceil c \rceil$ is the minimal integer greater than or equal to c.

$(z_k),(L_k)$ and (B_k) by the formulas

$$z_0 = Z, B_0 = 1,$$
$$\lambda_k = \lambda(z_k), L_k = L(z_k), B_{k+1} = B_k + L_k,$$
$$z_{k+1} = \max\left\{Z, z_k + \lambda_k\left(L_k\widehat{r}_k - \sum_{B_k\le n<B_{k+1}} w_{\lambda_k}(s_n)\right) + \frac{\varepsilon}{2}\right\}.$$

Let (I_k) be a $\{0,1\}$-valued process, where I_k is $\mathcal{H}^1_{B_{k-1}}$-measurable and $I_k = 1$ for at most one value of k. This process does not appear in MN (1981) formulation.

Theorem 9 *Let (σ,τ) be a strategy pair. Assume that for every $k \ge 0$,*

$$\mathbf{E}_{s,\sigma,\tau}\left[\lambda_k L_k \widehat{r}_k + \varepsilon I_k + (1-\lambda_k L_k) w_{\lambda_k}(\mathbf{s}_{B_{k+1}}) | \mathcal{H}^1_{B_k}\right] \quad (4)$$
$$\ge w_{\lambda_k}(\mathbf{s}_{B_k}) - \frac{\varepsilon}{12}\lambda_k L_k. \quad (5)$$

Then there exists $N_0 \in \mathbf{N}$, independent of (σ,τ), such that the following holds for every $n \ge N_0$:

$$\mathbf{E}_{s,\sigma,\tau}\left[\frac{1}{n}\sum_{p=1}^{n} \widehat{R}_n\right] \ge w(s) - 2\varepsilon, \quad (6)$$

where $\widehat{R}_n = \widehat{r}_k$ whenever $B_k \le n < B_{k+1}$. Moreover,

$$\mathbf{E}_{s,\sigma,\tau}\left[\sum_{k=1}^{\infty} \lambda_k L_k\right] < +\infty. \quad (7)$$

The result also holds when replacing in (5) and (6) $\ge$ by $\le$, and the '+' sign on the right-hand side by a '-' sign.

Proof. This statement differs from the statement in MN (1981) through the additional process I_k.

To handle the term εI_k on the left-hand side, it is enough to introduce the following changes in Mertens and Neyman (1981). First, add εI_k in Lemmas 3.4 and 3.5. Second, define Y_k as $Y_k = l_k - t_k + \varepsilon \sum_{p<k} I_p$, and rewrite the second consequence of Proposition 3.6 as $E(\bar{l}_i) \ge \bar{l}_0 - 12\delta - \varepsilon E\left[\sum_{p<k(i)} I_k\right] \ge \bar{l}_0 - 12\delta - \varepsilon$. Finally, rewrite the subsequent computation taking into account the additional term $-\varepsilon$. ■

In MN's analysis, $\widehat{r}_k$ is the average payoff received by player 1 during stages $B_k, B_{k-1}+1, \ldots, B_{k+1}-1$. Here, this variable is not observable, and $\widehat{r}_k$ will be defined as the worst payoff to player 1, consistent with his observations on block k. Since signals depend randomly on actions, many stages are needed to provide a reliable estimator of the moves selected by player 2. This is why strategies (for player 1's side) will be defined in blocks. In the case of absorbing games (see later for a definition), early absorption on a block may prevent such a computation. In fact, for all blocks during which the probability of reaching an absorbing state is non negligible, it is very difficult to relate $\widehat{r}_k$ to the moves used by player 2. Therefore, we introduce a *flag* I_k, that will be interpreted as "absorption occurs during block k". The condition on I_k arises from the fact that absorption may occur at most once during the play.

5 Player 2 can defend v

We prove in this section that player 2 can defend v. Since the value of v_λ^ε, hence of v, is independent of the signaling structure for player 2, we need to prove the result for the least favorable situation for player 2, i.e., the case where player 2 is only told the current state.

Let $s \in S$ and $\varepsilon > 0$ be given. We shall prove that player 2 can defend $\lim_{\lambda \to 0} v_\lambda^\varepsilon(s)$. Let a strategy σ of player 1 be given. Given a strategy τ, for every $n \in \mathbf{N}$ and every $\lambda \in (0,1)$, we let $\mathbf{y}_n^\lambda \in \Delta(S)$ satisfy

$$\lambda \widetilde{r}_\lambda^\varepsilon(\mathbf{s}_n, \xi_n, \mathbf{y}_n^\lambda) + (1-\lambda)\mathbf{E}\left[v_\lambda^\varepsilon | \mathbf{s}_n, \xi_n, \mathbf{y}_n^\lambda\right] \le v_\lambda^\varepsilon(\mathbf{s}_n) + \frac{\varepsilon}{12}\lambda,$$

where $\xi_n = \mathbf{E}_{s,\sigma,\tau}\left[\mathbf{a}_n | \mathbf{s}_1, .., \mathbf{s}_n\right]$ is the expected mixed action of player 1 given the information of player 2 at stage n. We also choose $\overline{\mathbf{y}}_n^\lambda \sim_{\lambda,\varepsilon,\mathbf{s}_n,\xi_n} \mathbf{y}_n^\lambda$ such that

$$r(\mathbf{s}_n, \xi_n, \overline{\mathbf{y}}_n^\lambda) = \widetilde{r}_\lambda^\varepsilon(\mathbf{s}_n, \xi_n, \mathbf{y}_n^\lambda).$$

Since player 2 knows σ, he can calculate ξ_n, hence also $\mathbf{y}_n^\lambda$ and $\overline{\mathbf{y}}_n^\lambda$.

We now apply the result stated in Section 4, **Case 2** ($L(z) = 1$, $\lambda(z) = 1/(z(\ln(z))^2)$), $\widehat{r}_n = \widetilde{r}_{\lambda_n}^\varepsilon(\mathbf{s}_n, \xi_n, \mathbf{y}_n^{\lambda_n})$ and $I_k = 0$ for every $k \ge 0$.

We let τ be the strategy of player 2 that plays $\mathbf{y}_n^{\lambda_n}$ in stage n, where ξ_n is computed using τ. The computation of ξ_n involves only the restriction of τ to the first $n-1$ stages, hence there is no circularity in this definition. Both conclusions (6) (with inequalities reversed) and (7) of Theorem 9 are satisfied w.r.t. (σ, τ).

By (7), there is $N \in \mathbf{N}$ such that

$$\mathbf{E}_{s,\sigma,\tau}\left[\sum_{n=N}^{\infty} \lambda_n\right] < \frac{\varepsilon^2}{|A|}. \tag{8}$$

We define a strategy $\overline{\tau}$ for player 2 as: play $\mathbf{y}_n^{\lambda_n}$ in each stage $n < N$, and $\overline{\mathbf{y}}_n^{\lambda_n}$ in each stage $n \ge N$. We prove that $\overline{\tau}$ defends $\lim_\lambda v_\lambda^\varepsilon(s) + 2\varepsilon$. The definition of $\overline{\tau}$ is reminiscent of the type of replies defined by Coulomb (2001). Loosely speaking, under $\overline{\tau}$, player 2 plays for good transitions up to stage N, and for low payoffs afterwards.

By definition,

$$\mathbf{P}_{s,\sigma,\tau}\left(\psi^1(\mathbf{s}_n, \mathbf{a}_n, \mathbf{y}_n^{\lambda_n}) \ne \psi^1(\mathbf{s}_n, \mathbf{a}_n, \overline{\mathbf{y}}_n^{\lambda_n}) | \mathcal{H}_n^1\right) \le |A|\frac{\lambda_n}{\varepsilon}, \text{ for each } n \ge N.$$

Using (8), this implies that

$$\mathbf{P}_{s,\sigma,\tau}\left(\overline{H}_\infty\right) < \varepsilon, \tag{9}$$

where $\overline{H}_\infty$ is the $\mathcal{H}_\infty^1$-measurable set

$$\overline{H}_\infty = \left\{\psi^1(\mathbf{s}_n, \mathbf{a}_n, \mathbf{y}_n^{\lambda_n}) \ne \psi^1(\mathbf{s}_n, \mathbf{a}_n, \overline{\mathbf{y}}_n^{\lambda_n}) \text{ for some } n \ge N\right\}.$$

Therefore, $\sup_{F \in \mathcal{H}_\infty^1} |\mathbf{P}_{s,\sigma,\overline{\tau}}(F) - \mathbf{P}_{s,\sigma,\tau}(F)| < \varepsilon$. Since ξ_n is $\mathcal{H}_n^1$-measurable, one obtains

$$\left|\mathbf{E}_{s,\sigma,\tau}\left[\widetilde{r}_{\lambda_n}^\varepsilon(\mathbf{s}_n, \xi_n, \mathbf{y}_n^{\lambda_n})\right] - \mathbf{E}_{s,\sigma,\overline{\tau}}\left[\widetilde{r}_{\lambda_n}^\varepsilon(\mathbf{s}_n, \xi_n, \mathbf{y}_n^{\lambda_n})\right]\right| \le \varepsilon \text{ for each } n \in \mathbf{N}.$$

By the choice of $\overline{\mathbf{y}}_n^{\lambda_n}$, $\mathbf{E}_{s,\sigma,\overline{\tau}}\left[\widetilde{r}^{\varepsilon}_{\lambda_n}(\mathbf{s}_n,\xi_n,\mathbf{y}_n^{\lambda_n})\right] = \mathbf{E}_{s,\sigma,\overline{\tau}}\left[r(\mathbf{s}_n,\mathbf{a}_n,\mathbf{b}_n)\right]$, for every $n \geq N$. By summation, one obtains for every $n \geq N/\varepsilon$,

$$\gamma_n(s,\sigma,\overline{\tau}) = \mathbf{E}_{s,\sigma,\overline{\tau}}\left[\frac{1}{n}\sum_{k=1}^{n} r(\mathbf{s}_n,\mathbf{a}_n,\mathbf{b}_n)\right] \leq \mathbf{E}_{s,\sigma,\tau}\left[\frac{1}{n}\sum_{k=1}^{n}\widetilde{r}^{\varepsilon}_{\lambda_n}(\mathbf{s}_n,\xi_n,\mathbf{y}_n^{\lambda_n})\right] + \varepsilon.$$

By (6), this yields

$$\gamma_n(s,\sigma,\overline{\tau}) \leq \lim_{\lambda\to 0} v^{\varepsilon}_{\lambda}(s) + 2\varepsilon,$$

for every n sufficiently large, as desired.

6 Absorbing games: player 1 can guarantee v

In this section we restrict ourselves to absorbing games. An *absorbing state* is a state $s \in S$ such that $q(s|s,a,b) = 1$, for each $(a,b) \in A \times B$. Since the current state $\mathbf{s}_n$ is known to both players, and since in repeated games with imperfect monitoring the max-min value coincides with the value of the game with perfect monitoring, we may assume that for every absorbing state $s \in S$, $r(s) := r(s,a,b)$ is independent of $(a,b) \in A \times B$.

A game is *absorbing* if all states but one are absorbing. Clearly, the unique interesting starting state is the non absorbing one. From now on, we assume that the unique non absorbing state is the initial state, and we shall make no reference to it, unless when useful. We denote by θ the stage of absorption. If absorption never occurs, $\theta = +\infty$. Thus, $\theta + 1$ is the first stage in which the play is not in the non absorbing state.

We let $\omega > 0$ satisfy $\omega < \psi^1(m^1 \mid a,b)$, whenever the right hand side is strictly positive.

Throughout the section, $\varepsilon \in (0, 1/300|B|)$ is given. We let $\eta < 1/25$ be given by Lemma 4 w.r.t. ε, and we fix $\delta < \min\{\eta^2,\varepsilon^2\}$.

6.1 The strategy

Recall from Section 3.3 that there is a semi-algebraic function $\varepsilon : (0,1) \to (0,1)$ such that $\lambda \leq (\varepsilon(\lambda))^2$ for every $\lambda \in (0,1)$, $\lim_{\lambda\to 0}\varepsilon(\lambda) = 0$ and $\lim_{\lambda\to 0} v_{\lambda}^{\varepsilon(\lambda)} = v$. We there defined $d \in (0,\frac{1}{2}]$ to be the degree in λ of $\varepsilon(\lambda)$. For notational simplicity, we write v_λ and $\widetilde{r}_\lambda$ instead of the more cumbersome $v_{\lambda}^{\varepsilon(\lambda)}$ and $\widetilde{r}_{\lambda}^{\varepsilon(\lambda)}$. For $\lambda \in (0,1)$, we let $x_\lambda \in \Delta(A)$ be a mixed move of player 1 that satisfies:

$$\lambda\widetilde{r}_\lambda(x_\lambda,y) + (1-\lambda)\mathbf{E}\left[v_\lambda(\cdot)|x_\lambda,y\right] \geq v_\lambda - \varepsilon\lambda,\ \ \forall y \in \Delta(B). \tag{10}$$

We apply the result of Section 4, with **Case 1** ($\lambda(z) = z^{-\beta}$, $L(z) = \lceil\lambda(z)^{-\alpha}\rceil$), where $\alpha \in (1-d,1)$ is close enough to one so that $\|v_\lambda - v\| < \lambda^{1-\alpha}$ in a neighborhood of zero. We will let σ be the strategy that, for every $k \in \mathbf{N}$, plays x_{λ_k}on the k-th block, i.e. in stage $B_k, ..., B_{k+1}-1$. We set $I_k = 1$ if absorption occurs during the kth block, and 0 otherwise. To complete the definition of σ, it remains to define the variable $\widehat{r}_k$ that is used to update z_k, and the parameter Z.

If absorption occurred prior to stage B_k; that is, before the current block started, then $\widehat{r}_k$ is equal to the payoff at the absorbing state.

Otherwise, the value of $\widehat{r}_k$ depends only on the sequence of signals received during block k. For notational simplicity, we drop the subscript k; we thus write L instead of L_k, λ instead of λ_k, etc. We also relabel the stages of block k from 1 to L, so that $B_{k+1} = L + 1$.

We let $\overline{A} = \{a \in A, x_\lambda(a) \geq \lambda/\varepsilon(\lambda)\}$. For $a \in \overline{A}$, we let $\rho_a \in \Delta(M^1)$ stand for the empirical distribution of signals received by player 1 in the stages where a was played:

$$\rho_a[m] = \frac{\left|\{n \leq L, \mathbf{m}^1_n = m\}\right|}{|\{n \leq L, \mathbf{a}_n = a\}|}$$

We let $\rho \in (\Delta(M^1))^{\overline{A}}$ be the vector with coordinates $(\rho_a)_{a\in\overline{A}}$. Following the notations used in Section 3.1, we set $\Psi(y) = (\psi(a,y))_{a\in\overline{A}}$, for $y \in \Delta(B)$.

We let $\widehat{\mathbf{y}}$ minimize

$$\|\rho - \Psi(y)\|_\infty \tag{11}$$

among $y \in \Delta(B)$, and define

$$\widehat{r} = \begin{cases} 0 & \theta \leq L+1 \\ \widetilde{r}_\lambda(x_\lambda, \widehat{\mathbf{y}}) & \theta > L+1 \end{cases}$$

Thus, given a sequence of signals, player 1 computes a stationary strategy that is most consistent with the sequence of signals, and $\widehat{r}$ is the corresponding worst payoff.

The parameter Z that appears in the definition of σ will be fixed in Section 6.2.

6.2 A large number estimate

Here and later, it is convenient, and not restrictive, to assume that the play proceeds up to stage $L+1$, with players choosing actions, even if $\theta < L+1$. We denote by $\overline{\mathbf{y}} = \frac{1}{L}\sum_{n=1}^L \mathbf{b}_n \in \Delta(B)$ the empirical distribution of player 2's moves.

We prove that, with high probability, the empirical distribution of signals to player 1 is close to the (theoretical) distribution, given the sequence of moves actually played by player 2.

Lemma 10 *For every $a \in \overline{A}$ and every strategy τ, one has*

$$\mathbf{P}_{x_\lambda,\tau}(\left\|\psi^1(a,\overline{\mathbf{y}}) - \rho_a\right\| \geq \eta) \leq \varepsilon,$$

provided Z is sufficiently large.

Proof. Let $0 < \widehat{\eta} < \eta$ satisfy $\eta = 2\widehat{\eta}/(1-\widehat{\eta})$.

For $b \in B$, we denote by $N_b = |\{n \leq L, \mathbf{b}_n = b\}|$ the number of times the action b is played by player 2. For $a \in A$ and $m^1 \in M^1$, the meaning of N_a and N_{m^1} is similar. Also, $N_{b,m^1} = \left|\{n \leq L, \mathbf{b}_n = b, \mathbf{m}^1_n = m^1\}\right|$.

Fix $b \in B$, and let $a \in \overline{A}$ and $m^1 \in M^1$ satisfy $\psi^1(m^1|a,b) > 0$.

In every stage $n \leq L$ such that $\mathbf{b}_n = b$, the r.v. X_n that is equal to 1 if $\mathbf{m}^1_n = m^1$ and 0 otherwise is a Bernouilli r.v. with parameter $p = \psi^1(m^1 \mid x_\lambda, b) = x_\lambda(a)\psi^1(m^1|a,b) \geq \omega x_\lambda(a)$, where the second equality holds since the signal m^1 contains the action a. Moreover, the r.v.s $(X_n)_{n\leq L, \mathbf{b}_n=b}$ are *iid*. Choose $\xi \in (0, 1-\frac{1-d}{\alpha})$. Provided L is sufficiently large, $Lx_\lambda(a) > L^\xi$ for every $a \in \overline{A}$. For every $\eta \in (0,1)$ set $c_\eta = \min\{\eta^2/2, -\ln((1+\eta)^{1+\eta}\exp(\eta))\}$. By Alon et al (2001, Corollary A.14)

$$\mathbf{P}_{x_\lambda,\tau}\left(\left|\frac{N_{m^1,b}}{N_b} - \psi^1(m^1|x_\lambda,b)\right| > \frac{\widehat{\eta}}{|B|}\psi^1(m^1 \mid x_\lambda, b) \text{ and } N_b \geq L^\xi\right)$$
$$\leq \sum_{n=L^\xi}^{L} \exp(-c_{\frac{\widehat{\eta}}{|B|}} n x_\lambda(a)) \leq \frac{\exp(-c_{\frac{\widehat{\eta}}{|B|}} L^\xi x_\lambda(a))}{1-\exp(-c_{\frac{\widehat{\eta}}{|B|}} x_\lambda(a))} \leq \frac{\varepsilon}{2|B|},$$

provided Z is chosen large enough (recall that $x_{\lambda(z)}(a) \geq c(\lambda(z))^{1-d}$ for some constant c and every z large enough).

Note that, if $N_b < L^\xi$, the inequality $\left|N_{m^1,b} - N_b\psi^1(m^1|x_\lambda, b)\right| \leq \frac{\eta}{|B|}Lx_\lambda(a)$ holds trivially for Z large enough.

Since $N_b\psi^1(m^1 \mid x_\lambda, b) = x_\lambda(a)N_b\psi^1(m^1 \mid a, b) \leq Lx_\lambda(a)$, we have with probability at least $1 - \varepsilon/2B$,

$$\left|N_{m^1,b} - N_b\psi^1(m^1|x_\lambda, b)\right| \leq \frac{\widehat{\eta}}{|B|}Lx_\lambda(a). \tag{12}$$

By summing (12) over $b \in B$, this yields

$$\left|N_{m^1} - Lx_\lambda(a)\psi^1(m^1|a, \overline{\mathbf{y}})\right| \leq \widehat{\eta}Lx_\lambda(a),$$

with probability at least $1 - \varepsilon/2$.

Similarly, provided Z is sufficiently large, for every $a \in \overline{A}$, with probability at least $1 - \varepsilon/2$, $|N_a/L - x_\lambda(a)| \leq \widehat{\eta}x_\lambda(a)$, which implies that $Lx_\lambda(a)/N_a < 1/(1-\widehat{\eta})$. Hence, with probability at least $1 - \varepsilon$, one has

$$\left|\frac{N_{m^1}}{N_a} - \psi^1(m^1|a, \overline{\mathbf{y}})\right| \leq \frac{2\widehat{\eta}}{1-\widehat{\eta}} = \eta,$$

and the result follows. ■

We are now able to list the properties that Z should satisfy. We choose Z large enough so that:

D1 **C1-C5** of Section 4 are satisfied w.r.t. $w_\lambda = v_\lambda$.

D2 $2\left|A \times B \times M^1\right|^2 (\lambda(z))^{\alpha-(1-d)} < \eta\varepsilon$, for each $z \geq Z$.

D3 The conclusion of Lemma 10 holds.

D4 $\lambda(z)L(z) \leq \delta\varepsilon$ for each $z \geq Z$.

6.3 Applying the technique of Mertens and Neyman

The rest of this section is devoted to the proof of the next Proposition.

Proposition 11 *Provided Z satisfies **D1-D4**, there exists $N_0 \in \mathbf{N}$ such that for every $n \geq N_0$ and every strategy τ,*

$$\gamma_n(\sigma, \tau) \geq v - (12 \times 2 \times 18 + 3)\varepsilon.$$

The proof of the Proposition goes as follows. Consider a single block k. All the computations are conditional on $\mathcal{H}^1_{B_k}$. For notational simplicity, we omit any reference to k and relabel the stages of that block from 1 to $L := L(z_k)$. We let $I = \mathbf{1}_{\theta \leq L+1}$ and $w(\tau) = \mathbf{P}_{x_\lambda,\tau}(\theta \leq L+1) = \mathbf{E}_{x_\lambda,\tau}[I]$ denote the probability that absorption occurs within the block.

The proof follows from two lemmas.

Lemma 12 *For every strategy τ of player 2, one has*

$$\mathbf{E}_{x_\lambda,\tau}\left[\lambda L\widehat{\mathbf{r}} + \varepsilon I + (1-\lambda L)v_\lambda(\mathbf{s}_{L+1})\right] \geq v_\lambda - 18\varepsilon\lambda L. \tag{13}$$

Lemma 12 enables us to invoke Theorem 9. It is proved in the next section. To complete the proof of Proposition 11, it is sufficient to prove the following.

Lemma 13 *There exists N_0 such that, for every $n \geq N_0$ and every τ,*

$$\mathbf{E}_{\sigma,\tau}\left[\frac{1}{N_n}\sum_{p=1}^{N_n}\widehat{R}_p\right] \leq \gamma_n(\sigma,\tau) + 3\varepsilon,$$

where $N_n = \max\{k \in \mathbf{N}, B_k - 1 \leq n\}$ is the number of blocks that have been completed up to stage n.

Proof. The proof is standard, hence details omitted. We first compare $\mathbf{E}_{\sigma,\tau}[\widehat{r}_p]$ to the average payoff in block p. For notational ease, we relabel the stages of the pth block from 1 to $L := L_p$ (and condition on the history up to stage B_p). In the sequel, probabilities and expectations are with respect to (σ,τ). By Lemma 10, $\mathbf{P}(|\widehat{\mathbf{r}} - \widetilde{r}_\lambda(x_\lambda, \overline{\mathbf{y}}_L)| \geq \varepsilon \text{ and } \theta > B) \leq \varepsilon$. Since $\widehat{\mathbf{r}} = 0$ if $\theta \leq B$, this implies

$$\mathbf{E}[\widehat{\mathbf{r}}] \leq \mathbf{E}[\widetilde{r}_\lambda(x_\lambda, \overline{\mathbf{y}}_L)] \leq \mathbf{E}[r(x_\lambda, \overline{\mathbf{y}}_L)].$$

Next, we show that $\mathbf{E}[r(x_\lambda, \overline{\mathbf{y}}_L)]$ differs from $\mathbf{E}\left[\frac{1}{L}\sum_{n=1}^{L} r(\mathbf{a}_n, \mathbf{b}_n)\right]$ by at most $\frac{\varepsilon}{2}$. Indeed, the contribution of stages $n > \theta$ to both summations is the same. By the choice of Z (see the proof of Lemma 10), with probability at least $1-\varepsilon$, either (i) $\theta \leq \varepsilon L$, or for every $b \in B$, (ii) b is played less than $\varepsilon L/|B|$ times, or (iii) $\frac{1}{L}\sum_{n<\theta, \mathbf{b}_n = b} r(\mathbf{a}_n, b)$ and $\frac{N_b}{L} r(x_\lambda, b)$ differ by at most $\varepsilon/|B|$. In each case, $r(x_\lambda, \overline{\mathbf{y}}_L)$ differs from $\frac{1}{L}\sum_{n=1}^{L} r(\mathbf{a}_n, \mathbf{b}_n)$ by at most ε. Details are standard.

Therefore, for every n,

$$\mathbf{E}\left[\frac{1}{N_n}\sum_{p=1}^{N_n}\widehat{R}_p\right] \leq \mathbf{E}\left[\frac{1}{L_1 + ... + L_{N_n}}\sum_{p=1}^{L_1+...+L_{N_n}} r(\mathbf{s}_p, \mathbf{a}_p, \mathbf{b}_p)\right] + 2\varepsilon$$

By the definition of the strategy of Mertens and Neyman it follows that (i) $\lim_{k\to\infty} L_k/(L_1 + L_2 + \cdots + L_{k-1}) = 0$, uniformly in τ, (ii) for each k, L_k is bounded. It follows that $\mathbf{E}\left[\frac{1}{L_1+...+L_{N_n}}\sum_{p=1}^{L_1+...+L_{N_n}} r(\mathbf{s}_p, \mathbf{a}_p, \mathbf{b}_p)\right]$ differs from $\gamma_n(\sigma,\tau)$ by at most ε, provided n is large enough. ■

6.4 Proof of Lemma 12

The proof is different when $w(\tau) \leq \delta$, and when $w(\tau) > \delta$.

Lemma 14 *If $w(\tau) > \delta$ then (13) holds.*

Proof. Since payoffs are between 0 and 1, and by (10),

$$\begin{aligned}\mathbf{E}_{x_\lambda,\tau}[\lambda L\widehat{\mathbf{r}} + (1-\lambda L)v_\lambda(\mathbf{s}_{L+1})] &\geq \mathbf{E}_{x_\lambda,\tau}[v_\lambda(\mathbf{s}_{L+1})] - \lambda L \\ &\geq v_\lambda - \varepsilon\lambda L - \lambda L,\end{aligned}$$

and the result follows, since $\mathbf{E}_{x_\lambda,\tau}[\varepsilon I] \geq \varepsilon\delta \geq \lambda L$. ■

It remains to deal with the case $w(\tau) \le \delta$. It is more convenient to reduce first the lemma to the case of Markov strategies, and then to prove it for Markov strategies.

Given τ, define for every $n \le L$, $y_n^\tau \in \Delta(B)$ by

$$y_n^\tau = \mathbf{E}_{x_\lambda,\tau}\left[\mathbf{y}_n | \theta > n\right].$$

and denote by y^τ the Markov strategy that plays y_n^τ in stage n, irrespective of past play.

Plainly,

$$w(\tau) = w(y^\tau) \text{ and } \mathbf{E}_{x_\lambda,\tau}\left[v_\lambda(\mathbf{s}_{L+1})\right] = \mathbf{E}_{x_\lambda,y^\tau}\left[v_\lambda(\mathbf{s}_{L+1})\right]. \tag{14}$$

Lemma 15 *If $w(\tau) \le \delta$, then*

$$\mathbf{E}_{x_\lambda,\tau}\left[\widehat{\mathbf{r}}\right] \ge \mathbf{E}_{x_\lambda,y^\tau}\left[\widehat{\mathbf{r}}\right] - 9\varepsilon.$$

Proof. Since $w(\tau) \le \delta \le \varepsilon$, by Lemma 10 we have

$$\left|\mathbf{E}_{x_\lambda,\tau}\left[\widehat{\mathbf{r}}\right] - \mathbf{E}_{x_\lambda,\tau}\left[\widetilde{r}_\lambda(x_\lambda, \overline{\mathbf{y}})\right]\right| \le 3\varepsilon. \tag{15}$$

and

$$\left|\mathbf{E}_{x_\lambda,y^\tau}\left[\widehat{\mathbf{r}}\right] - \mathbf{E}_{x_\lambda,y^\tau}\left[\widetilde{r}_\lambda(x_\lambda, \overline{\mathbf{y}})\right]\right| \le 3\varepsilon. \tag{16}$$

Denoting $\overline{y}^\tau = \frac{1}{L}\sum_{n=1}^L y_n^\tau$, one has $\mathbf{P}_{x_\lambda,y^\tau}\left(\|\overline{\mathbf{y}} - \overline{y}^\tau\| \ge \sqrt{\delta} \text{ and } \theta > L+1\right) \le \sqrt{\delta} \le \varepsilon$, hence

$$\left|\mathbf{E}_{x_\lambda,y^\tau}\left[\widetilde{r}_\lambda(x_\lambda, \overline{\mathbf{y}})\right] - \widetilde{r}_\lambda(x_\lambda, \overline{y}^\tau)\right| \le 2\varepsilon.$$

This implies that

$$\left|\mathbf{E}_{x_\lambda,y^\tau}\left[\widehat{\mathbf{r}}\right] - \widetilde{r}_\lambda(x_\lambda, \overline{y}^\tau)\right| \le 5\varepsilon. \tag{17}$$

Since $w(\tau) \le \delta$,

$$\|y_n^\tau - \mathbf{E}_{x_\lambda,\tau}\left[\mathbf{y}_n\right]\| = \|\mathbf{E}_{x_\lambda,\tau}\left[\mathbf{y}_n \,|\theta > L|\right] - \mathbf{E}_{x_\lambda,\tau}\left[\mathbf{y}_n\right]\| \le 2\delta.$$

By summation over $n = 1, \ldots, L$ we get

$$\|\overline{y}^\tau - \mathbf{E}_{x_\lambda,\tau}\left[\overline{\mathbf{y}}_L\right]\| \le 2\delta.$$

Therefore,

$$\mathbf{E}_{x_\lambda,\tau}\left[\widetilde{r}(x_\lambda, \overline{\mathbf{y}})\right] \ge \widetilde{r}(x_\lambda, \mathbf{E}_{x_\lambda,\tau}\left[\overline{\mathbf{y}}\right]) \ge \widetilde{r}_\lambda(x_\lambda, \overline{y}^\tau) - \varepsilon,$$

where the first inequality follows by convexity of $\widetilde{r}$, and the second one from the choice of δ. Using (15) and (17) we get

$$\mathbf{E}_{x_\lambda,\tau}\left[\widehat{\mathbf{r}}\right] \ge \mathbf{E}_{x_\lambda,y^\tau}\left[\widehat{\mathbf{r}}\right] - 9\varepsilon.$$

■

Lemma 16 *Let y be a Markov strategy such that $w(y) \le \delta$. Then*

$$\mathbf{E}_{x_\lambda,y}\left[\lambda L\widehat{\mathbf{r}} + (1 - \lambda L)v_\lambda(\mathbf{s}_{L+1})\right] \ge v_\lambda - \varepsilon w(y) - 9\varepsilon\lambda L$$

Proof. Call $\overline{y}$ the stationary strategy that plays $\overline{y} = \frac{1}{L}\sum_{n=1}^{L} y_n$ in every stage. Corollary 20 (proved in the appendix) implies that

$$|\mathbf{E}_{x_\lambda,y}[v_\lambda(\mathbf{s}_{L+1})] - \mathbf{E}_{x_\lambda,\overline{y}}[v_\lambda(\mathbf{s}_{L+1})]| \leq \varepsilon w(y). \tag{18}$$

Indeed we set X_i to be 1 if absorption occurs at stage i through some fixed action b under the Markov strategy y and Y_i to be 1 if absorption occurs at stage i through some fixed action b under the average strategy $\overline{y}$. W_i is 1 if absorption occurs at stage i through some action other than b under the Markov strategy y and Z_i to be 1 if absorption occurs at stage i through some action other than b under the average strategy $\overline{y}$. Similarly, Lemma 19 (in the appendix) implies that

$$|w(y) - w(\overline{y})| \leq \varepsilon w(y). \tag{19}$$

Substituting $\tau = y$ in Lemma 15, Eq. (17) yields

$$|\mathbf{E}_{x_\lambda,y}[\widehat{\mathbf{r}}] - \widetilde{r}_\lambda(x_\lambda, \overline{y})| \leq 5\varepsilon. \tag{20}$$

By (19), $w(\overline{y}) \leq w(y)(1+\varepsilon) \leq \delta(1+\varepsilon) \leq 2\varepsilon$, hence

$$\left|\widetilde{r}_\lambda(x_\lambda, \overline{y}) - \mathbf{E}_{x_\lambda,\overline{y}}\left[\frac{1}{L}\sum_{n=1}^{L}\widetilde{r}_\lambda(\mathbf{s}_n, x_\lambda, \overline{y})\right]\right| \leq 2\varepsilon. \tag{21}$$

By the choice of L and λ, $1 > (1-\lambda)^L > 1-\varepsilon$. By (20), (21), and the $\varepsilon\lambda$-optimality of x_λ, one obtains finally

$$\begin{aligned}
&\mathbf{E}_{x_\lambda,y}[\lambda L\widehat{\mathbf{r}} + (1-\lambda L)v_\lambda(\mathbf{s}_{L+1})] \\
&\quad\geq \mathbf{E}_{x_\lambda,y}[\lambda L\widetilde{r}_\lambda(x_\lambda,\overline{y}) + (1-\lambda L)v_\lambda(\mathbf{s}_{L+1})] - 5\varepsilon\lambda L \\
&\quad\geq \mathbf{E}_{x_\lambda,\overline{y}}\left[\lambda\sum_{n=1}^{L}\widetilde{r}_\lambda(\mathbf{s}_n, x_\lambda, \overline{y}) + (1-\lambda L)v_\lambda(\mathbf{s}_{L+1})\right] - 7\varepsilon\lambda L - \varepsilon w(y) \\
&\quad\geq \mathbf{E}_{x_\lambda,\overline{y}}\left[\sum_{n=1}^{L}\lambda(1-\lambda)^{n-1}\widetilde{r}_\lambda(\mathbf{s}_n, x_\lambda, \overline{y}) + (1-\lambda)^L v_\lambda(\mathbf{s}_{L+1})\right] - 8\varepsilon\lambda L - w(y) \\
&\quad\geq v_\lambda - 9\varepsilon\lambda L - \varepsilon w(y).
\end{aligned}$$

■

Corollary 17 *If $w(\tau) \leq \delta$ then (13) holds.*

Proof. By Lemma 15, (14), Lemma 16 and (14) once again,

$$\begin{aligned}
\mathbf{E}_{x_\lambda,\tau}[\lambda L\widehat{\mathbf{r}} + (1-\lambda L)v_\lambda(\mathbf{s}_{L+1})] &\geq \mathbf{E}_{x_\lambda,y^\tau}[\lambda L\widehat{\mathbf{r}} + (1-\lambda L)v_\lambda(\mathbf{s}_{L+1})] - 9\varepsilon\lambda L \\
&\geq v_\lambda - \varepsilon w(y^\tau) - 18\varepsilon\lambda L \\
&\geq v_\lambda - \varepsilon w(\tau) - 18\varepsilon\lambda L,
\end{aligned}$$

and the result follows. ■

7 Absorbing games: the case of no signals

Coulomb (1992) proves that if the game is absorbing and player 1 knows only the current state, then player 1 can guarantee the max-min value by *stationary* strategies. The purpose of this section is to sketch a proof of this earlier result using the above tools. Details are standard and will be omitted.

We let $\varepsilon > 0$ be given. We define $\widetilde{r}_\lambda$, v_λ, x_λ and v as in Section 6.1, $d(a)$ to be the degree in λ of $\lambda \mapsto x_\lambda[a]$, and $\overline{A} = \{a \in A, d(a) < 1\}$. Thus, $\overline{A} = \{a \in A, \lim_{\lambda\to 0} \lambda/x_\lambda(a) = 0\}$. We denote by $\overline{x}_\lambda \in \Delta(A)$ the distribution on $\overline{A}$ induced by x_λ. We shall prove that, depending on cases, x_λ or $\overline{x}_\lambda$ guarantees $v - 4\varepsilon$, for λ sufficiently small. We choose $\lambda_0 > 0$ small enough so that for every $\lambda < \lambda_0$ (i) the support of x_λ is independent of λ, (ii) $\|x_\lambda - \overline{x}_\lambda\|_\infty \leq \varepsilon$, and (iii) $|v_\lambda - v| < \varepsilon$.

For $(x, y) \in \Delta(A) \times \Delta(B)$ we denote by $w(x, y)$ the probability of absorption in a single stage under (x, y), and, if $w(x, y) > 0$, by $u(x, y)$ the expected absorbing payoff.

For every stationary strategy $y \in \Delta(B)$, denote $\alpha_\lambda(y) = \frac{\lambda}{\lambda+(1-\lambda)w(x_\lambda,y)}$. Eq. (10) implies that

$$\alpha_\lambda(y)\widetilde{r}_\lambda(x_\lambda, y) + (1 - \alpha_\lambda(y))u(x_\lambda, y) \geq v_\lambda - \alpha_\lambda(y)\varepsilon > v - 2\varepsilon \tag{22}$$

(see Vrieze and Thuijsman (1989) or Solan (1999) for this formula).

Since player 1 is informed only of transitions, and since $\overline{A}$ contains all actions that matter for the indistinguishability relation, $\widetilde{r}_\lambda(x_\lambda, b)$ is constant for all b such that $w(\overline{x}_\lambda, b) = 0$.

On the other hand, if $w(\overline{x}_\lambda, b) > 0$ then $\lim_{\lambda\to 0} \alpha_\lambda(b) = 0$, hence by (22) $u(x_\lambda, b) \geq v - 3\varepsilon$ for every λ sufficiently close to 0.

Note that in both cases, $|r(x_\lambda, b) - r(\overline{x}_\lambda, b)| \leq \varepsilon$ for λ close enough to 0.

Case 1: For some $b_0 \in B$, $w(\overline{x}_\lambda, b_0) = 0$ and $\widetilde{r}_\lambda(x_\lambda, b_0) < v - 3\varepsilon$.

We claim that, for λ close enough to zero, x_λ guarantees $v_\lambda - 3\varepsilon$. Indeed, let $b \in B$. If $w(\overline{x}_\lambda, b) = 0$, then $b_0 \sim b$, and $\widetilde{r}_\lambda(x_\lambda, b) < v - 3\varepsilon$. By (22), $w(x_\lambda, b) > 0$ and $u(x_\lambda, b) \geq v - 3\varepsilon$. On the other hand, we already know that $u(x_\lambda, b) \geq v - 3\varepsilon$ if $w(\overline{x}_\lambda, b) > 0$. In particular, for every $b \in B$, $w(x_\lambda, b) > 0$ and $u(x_\lambda, b) \geq v - 3\varepsilon$. It follows that for every strategy τ of player 2 $\lim_{n\to\infty} \gamma_n(x_\lambda, \tau) \geq v - 3\varepsilon$.

Case 2: For every $b \in B$, one has $w(\overline{x}_\lambda, b) > 0$ or $r(x_\lambda, b) \geq \widetilde{r}_\lambda(x_\lambda, b) \geq v - 3\varepsilon$.

We claim that, for λ close enough to zero, $\overline{x}_\lambda$ guarantees $v_\lambda - 4\varepsilon$. Indeed, let $b \in B$. If $w(\overline{x}_\lambda, b) = 0$, then $\lim_{n\to\infty} \gamma_n(\overline{x}_\lambda, b) = r(\overline{x}_\lambda, b) \geq r(x_\lambda, b) - \varepsilon \geq v - 4\varepsilon$. If $w(\overline{x}_\lambda, b) > 0$ then for λ sufficiently small $u(\overline{x}_\lambda, b) \geq u(x_\lambda, b) - \varepsilon \geq v - 4\varepsilon$. Thus, for every strategy τ of player 2 $\lim_{n\to\infty} \gamma_n(x_\lambda, \tau) \geq v - 4\varepsilon$, provided λ is sufficiently small.

A Appendix

For any given integer $L \in \mathbf{N}$, the function $1 - \prod_{i=1}^L (1 - x_i)$ is equivalent to $1 - \sum_{i=1}^L x_i$ at zero. The purpose of the next lemma is to verify that this equivalence holds uniformly over L.

Lemma 18 *Consider the function $f_L : [0, 1]^L \to \mathbf{R}$ defined by $f_L(x) = 1 - \prod_{i=1}^L (1 - x_i)$. For every $L \in \mathbf{N}$ and every $x \in [0, 1]^L$ such that $f_L(x) < 1/20$ we have*

$$0 \leq \sum_{i=1}^{L} x_i - f_L(x) \leq 20(f_L(x))^2. \tag{23}$$

Proof. By induction, one has for every $L \in \mathbf{N}$ and every $x_1, ..., x_L \in [0,1]$:

$$\sum_{i=1}^{L} x_i - \left(\sum_{i=1}^{L} x_i\right)^2 \leq f_L(x) \leq \sum_{i=1}^{L} x_i. \tag{24}$$

Write $s = \sum_{i=1}^{L} x_i$ and $f = f_L(x)$. We claim that as long as $s < 1/4$,

$$0 \leq s - f \leq \sum_{n=1}^{\infty} c_n f^{2^n}, \tag{25}$$

where $c_n = 2^1 \times 2^3 \times 2^7 \times \cdots \times 2^{2^n - 1} = 2^{1+3+7+\cdots+(2^n-1)} = 2^{2\times 2^n - (n+1)} = 4^{2^n - \frac{n+1}{2}}$.

Indeed, by (24) we have $s - s^2 \leq f \leq s$, hence $f \leq s \leq f + s^2$. In particular, $0 \leq s - f$. For every $n \in \mathbf{N}$ the function $x \mapsto x^n$ is convex, hence for every $a, b \geq 0$, $(a+b)^n \leq 2^{n-1}a^n + 2^{n-1}b^n$.

Therefore

$$0 \leq s - f \leq s^2 \leq (f + s^2)^2 \leq 2^1 f^2 + 2^1 s^4 \leq 2^1 f^2 + 2^1 (f + s^2)^4 \leq 2^1 f^2 + 2^1 2^3 f^4 + 2^1 2^3 s^8.$$

Continuing inductively we get that for every $N \in \mathbf{N}$,

$$s - f \leq \sum_{n=1}^{N} c_n f^{2^n} + c_N s^{2^{N+1}}.$$

Note that for $s < 1/4$ the last term goes to 0 with N, and the radius of convergence of the power series $\sum_{n=1}^{\infty} c_n t^{2^n}$ is 1/4.

For $f < 1/20$, $\sum_{n=1}^{\infty} c_n f^{2^n} \leq 16f^2 + 1/(1 - 4f^2) \leq 20f^2$, and then $s \leq f + 20f^2 < 1/4$. The lemma follows. ■

Lemma 19 *Let $h(t) = 100t^2$. For every $L \in \mathbf{N}$, and every sequence $(X_i)_{i=1,...,L}$ of independent Bernouilli r.v.s with parameters $p_i \in [0,1]$ the following holds. Set $p = \frac{1}{L}\sum_{i=1}^{L} p_i$, and let $(Y_i)_{=1,...,L}$ be iid Bernouilli variables with parameter p. If $\mathbf{P}\left(\sum_i X_i \geq 1\right) \leq 1/20$ then*

$$\left|\mathbf{P}\left(\sum_i X_i \geq 1\right) - \mathbf{P}\left(\sum_i Y_i \geq 1\right)\right| \leq h\left(\mathbf{P}\left(\sum_i X_i \geq 1\right)\right).$$

Proof. Set $\omega = \mathbf{P}\left(\sum_i X_i \geq 1\right) = 1 - \prod_{i=1}^{L}(1-p_i) \leq 1/20$, and $\omega' = \mathbf{P}\left(\sum_i Y_i \geq 1\right) = 1 - (1-p)^L$. By Lemma 18,

$$\left|\mathbf{P}\left(\sum_i X_i \geq 1\right) - \mathbf{E}\left[\sum_i X_i\right]\right| \leq 20\omega^2, \text{ and}$$
$$\left|\mathbf{P}\left(\sum_i Y_i \geq 1\right) - \mathbf{E}\left[\sum_i Y_i\right]\right| \leq 20\omega'^2.$$

By Lemma 18 and since $1 - (1-p)^L \leq Lp$, $\omega' \leq \omega + 20\omega^2$. Since $\mathbf{E}\left[\sum_i X_i\right] = \mathbf{E}\left[\sum_i Y_i\right]$, the result follows, with the right hand side $h'(t) = 20(t^2 + (t + 20t^2)^2) < 100t^2 = h(t)$ on (0,1/20). ■

Two real valued r.v.s X and Y are *exclusive* if $XY = 0$ a.s. Note that if X and Y are exclusive Bernouilli r.v.s, then $X + Y$ is a Bernouilli r.v. as well.

Corollary 20 *Let $X_1,...,X_L$ (resp. $W_1,...,W_L$) Bernouilli r.v.s with parameter $p_1,...,p_L$ (resp. $q_1,...q_L$). Assume that for every $n \leq L$, (a) X_n and W_n are exclusive, and (b) (X_n, W_n) is independent of $(X_k, W_k)_{k<n}$. Let $Y_1,...,Y_L$ (resp. $Z_1,...,Z_L$) be iid Bernouilli r.v.s with parameter $p = \frac{1}{L}\sum_{n=1}^{L} p_n$ (resp. $q = \frac{1}{L}\sum_{n=1}^{L} q_n$). Assume that (Y_n) and (Z_n) satisfy (a) and (b) as well.*

Define $S_X = \min\{n : X_n = 1\}$, and S_Y, S_W, S_Z similarly. If $\mathbf{P}(\min(S_X, S_W) \leq L) < 1/20$ then

$$|\mathbf{P}(S_X \leq \min(S_W, L)) - \mathbf{P}(S_Y \leq \min(S_Z, L))| \leq h_1(\mathbf{P}(\min(S_X, S_W) \leq L)),$$

where $h_1(t) = 201t^2$.

Proof. Note that

$$\mathbf{P}(S_X \leq L) - \mathbf{P}(S_X \leq \min(S_W, L)) = \mathbf{P}(S_W < S_X \leq L). \tag{26}$$

Note that since $1 - p_i - q_i \leq (1-p_i)(1-q_i)$, $\mathbf{P}(X_i = W_i = 0) \leq \mathbf{P}(X_i = 0)\mathbf{P}(W_i = 0)$, for each $i = 1, \ldots, L$. Therefore, since (X_n, W_n) satisfy (a) and (b),

$$\begin{aligned}
\mathbf{P}(S_W = k, L \geq S_X > k) &= \mathbf{P}(W_k = 1)\mathbf{P}\left(\sum_{i=k+1}^{L} X_i \geq 1\right)\prod_{i=1}^{k-1}\mathbf{P}(X_i = W_i = 0) \\
&\leq \mathbf{P}(W_k = 1)\mathbf{P}\left(\sum_{i=k+1}^{L} X_i \geq 1\right)\prod_{i=1}^{k-1}\mathbf{P}(X_i = 0)\mathbf{P}(W_i = 0) \\
&= \mathbf{P}(S_W = k)\mathbf{P}(L \geq S_X > k) \\
&\leq \mathbf{P}(S_W = k)\mathbf{P}(S_X \leq L).
\end{aligned}$$

Summing up this inequality over $k = 1, \ldots, L$ yields

$$\mathbf{P}(S_W < S_X \leq L) \leq \mathbf{P}(S_W \leq L)\mathbf{P}(S_X \leq L). \tag{27}$$

Plainly, this result applies to Y and Z and one has also

$$\mathbf{P}(S_Z < S_Y \leq L) \leq \mathbf{P}(S_Y \leq L)\mathbf{P}(S_Z \leq L). \tag{28}$$

By (26), (27) and (28),

$$\begin{aligned}
&|\mathbf{P}(S_X \leq \min(S_W, L)) - \mathbf{P}(S_Y \leq \min(S_Z, L))| \\
&\quad \leq |\mathbf{P}(S_X \leq L) - \mathbf{P}(S_Y \leq L)| + \mathbf{P}(S_W \leq L)\mathbf{P}(S_X \leq L) + \mathbf{P}(S_Y \leq L)\mathbf{P}(S_Z \leq L).
\end{aligned} \tag{29}$$

Set $h(t) = 100t^2$. By Lemma 19 the first term in (29) is bounded by

$$|\mathbf{P}(S_X \leq L) - \mathbf{P}(S_Y \leq L)| \leq h(\mathbf{P}(S_X \leq L)) \leq h(\mathbf{P}(\min(S_X, S_W) \leq L)).$$

The second term in (29) satisfies

$$\mathbf{P}(S_W \leq L)\mathbf{P}(S_X \leq L) \leq \mathbf{P}(\min(S_X, S_W) \leq L) \times \mathbf{P}(\min(S_X, S_W) \leq L).$$

Consider the third term in (29). Note that $\mathbf{P}(\min(S_X, S_W) \leq L) = \mathbf{P}(X_1 + W_1 + ... + X_L + W_L \geq 1)$. Applying Lemma 19 to the sequence $(X_k + Y_k)_k$, one has $|\mathbf{P}(\min(S_Y, S_Z) \leq L) - \mathbf{P}(\min(S_X, S_W) \leq L)| \leq h(\mathbf{P}(\min(S_X, S_W) \leq L))$. The proof is complete. ∎

References

[1] Alon N., Spencer J.H. and Erdos P. (2000) The Probabilistic Method, John Wiley and Sons

[2] Aumann, R.J. and M. Maschler, with the collaboration of R.E. Stearns (1995), Repeated games with incomplete information, MIT Press

[3] Aubin J.P. and Cellina A. (1984) Differential Inclusions, Springer Verlag

[4] Aubin J.P. and Frankowska H. (1990) Set-Valued Analysis, Birkhauser

[5] Bochnak J., Coste M. and Roy M. F. (1998) Real Algebraic Geometry, Springer Verlag, Berlin

[6] Coulomb J.M. (1992) Repeated Games with Absorbing States and No Signals, *Int. J. Game Th.*, **21**, 161-174

[7] Coulomb J.M. (1999) Generalized Big-Match, *Math. Oper. Res.*, **24**, 795-816

[8] Coulomb J.M. (2001) Absorbing Games with a Signaling Structure, *Math. Oper. Res.*, **26**, 286-303

[9] Lehrer E. (1989) Lower Equilibrium Payoffs in Two-Player Repeated Games with Nonobservable Actions, *Int. J. Game Th.*, **18**, 57-89

[10] Lehrer E. (1990) Nash Equilibria of n-Player Repeated Games with Semi-Standard Information, *Int. J. Game Th.*, **19**, 191- 217

[11] Lehrer E. (1992a) Correlated Equilibria in Two-Player Repeated Games with Nonobservable Actions, *Math. Oper. Res.* **17**, 175-199

[12] Lehrer E. (1992b) On the Equilibrium Payoffs Set of Two Player Repeated Games with Imperfect Monitoring, *Int. J. Game Th.*, **20**, 211-226

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

[14] Radner R. (1981) Monitoring Cooperative Agreements in Repeated Principle-Agent Relationship, *Econometrica*, **49**, 1127- 1148

[15] Rubinstein A. and Yaari M. (1983) Repeated Insurance Contracts and Moral Hazard, *J. Econ. Th.*, **30**, 74-97

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

[17] Solan E. and Vohra R. (1999), Correlated Equilibrium, Public Signalling and Absorbing Games, D.P. # 1272, The Center for Mathematical Studies in Economics and Management Science, Northwestern University

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