האוניברסיטה העברית בירושלים

THE HEBREW UNIVERSITY OF JERUSALEM

STOCHASTIC GAMES WITH INFORMATION LAG

by

YEHUDA (JOHN) LEVY

Discussion Paper # 499 January 2009

מרכז לחקר הרציונליות

CENTER FOR THE STUDY OF RATIONALITY

Feldman Building, Givat-Ram, 91904 Jerusalem, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

Stochastic Games with Information Lag

Yehuda (John) Levy*

December 28, 2008

Abstract

Two-player zero-sum stochastic games with finite state and action spaces, as well as two-player zero-sum absorbing games with compact metric action spaces, are known to have undiscounted values. We study such games under the assumption that one or both players observe the actions of their opponent after some time-dependent delay. We develop criteria for the rate of growth of the delay such that a player subject to such an information lag can still guarantee himself in the undiscounted game as much as he could have with perfect monitoring. We also demonstrate that the player in the Big Match with the absorbing action subject to information lags which grow too rapidly, according to certain criteria, will not be able to guarantee as much as he could have in the game with perfect monitoring.

1 Introduction

1.1 Background

Stochastic games were introduced by Shapley [28]. In a stochastic game, the players play in stages. At each stage, the game is in one of the available states, and each player chooses an action from the action spaces in that state. The actions chosen then determine a probability distribution on the state space which is used to determine the state at the next stage.

*Center for the Study of Rationality, and Department of Mathematics, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel. Research supported in part by Israel Science Foundation grant 1123/06. This work is the author's M.Sc. thesis (2008). The author expresses gratitude to Prof. Abraham Neyman for his guidance and patience. E–mail: john.calculus@gmail.com

Stochastic games are usually assumed to have a rule which assigns payoffs to the players at each stage, depending on the state and actions played at that stage. In a related model, it is assumed that there is a payoff that is a Borel-measurable function of the entire infinite history of the game. Often, a combination of the two is employed: a Borel payoff is defined in terms of the stage-by-stage payoffs (important examples being the λ-discounted payoff, liminf payoff, and limsup payoff).

A strategy for a player is a rule by which a player chooses, at each stage, a mixed action from his action space in that state. Much of the study of stochastic games has made the assumption that the game has perfect monitoring; that is, the choice of mixed action of a player can depend, at any given time, on the entire history of the game up to that point.

Shapley proved [28] the existence of the λ-discounted value of two-player zero-sum stochastic games with finitely many states and actions, as well as the existence of stationary optimal strategies for each player. (Since the strategies are stationary, this result does not require the perfect monitoring assumption, or for that matter any monitoring of the history of the game except for the knowledge of the current state.)

The existence of the undiscounted value in two-player zero-sum stochastic games with finitely many states and actions was proved, under the assumption of perfect monitoring, in [16], and this value is the limit of the λ-discounted values as $\lambda \to 0$ (the existence of which was proved earlier in [2]).

Stochastic games with more general state and action spaces have been studied (a very general approach is presented in [22]). One such generalization is when the available action spaces are compact metric, and appropriate continuity conditions on the payoffs and transition probabilities are assumed. In this case, the techniques used by Shapley can be generalized to show the existence of the λ-discounted values and stationary optimal strategies.

There is no general result for the existence of the value (or even convergence of the λ-discounted values) in games with compact metric action spaces and finite state space. However, the existence of the undiscounted value was proved for two-player zero-sum absorbing games with compact metric action spaces, with perfect monitoring, in [17].

In [4], a specific example of an absorbing stochastic game, called the *Big*

Match, was discussed. It was demonstrated there that when a player does not view any information about his opponent's previous actions, he may not be able to guarantee in the undiscounted game as much as he could have with perfect monitoring; in particular, the undiscounted value there no longer exists.

It is also of interest to study stochastic games with various types of partial monitoring. Examples include games where the current state is not publicly known [30], games in which a player is not informed of his opponents actions but does receive stochastic signals according to some rule [6],[23], games in which a player observes only his own actions but neither the states or actions of his opponent [24], and stochastic games where a player is informed of his opponent's actions and the states but after some delay [27],[8].

1.2 Objectives

In this thesis, we examined two-player zero-sum stochastic games with a particular type of partial monitoring, called an *information lag*; that is, games such that for a player, there is a function f such that after the n-th stage, that player observes only the first $n - f(n)$ actions of his opponent (however, he observes all the states of the game up through the current one, and his own previous actions[1].)

We posed the following question: What is a sufficient condition on the growth of the information lag function f that a player is subject to, such that he would be able to guarantee in the undiscounted game as much as he could have in the game with perfect monitoring? We established that $f(n) = o(n)$ is a necessary condition for this to be true for the player with the absorbing action in the Big Match (Proposition 4.2.1).

For some time, we had conjectured that this was also the sufficient condition we had been seeking. However, eventually we demonstrated, also for the player with the absorbing action in the Big Match, that $f(n) = o(n)$ is not a sufficient condition (Proposition 4.2.4), and, in fact, if $f(n) >> \frac{n}{\log(\log(n))}$, then this player cannot guarantee, in the undiscounted game, any more than the minimum payoff in the game (Proposition 4.2.5).

[1]The delay in the receipt of information discussed in [27],[8] is of a different kind; the games there can be modeled as stochastic games in which a player also has a delay on the information he receives about the states. These delays in information are also referred to in these papers as information lags.

A sufficient condition, we proved, is given by $f(n) = O(\frac{n}{(\log(n))^\beta})$ for arbitrary real $\beta > 1$ (Theorem 4.1.1). (In Theorem 4.1.2, a slightly more general sufficient condition is presented, of which Theorem 4.1.1 is a prototypical case.) The method of this proof is to generalize the proofs in [16], which showed the existence of the value in the case of games with finite state and action spaces, as well as the similar proof in [17], which showed the existence of the undiscounted value in the case of absorbing games with compact metric action spaces, so as to encompass games with information lags whose growth satisfies the conditions given in Theorem 4.1.2.

Section 4.3 discusses the results, including remarks on the minmax of multi-player games and an application of theorems pertaining to games with Borel payoffs. Section 4.4 demonstrates that it was necessary to modify the proofs in [16] and [17] to encompass games with information lags.

1.3 Acknowledgements

This work would not have been possible without the guidance of Prof. Abraham Neyman, whose patience continues to amaze me. His introducing me to the field of game theory opened for me previously unrealized windows of opportunity, and his extraordinary mathematical overview in this and other fields guided me well in gaining familiarity with the relevant research topics.

Special thanks to my wife, Lois, for her support and encouragement during this past year. As I prepare to finish this thesis, she has been in hospitals for several months, recovering from a horrific traffic accident. I dedicate this work to her complete recovery.

I also wish to thank Orna Barak, Orly Koka, and Zahava Nissim of the administrative staff of the Einstein Institute of Mathematics for helping me deal smoothly and quickly with administrative and financial issues during my M.Sc. studies, and fellow student Andrei Osipov for his help with LaTeX, much encouragement, and useful mathematical insights.

This research was supported in part by Israel Science Foundation grant 1123/06.

2 Notations

Notation 1. *For a bounded real-valued function f on a set A, we denote the supremum norm $||f|| = sup_{x\in A}|f(x)|$.*

Notation 2. $\mathbb{N}$ *denotes the positive integers (without 0).*

Notation 3. *Given a set A and two vectors $u, v \in \mathbb{R}^A$, $u \leq v$ denotes inequality in each coordinate.*

Notation 4. *Given a compact metric space X, $C(X)$ denotes the Banach space of continuous complex-valued functions on X with the supremum norm, and $\Delta(X)$ denotes the space of regular Borel probability measures on X; if X is a compact metric action space, $\Delta(X)$ is the space of mixed actions.*

Notation 5. *The value of the λ-discounted game, if it exists, will be denoted v_λ or $v(\lambda)$. (This is a vector in $\mathbb{R}^S$, i.e., a value for each initial state.)*

Notation 6. *If $f(n)$,$g(n)$ are positive real-valued functions defined for large enough positive integers, $f(n) >> g(n)$ and $g(n) = o(f(n))$ denote $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$.*

Notation 7. *If $f(n)$,$g(n)$ are positive real-valued functions defined for large enough positive integers, $f(n) \sim g(n)$ denotes $\lim_{n\to\infty} \frac{g(n)}{f(n)} = 1$.*

Notation 8. *Given a finite set A, $|A|$ and $\#A$ denote the number of elements in A.*

Notation 9. *The value of a game in normal form, if it exists, will be denoted $val_{X\times Y}(g)$, where X, Y are the action spaces for Players 1 and 2 respectively, and g is the payoff function.*

3 Definitions and Previous Results

3.1 Definitions

Definition 3.1.1. A two-player zero-sum *stochastic game* (with compact or finite action spaces) consists of the following:

- A set S (the collection of states) assumed to be finite unless otherwise specified.[2]

[2] If the state space is not discrete, but rather endowed with a measurable structure, certain measurability conditions must be assumed on the payoff functions and probability transitions defined here; see [22]. Such matters will not concern us here.

- For each $z \in S$, compact metric or finite sets A_z, B_z of actions for Player 1 and Player 2, respectively.
- For each $z \in S$, a separately continuous payoff function, $g(z, \cdot, \cdot) : A_z \times B_z \to \mathbb{R}$, such that $g(\cdot, \cdot, \cdot)$ is bounded.
- For each $z \in S$, $a \in A_z$,$b \in B_z$, a probability distribution $q(z, a, b)$ on S, such that for each $z, z' \in S$, $q(z, \cdot, \cdot)(z')$ is separately continuous.[3]

By *stochastic game* I refer to the two-player, zero-sum case defined here, unless otherwise specified.

Definition 3.1.2. A state $z \in S$ in a stochastic game is called *absorbing* if for every $(a, b) \in A_z \times B_z$, $q(z, a, b)(z) = 1$. An *absorbing game* is a stochastic game in which every state but one is absorbing.

Remark 3.1.3. The continuity conditions on the transition probability and the payoff function imply measurability ([19, I.1, Ex.7a], [15, p. 238]).

Remark 3.1.4. A reference to a stochastic game is technically a reference to a collection of games: each $z_1 \in S$ defines a game in which z_1 is the *initial state*.

A stochastic game is played in stages. At stage $n \in \mathbb{N}$, each player chooses an action in the set of actions available to him in the current state z_n (Player 1 chooses $a_n \in A_{z_n}$, Player 2 chooses $b_n \in B_{z_n}$). The payoff for that stage is then $g(z_n, a_n, b_n)$ (Player 2 pays Player 1) and the next state is chosen according to the probability distribution $q(z_n, a_n, b_n)$.

In general we will assume a stochastic game to have infinitely many stages, unless otherwise specified. But we will also refer to games of finite length; that is, games consisting of finitely many stages.

Given a stochastic game, we find it is convenient to define several sets and σ-algebras pertaining to the space of plays of the game:

- H_n is the set of finite histories of the form

$$h_n := (z_1, a_1, b_1, \ldots, z_{n-1}, a_{n-1}, b_{n-1}, z_n),$$

  where $z_k \in S$, $(a_k, b_k) \in A_{z_k} \times B_{z_k}$.

[3]This continuity condition is also stronger than is usually used in the study of stochastic games with infinite state spaces; for example, [22]. However, in the case of a finite state space, the various continuity conditions coincide.

- H_∞ is the set of infinite histories of the form $h = (z_1, a_1, b_1, z_2, \ldots)$, where $z_k \in S$, $(a_k, b_k) \in A_{z_k} \times B_{z_k}$.

- $\mathscr{H}_n$ denotes the product σ-algebra of $\left(\bigcup_{z\in S}(\{z\} \times A_z \times B_z)\right)^{n-1} \times S$ on H_n, where A_z, B_z are endowed with the Borel σ-algebra, and also its inverse image on H_∞ via the projections $H \to H_n$ given by $(z_1, a_1, b_1, \ldots) \to (z_1, a_1, b_1, \ldots, z_{n-1}, a_{n-1}, b_{n-1}, z_n)$. (It will always be clear from the context which σ-algebra we are referring to.)

- $\mathscr{H}_\infty$ is the σ-algebra on H_∞ generated by $\bigcup_{n\in\mathbb{N}} \mathscr{H}_n$.

Definition 3.1.5. Given a compact metric space X, we define the weak topology on $\Delta(X)$ to be the topology induced on it as a subset of the dual space of $C(X)$ with the weak-* topology; that is, the weak topology on $\Delta(X)$ is the weakest topology such that all the linear functionals of the form $\mu \to \int_X f d\mu$, where $f \in C(X)$, are continuous.

If X is a compact metric space, then reference to a measurable mapping from or into $\Delta(X)$ assumes $\Delta(X)$ to be endowed with the Borel σ-algebra generated by the weak topology; reference to a continuous mapping from $\Delta(X)$ assumes $\Delta(X)$ to be endowed with the weak toplogy.

In the study of games with *perfect monitoring* (see Remark 3.1.7), the following class of strategies is used:

Definition 3.1.6. A *behavioral strategy*[4] σ for Player 1 is a collection of measurable mappings, one for each $(H_n, \mathscr{H}_n)$, such that $\sigma((z_1, a_1, b_1, \ldots, z_n)) \in \Delta(A_{z_n})$. Similarly we define a behavioral strategy τ for Player 2. A behaviorial strategy is said to be *stationary* if it is dependent only on the current state; it is said to be *Markov stationary* (or just *Markov*) if it is dependent only on the current state and the number of the stages so far. Stationary strategies are often described by a listing of mixed actions, one for each state in S. A strategy is *pure* if for every finite history h the support of $\sigma(h)$ is a single element of the available action space.

For brevity, behavioral strategies will be referred to simply as *strategies.*

[4]There is also a class of strategies known as *mixed strategies.* Due to the complexity of the definition of this class of strategies in games with non-discrete action spaces, and due to their equivalence to behavioral strategies in the class of games considered in this thesis [1], we will not concern ourselves with such matters here.

Remark 3.1.7. If a player in a game has no further restrictions placed on the strategies that he can choose from, with respect to the dependency of the strategy on the history of the game, he is said to have *perfect monitoring.* Intuitively, a player has perfect monitoring if he can "see", at any point in the game, the entire history of the game up to this point.

Games in which a player does not observe all the information pertaining to the history of the game up to the present do not have perfect monitoring.

An initial state $z_1 \in S$ and profile of behavioral strategies (σ, τ) define a probability distribution on $(H_\infty, \mathscr{H}_\infty)$, $P^{z_1}_{\sigma,\tau}$, and the expected value operator with respect to this distribution, $E^{z_1}_{\sigma,\tau}$, as previously described.

Remark 3.1.8. If as in Definition 3.1.1, $q(\cdot,\cdot,\cdot)$ is a probability transition function, given a bounded measurable function u so that for each $z \in S$, $u(z,\cdot,\cdot,\cdot) : A_z \times B_z \times S \to \mathbb{R}$, we will often write an expression of the form

$$E_{x,y,q}[u(z,a,b,z')],$$

where z is a given state and either $x = (x_s)_{s\in S}, y = (y_s)_{s\in S}$ are stationary strategies for the players or $x = x_z, y = y_z$ are mixed actions in state z for the players. Informally, this means that (a,b) distributes according to the product measure $x_z \otimes y_z$ on $A_z \times B_z$, and z' then distributes according to $q(z,a,b)$. Formally, this means

$$\begin{aligned} &E_{x_z,y_z}\left[E_{q(z,a,b)}[u(z,a,b,z')]\right] \\ &\qquad = \int_{A_z\times B_z} d(x_z \otimes y_z)(a,b) \int_S u(z,a,b,z')d(q(z,a,b))(z') \end{aligned}$$

(This is well defined by Remark 3.1.3 and Fubini's theorem.)

We now address the issue of payoffs in the infinite-stage game and the relevant value concepts.

Definition 3.1.9. In a stochastic game with a countable discrete state space and finite action spaces, let $\varphi : H_\infty \to \mathbb{R}$ be a bounded function which is measurable with respect to $\mathscr{H}_\infty$. Given a strategy profile (σ,τ) and an initial state $z_1 \in S$, we define

$$E^{z_1}_{\sigma,\tau}[\varphi]$$

to be the φ-payoff (Player 2 pays Player 1). If there is $v \in \mathbb{R}^S$ such that for all $\varepsilon > 0$ there is a strategy σ_ε of Player 1 such that for every strategy τ of Player 2 and all $z_1 \in S$,

$$E^{z_1}_{\sigma_\varepsilon,\tau}[\varphi] \geq v(z_1) - \varepsilon$$

then we say that Player 1 can *guarantee* v in the φ-game. Similarly, if for all $\varepsilon > 0$ there is a strategy τ_ε of Player 2 such that for every strategy σ of Player 1 and all $z_1 \in S$,

$$E^{z_1}_{\sigma,\tau_\varepsilon}[\varphi] \leq v(z_1) + \varepsilon$$

then we say that Player 2 can *guarantee* v in the φ-game. If there is $v \in \mathbb{R}^S$ that both players can guarantee, it is referred to as the *value* of the φ-game.

A few important examples of Borel payoffs are given in Definitions 3.1.10, 3.1.11, and 3.1.12 below.

Definition 3.1.10. For $\lambda \in (0,1)$, given a strategy profile (σ, τ) and an initial state $z_1 \in S$, define the λ-*discounted payoff* (to Player 1) by

$$E^{z_1}_{\sigma,\tau}\Big[\sum_{k=1}^{\infty} \lambda(1-\lambda)^{k-1} g_k\Big]$$

where $g_k = g(z_k, a_k, b_k)$ is the payoff to Player 1 at stage k.

Theorem 3.3.6 will show that under our assumptions, the value of the λ-discounted game, v_λ, exists and can be characterized.

Definition 3.1.11. Given a strategy profile (σ, τ) and initial state an z_1, we will refer to

$$E^{z_1}_{\sigma,\tau}\Big[\liminf_{n\to\infty} \frac{\sum_{k=1}^{n} g_k}{n}\Big]$$

where $g_k = g(z_k, a_k, b_k)$, as the *liminf* payoff (to Player 1). Similarly, we can define the *limsup* payoff.

Definition 3.1.12. Given a strategy profile (σ, τ), an initial state z_1, and $N \in \mathbb{N}$, we will refer to

$$E^{z_1}_{\sigma,\tau}\Big[\frac{\sum_{k=1}^{N} g_k}{N}\Big]$$

where $g_k = g(z_k, a_k, b_k)$, as the *average* payoff (to Player 1) in the *finite N-stage game.*

Definition 3.1.12 also defines a payoff for games of finite length N.

Our primary value concept is given in the following definition, which was given in [16]:

Definition 3.1.13. Player 1 can *guarantee* $v \in \mathbb{R}^S$ (in the undiscounted game) if for every $\varepsilon > 0$ there is a strategy σ_ε for Player 1 and $N \in \mathbb{N}$ such that for every strategy τ of Player 2 and all $z_1 \in S$,

$$E^{z_1}_{\sigma_\varepsilon,\tau}\left[\liminf_{n\to\infty} \frac{\sum_{k=1}^n g_k}{n}\right] \geq v(z_1) - \varepsilon$$

and for all $n \geq N$,

$$E^{z_1}_{\sigma_\varepsilon,\tau}\left[\frac{\sum_{k=1}^n g_k}{n}\right] \geq v(z_1) - \varepsilon$$

where $g_k = g(z_k, a_k, b_k)$.

Player 2 can *guarantee* $v \in \mathbb{R}^S$ (in the undiscounted game) if for every $\varepsilon > 0$ there is a strategy τ_ε for Player 2 and $N \in \mathbb{N}$ such that for every strategy σ of Player 1 and all $z_1 \in S$,

$$E^{z_1}_{\sigma,\tau_\varepsilon}\left[\limsup_{n\to\infty} \frac{\sum_{k=1}^n g_k}{n}\right] \leq v(z_1) + \varepsilon$$

and for all $n \geq N$,

$$E^{z_1}_{\sigma,\tau_\varepsilon}\left[\frac{\sum_{k=1}^n g_k}{n}\right] \leq v(z_1) + \varepsilon$$

A stochastic game has an *undiscounted value* $v \in \mathbb{R}^S$ if both players can guarantee v.

Remark 3.1.14. Strategies $\sigma_\varepsilon,\ \tau_\varepsilon$, described in Definition 3.1.9 and Definition 3.1.13, are referred to as ε-*optimal.*

Other value concepts and examples can be found in [21]. Other payoff concepts with examples can also be found in [3].

Remark 3.1.15. In all models that we will consider, both players have perfect monitoring of sequence of states. Therefore, in an absorbing game, both players know when an absorbing state z is reached; the rest of the game is a repeated zero-sum game with value $v(z) = val_{A_z \times B_z}(g(z,\cdot,\cdot))$, where g denotes the payoff function for the game (see Theorem 3.3.4). Therefore, by replacing all the absorbing states with a pair of absorbing states s^+, s^- with values $\pm||g||$, and then replacing the probability of a transition to state z with transitions to s^+ or s^- with probabilities $\frac{1}{2}(1 \pm \frac{v(z)}{||g||})$, we can always reduce a game with infinitely many absorbing states to a game with two absorbing states.

Definition 3.1.16. A *lag function* is a function $f : \mathbb{N} \bigcup \{0\} \to \mathbb{N} \bigcup \{0\}$ which satisfies $f(0) = 0$ and

$$f(n+1) \leq f(n) + 1 \tag{3.1.1}$$

A strategy is *subject to an information lag* f if, at the $n+1$-th stage, the choice of mixed action is independent of the actions of the opponent from stages $n - f(n) + 1$ through n.

If f is a lag function, we say that a player is *subject to an information lag* f if he can only play behavioral strategies subject to this information lag.[5]

Explicitly, a behavioral strategy – say, σ for Player 1 – is subject to an information lag f if it is dependent, at the $n+1$-th stage, only on the sequence of previous and current states $(z_1, \ldots, z_{n+1})$, the sequence of Player 1's previous actions $(a_1, \ldots, a_n)$, and the sequence of the first $n - f(n)$ actions of Player 2 $(b_1, \ldots, b_{n-f(n)})$.

Intuitively, we see that if a player is subject to an information lag f, then after stage n, he does not "see" the last $f(n)$ of his opponent's actions. The condition (3.1.1) implies that a player subject to an information lag doesn't forget what he knows. (Since a player also recalls his own actions, and the sequence of states, the game has perfect recall.)

Definition 3.1.17. Suppose as in Definition 3.1.1, we have a stochastic game with compact metric action spaces. Define, for each $\lambda \in [0, 1)$, $\Phi(\lambda, \cdot) : \mathbb{R}^S \to \mathbb{R}^S$, by

$$\Phi(\lambda, u)(z) = max_{x \in \Delta(A_z)} min_{y \in \Delta(B_z)} E_{x,y,q} \Big[\lambda g(z, a, b) + (1 - \lambda) u(z') \Big] \tag{3.1.2}$$

This is known as the *discounted Shapley operator*.

Remark 3.1.18. For games with compact action spaces, the discounted Shapley operator is well defined by Theorem 3.3.4.

Remark 3.1.19. Equivalently,

$$\Phi(\lambda, u)(z) = val_{A_z \times B_z}(\lambda g(z, a, b) + (1 - \lambda) E_{q(z,a,b)}(u(z')))$$

[5] As mentioned earlier, the game studied in [8] can be modeled as a stochastic game in which there is an information lag of a stronger type, such that a player does not have perfect monitoring over the sequence of states. Comparing the results there to the results that will be presented here (Theorem 4.1.1) demonstrates that the sort of information lag discussed there is significantly more severe than that discussed here. By their usage of the term, even a bounded information lag can ruin the existence of the undiscounted value.

Remark 3.1.20. Note that $\Phi(\lambda, \cdot)$ is monotonic: If $u \geq v$, then $\Phi(\lambda, u) \geq \Phi(\lambda, v)$.

Remark 3.1.21. For $\lambda \in (0, 1)$, $\Phi(\lambda, \cdot)$ is a contraction on $\mathbb{R}^S$ with the $||\cdot||$ norm, $||\Phi(\lambda, u) - \Phi(\lambda, v)|| \leq (1 - \lambda)||u - v||$.

Remark 3.1.22. As mentioned in Theorem 3.3.1 (for finite action spaces) and Theorem 3.3.6 (for compact action spaces), the λ-discounted value v_λ is the unique fixed point of the operator $\Phi(\lambda, \cdot)$.

3.2 The Big Match

A stochastic game of particular importance and interest is the Big Match, introduced by D. Gillette in 1957 [9]. The game is an absorbing game given by

	L	R
T	1	-1
B	-1^*	1^*

where the notation a^* means the game moves on to an absorbing state such that at this stage and every subsequent stage, the payoff is a. (Player 1 is the row player; Player 2 is the column player.) The Big Match is thus an absorbing game with two absorbing states.

Theorem 3.2.1 ([9]). *In the Big Match:*

- *If the players are restricted to using only stationary strategies, then the most Player 1 can guarantee in the undiscounted game is* -1*; the most Player 2 can guarantee is* 0 *(by using the strategy* $(\frac{1}{2}, \frac{1}{2})$*).*
- *The same result holds if the players are restricted to using Markov strategies.*

The question of the undiscounted value[6] of Big Match was eventually solved by D. Blackwell and T.S. Ferguson (1968).

Theorem 3.2.2 ([4]). *In the Big Match, in behavioral strategies the undiscounted value of the game is* 0*; for Player 2,* $(\frac{1}{2}, \frac{1}{2})$ *is an optimal strategy, while Player 1 has no optimal strategies (only* ε*-optimal strategies).*

[6]Their article worked with the limsup value, not with the stronger concept of the undiscounted value as defined in [16] and in Definition 3.1.13. However, the proof given in [4, Section 2] effectively shows the existence of the undiscounted value.

Remark 3.2.3. In the Big Match, we will refer to the event of Player 1 playing B as *absorption.*

It's also worth noting that in the λ-discounted game, the optimal stationary strategy for Player 1 is $(\frac{1}{1+\lambda}, \frac{\lambda}{1+\lambda})$ [31]; as before, the optimal stationary strategy for Player 2 is $(\frac{1}{2}, \frac{1}{2})$.

The results of [9] are [4] are presented neatly in [31].

3.3 Previous Results

I now state several known results in game theory and in the theory of discounted stochastic games.

The first result is that of Shapley, given in the paper in which he introduced stochastic games [28]. Shapley's formulation of this result was not for discounted games, but for games with positive stopping probabilities; that is, games with a positive probability of terminating at each stage, depending on the current state. Derivations of this result in the language of discounted games can be found, for example, in [29] and [2].

Theorem 3.3.1 ([28]). *In a stochastic game with finite state and action spaces, for each $\lambda \in (0,1)$ the value v_λ of the λ-discounted game exists and each player possesses stationary optimal strategies. Furtheremore,the value v_λ is the unique solution to the equation*

$$v_\lambda = \Phi(\lambda, v_\lambda) \tag{3.3.1}$$

where $\Phi(\lambda, \cdot)$ is the discounted Shapley operator.

A result of fundamental importance to the theory of stochastic games with finite state and actions spaces, due to T. Bewley and E. Kohlberg (1976), is as follows:[7]

Theorem 3.3.2 ([2]). *For a stochastic game with finite state and action spaces, there exists $M \in \mathbb{N}$,$\lambda_0 > 0$, and for each $z \in S$, $\{a_{k,z}\}_{k=0}^{\infty} \subseteq \mathbb{R}^S$, $\{p_{k,z}\}_{k=0}^{\infty} \subseteq \mathbb{R}^{A_z}$, $\{q_{k,z}\}_{k=0}^{\infty} \subseteq \mathbb{R}^{B_z}$ such that for $\lambda \in (0, \lambda_0)$ and $z \in S$,*

$$v_\lambda(z) = \sum_{k=0}^{\infty} a_{k,z} \lambda^{k/M} \tag{3.3.2}$$

[7] An alternative approach to these results is presented in [20].

and such that

$$p_\lambda(z) = \sum_{k=0}^{\infty} p_{k,z}\lambda^{k/M}, \; q_\lambda(z) = \sum_{k=0}^{\infty} q_{k,z}\lambda^{k/M} \tag{3.3.3}$$

are stationary optimal stationary strategies in the λ-discounted game for Player 1 and Player 2, respectively.

Remark 3.3.3. A series of the form

$$\sum_{k=-N}^{\infty} a_k\lambda^{k/M}$$

is known as a *Puiseux* series.

The following is a generalized minmax theorem about games in normal form with compact metric action spaces.

Theorem 3.3.4 ([15]; [19], Chapter I, Theorem 2.6)**.** *A two-person zero-sum game with compact metric action spaces,*[8] *such that the payoff function is upper semi-continuous in Player 1's strategy, lower semi-continuous in Player 2's strategy, and bounded either from above or below, has a value, and each player has optimal (mixed) strategies.*

Remark 3.3.5. It can be shown that under the conditions of Theorem 3.3.4, each player has ϵ-optimal strategies with finite support; see [15] and [19, Chapter I].

Using generalized minmax theorems like Theorem 3.3.4, the argument given by Shapley in [28] can be generalized:

Theorem 3.3.6 ([19], Chapter VI, Proposition 1.4)**.** *The existence of the λ-discounted value, the existence of stationary optimal strategies, and the characterization of v_λ by equation* (3.3.1) *hold for stochastic games with finite state space and compact action spaces.*

This theorem can also be deduced as a corollary of the much more general results given in [22].

[8] One can also speak, more generally, of compact Haussdorf action spaces; see [15] and [19, Chapter I].

The following result, due to D. Martin (1998), pertains to stochastic games with Borel payoffs.[9]

Theorem 3.3.7 ([14]). *A stochastic game with a countable state space, finite action spaces, and (bounded) Borel payoff, has a value.*

3.4 The Operator Approach to Stochastic Games

In this section, we present the results of an operator approach to two-player zero-sum games with compact action spaces discussed in [25], and demonstrate the measurability of the relevant selection of strategies that will be used in Section 5.3.

Proposition 3.4.1. *In a stochastic game with finite state space S and compact metric action spaces, for $z \in S,\ u \in \mathbb{R}^S$,*

$$\varphi^*(u)(z) := \lim_{\alpha \to 0^+} \frac{\Phi(\alpha, u)(z) - u(z)}{\alpha}$$

exists in $\mathbb{R} \bigcup \{\infty, -\infty\}$, where $\Phi(\lambda, \cdot)$ is the discounted Shapley operator.

In the following propositions, we will adopt the conventions that $\infty - (-\infty) = \infty$ and $-\infty - \infty = -\infty$.

Proposition 3.4.2. *In a stochastic game with finite state space S and compact metric action spaces, if $u, v \in \mathbb{R}^S$ and $z \in S$ satisfy*

$$u(z) - v(z) = \delta = max_{z' \in S}(u - v)(z') > 0,$$

then either $\varphi^*(u)(z) = \varphi^*(v)(z) = \pm\infty$, or

$$\varphi^*(v)(z) - \varphi^*(u)(z) \geq \delta.$$

Now, suppose the game of the preceding propositions is an absorbing stochastic game, with payoff function g, probability transition function q, and action spaces $(A_z)_{z \in S}$, $(B_z)_{z \in S}$. (By Remark 3.1.15, we needn't place any restrictions on the state space.) Let z_0 be the non-absorbing state. For each $u \in \mathbb{R}$, we define $\hat{u} \in \mathbb{R}^S$ by

[9] The results in [14] are proved there for a different class of games, and the corollary for stochastic games is deduced afterwards. An approach to these results directed explicitly at stochastic games is well presented in [13]. Note that this latter article makes for simplicity the assumption that all states have the same finite action spaces, but this is not a necessary assumption.

$$\hat{u}(z) = \begin{cases} u & \text{if } z = z_0 \\ val_{A_z \times B_z}(g(z,\cdot,\cdot)) & \text{if } z \neq z_0 \end{cases}$$

We can treat $\Phi(\lambda, \cdot)$ as an operator from $\mathbb{R}$ to $\mathbb{R}$ via

$$\Phi(\lambda, u) = \Phi(\lambda, \hat{u})(z_0)$$

and we can define $\varphi^*(\cdot) : \mathbb{R} \to \mathbb{R} \bigcup \{\infty, -\infty\}$ by

$$\varphi^*(u) = \varphi^*(\hat{u})(z_0)$$

The previous proposition thus becomes:

Proposition 3.4.3. *In an absorbing game with compact metric action spaces, if $u, v \in \mathbb{R}$, $u > v$, then either $\varphi^*(u) = \varphi^*(v) = \pm\infty$, or*

$$\varphi^*(v) - \varphi^*(u) \geq u - v$$

Since if $u \geq ||g||$ then $\varphi^*(u) \leq 0$, and if $u \leq -||g||$ then $\varphi^*(u) \geq 0$, we use the proposition to deduce:

Proposition 3.4.4. *In an absorbing game with compact metric action spaces, there exists a unique $w \in \mathbb{R}$ such that*

$$w' < w \Longrightarrow \varphi^*(w') > 0$$

$$w'' > w \Longrightarrow \varphi^*(w'') < 0$$

Furthermore, $w = \Phi(0, w)$.

Using other results from [25], it is demonstrated that:

Theorem 3.4.5. *In an absorbing game with compact metric action spaces, the limit $\lim_{\lambda \to 0^+} v_\lambda$ exists and is equal to $\hat{w}$, where w is the same as in the previous proposition.*

These last two results, along with the definition of φ^*, lead us to the following corollary (which was used in [17]):

Corollary 3.4.6. *In an absorbing game with compact metric action spaces, non-absorbing state z_0, payoff function g, and actions spaces $(A_z)_{z \in S}$, $(B_z)_{z \in S}$, let $\varepsilon > 0$, define*

$$w(z) = \begin{cases} \lim_{\lambda \to 0^+} v_\lambda(z_0) - \varepsilon & \text{if } z = z_0 \\ val_{A_z \times B_z}(g(z,\cdot,\cdot)) & \text{if } z \neq z_0 \end{cases}$$

Then, for $\lambda > 0$ small enough,

$$\Phi(\lambda, w) \geq w$$

This corollary implies that for this w, and each λ small enough, Player 1 has a mixed action x_λ in z_0 so that for all mixed action y in z_0 of Player 2,

$$E_{x_\lambda,y,q}[\lambda g(z_0,a,b)+(1-\lambda)w(z')] \geq w(z_0) \tag{3.4.1}$$

This, however, is not enough for us; we want this assignment $\lambda \to x_\lambda$ to be measurable.

Let λ_0 be such that the conclusion of Corollary 3.4.6 holds for $\lambda \in I := (0,\lambda_0)$. Let Ω be the set of pairs $(\lambda,x) \in I \times \Delta(A_{z_0})$ that satisfy (3.4.1) for all $y \in \Delta(B_{z_0})$, and for each $\lambda \in I$, denote $\Omega_\lambda = \{x \in \Delta(A_{z_0}) \big| (\lambda,x) \in \Omega\}$.

From Remark 3.1.8 and the fact that the payoff and the probability transition are separately continuous, it's easy to see that the mapping $(I \times \Delta(A_{z_0})) \times \Delta(B_{z_0}) \to \mathbb{R}$ given by

$$((\lambda,x),y) \to E_{x,y,q}[\lambda g(z_0,a,b)+(1-\lambda)w(z')]$$

is separately continuous in (λ,x) and in y, where we recall that $\Delta(A_{z_0})$ and $\Delta(B_{z_0})$ are endowed with the weak topology. $\Delta(A_{z_0})$ is well known to be compact metric (e.g., [19, Section I.1], or [26, Sections 3.15 and 3.16]) and therefore each Ω_λ is compact, and Ω is Borel.

We recall the following selector theorem:[10]

Theorem 3.4.7 ([5]). *Let U,V be complete, separable metric spaces, and let $\Omega \subseteq U \times V$ be a Borel set. Denote $Proj(\Omega) = \{u \in U | \exists v \in V, (u,v) \in \Omega\}$. Suppose that for each $u \in U$, $\Omega_u = \{v \in V \big| (u,v) \in \Omega\}$ is compact. Then there is a Borel-measurable function $f : Proj(\Omega) \to V$ such that for all $u \in Proj(\Omega)$, $(u,f(u)) \in \Omega$.*

Therefore, we deduce the following corollary:

Corollary 3.4.8. *Under the conditions of Corollary 3.4.6, there exists λ_0 such that there is a Borel-measurable mapping $(0,\lambda_0) \to \Delta(A_{z_0})$ (where $\Delta(A_{z_0})$ is endowed with the weak topology), $\lambda \to x_\lambda$, so that for all $\lambda \in (0,\lambda_0)$ and all $y \in \Delta(B_{z_0})$,*

$$E_{x_\lambda,y,q}[\lambda g(z_0,a,b)+(1-\lambda)w(z')] \geq w(z_0)$$

Remark 3.4.9. If $z \neq z_0$, by the definition of $w(z)$ there is $x \in \Delta(A_z)$ such that for all $\lambda \in (0,1)$ and all $y \in \Delta(B_z)$,

$$E_{x,y,q}[\lambda g(z,a,b)+(1-\lambda)w(z')] = \lambda E_{x,y}[g(z,a,b)]+(1-\lambda)w(z) \geq w(z)$$

[10] This theorem also results immediately from the general selector theorem in [12], using [10, Theorem 3], or from the selector theorem in [11, Section 14.2], using [11, Section 13.2].

3.5 Results in Undiscounted Stochastic Games

The following two results show the existence of the value in undiscounted stochastic games under proper conditions. (The aim of this thesis is to extend these results to the case where an information lag is imposed.)

The first is due to J.F. Mertens and A. Neyman (1981).

Theorem 3.5.1 ([16]). *A stochastic game with finite state and action spaces, and perfect monitoring, has a value.*

The second is due to J.F. Mertens, A. Neyman and D. Rosenberg (2007).

Theorem 3.5.2 ([17],[18]). *An absorbing game with compact metric action spaces, and perfect monitoring, has a value.*

Remark 3.5.3. By Remark 3.1.15, we needn't impose any restrictions on the state space in Theorem 3.5.2.

Remark 3.5.4. As demonstrated by the explicit constructions in the respective articles, the ϵ-optimal strategies in both these cases can be chosen such that they depend only, at stage $n+1$, on the sequence of states up to to that point, $(z_1, \ldots, z_{n+1})$, and the sequence of payoffs in the game up to that point, $(g_1, \ldots, g_n)$.

4 The Results

4.1 Conditions for the Existence of the Value

This theorem, the main result of this work, extends the results on the existence of the undiscounted value of those games mentioned in Section 3.5 to games with information lags that are not too severe.

Theorem 4.1.1. *Suppose we have a two-player zero-sum stochastic game with finite state and action spaces, or a two-player zero-sum absorbing game with compact metric action spaces. Suppose a player is subject to an information lag given by a lag function f such that there is $\beta > 1$ such that*[11] $f(n) = O\big(\frac{n}{(\log(n))^\beta}\big)$. *Then that player can guarantee in the undiscounted game whatever he can guarantee in the game with perfect monitoring. In particular, if one or both players have such information lags, the game has an undiscounted value.*

[11] Or $f(n) = O\big(\frac{n}{\log(n)(\log(\log(n)))^\beta}\big)$ or $O\big(\frac{n}{\log(n)\log(\log(n))(\log(\log(\log(n))))^\beta}\big)$, etc.

In Sections 5.3–5.5, we will prove the following more general result, and then deduce Theorem 4.1.1 as a corollary thereof in Section 5.6.

Theorem 4.1.2. *Suppose we have a two-player zero-sum stochastic game with finite state and action spaces, or a two-player zero-sum absorbing game with compact metric action spaces. Let $v(\lambda)$ denote the λ-discounted value. Suppose a player is subject to an information lag given by a lag function f such that for every $\varepsilon > 0$:*

1) There exists a non-decreasing lag function $g(n)$ so that $f(n) = o(g(n))$.

2) There exists a decreasing function $\lambda(s) : (0, \infty) \to (0, 1)$ satisfying the following conditions:

a) $\int_0^\infty \lambda(s)ds < \infty$.

b) For all $M > 0$, $\lambda(M + g(n)) \cdot n \to \infty$.

Moreover, either of the following pairs (c.I, d.I) or (c.II, d.II) of conditions hold:

c.I) For M large enough, and any k, if $s \geq M + g(k)$, then

$$\Big|\frac{\lambda(s \pm \max(1, f(k)))}{\lambda(s)} - 1\Big| \leq \varepsilon.$$

d.I) The game is absorbing, or for s large enough

$$||v(\lambda(s \pm 1)) - v(\lambda(s))|| \leq \varepsilon \cdot \lambda(s)$$

c.II) For every D small enough and s large,

$$\Big|\frac{\lambda((1 \pm D)s)}{\lambda(s)} - 1\Big| \leq \varepsilon.$$

d.II) The game is absorbing, or for every D small enough, and s large enough,

$$||v(\lambda((1 \pm D)s)) - v(\lambda(s))|| \leq \varepsilon \cdot s \cdot \lambda(s)$$

Then that player can guarantee in the undiscounted game whatever he could guarantee in the game with perfect monitoring. In particular, if one or both players have such information lags, the game has an undiscounted value.

We clarify that conditions (d.I) and (d.II) are of relevance in those games which are not absorbing. If the game is an absorbing game, then (d.I) and (d.II) are to be ignored.

The two pairs of conditions (c.I, d.I) and (c.II, d.II) correspond to two different proofs. The first will extend the proof in [16, Section 2], while the second will extend the proof in [16, Section 3].

4.2 Large Information Lags

These results deal with games in which a player has an information lag that grows fast enough so that the undiscounted value of the game no longer exists. All these results pertain to the Big Match (defined in Section 3.2).

This following proposition, proved in Section 5.1, shows that if the information lag is too large, the undiscounted value of the game no longer exists.

Proposition 4.2.1. *Suppose that in the Big Match, Player 1 is subject to an information lag $f(n)$ that is not $o(n)$. Then Player 1 cannot guarantee a liminf payoff of more than* $-\frac{1}{2}\limsup_{n\to\infty}\frac{f(n)}{n}$.

Player 2 may be able to guarantee a liminf payoff of much less. A technique provided by E. Shmaya (private communication) shows that:

Proposition 4.2.2. *Suppose that in the Big Match, Player 1 is subject to an information lag $f(n)$ so that there is a sequence of positive integers $n_1 < n_2 < n_3 < \ldots$ such that if $n_k \leq n_k + j < n_{k+1}$, then $f(n_k + j) \geq j$, and furthermore,*

$$\liminf_{k\to\infty}\frac{n_{k+1}-n_k}{n_{k+1}} > 0$$

Then Player 2 can guarantee a liminf payoff of -1.

The conditions of this proposition can be intuitively stated as follows: Suppose the stages of the game can be partitioned into blocks $(B_k)_{k=1}^{\infty}$ of consecutive stages so that while in block B_k Player 1 gets no information about Player 2's actions within the block. Suppose also that each block consists of at least a fixed percentage of the stages in the game up to that point. Then the result follows.

As a corollary to Proposition 4.2.2, one can deduce

Corollary 4.2.3. *Suppose that in the Big Match, Player 1 is subject to an information lag $f(n)$ such that*

$$\liminf_{n\to\infty} \frac{f(n)}{n} > 0$$

Then Player 2 can guarantee a liminf payoff of -1.

The above corollary is also immediate from Proposition 4.2.5 below, so I shall not state a separate proof.

The next proposition shows that even if the lag function $f(n)$ is $o(n)$, the undiscounted value may not exist.

Proposition 4.2.4. *There exists a lag function $f(n)$ which is $o(n)$ so that if in the Big Match Player 1 is subject to this information lag, then Player 2 can guarantee a liminf payoff of -1.*

A result provided by A. Neyman (private communication) gives a more explicit result:

Proposition 4.2.5. *Suppose that in the Big Match, Player 1 is subject to an information lag $f(n)$ such that*

$$f(n) >> \frac{n}{\log(\log(n))}$$

Then Player 2 can guarantee a liminf payoff of -1.

Although Proposition 4.2.4 is an immediate corollary of Proposition 4.2.5, the construction in the proof of Proposition 4.2.4 yields a lag function with the interesting property that there is a sequence $n_1 < n_2 < n_3 < \ldots$ such that for all k, $f(n_k) = 0$; that is, infinitely often Player 1 views the entire history of the game up to that point.

Propositions 4.2.2, 4.2.4, and 4.2.5 are proved in Section 5.2.

4.3 Discussion of the Information Lag Results

Remark I. For any fixed $N \in \mathbb{N}$, we observe the stochastic game (for simplicity, with finite state and action spaces) of finite length N (the payoff, given by Definition 3.1.12, is the average of the stage-by-stage payoffs). It results by an inductive argument that the value $v_N \in \mathbb{R}^S$ for such a game exists and the players possess optimal Markov strategies [29, Section 1.4].

Therefore, in such games of finite length, a player can certainly guarantee whatever he could in the perfect monitoring game even if he monitors nothing but the current state. If the undiscounted value of a stochastic game exists, then it must hold that $v_N(z) \underset{N\to\infty}{\longrightarrow} v(z)$ for each state z. The existence of this limit, and its value in the case that it does exist, is therefore independent of any information lags. In particular, in the propositions in Section 4.2, the undiscounted value of the game (Definition 3.1.13) is not merely shifted when an information lag is imposed on one player; the undiscounted value no longer exists.

It is interesting to observe the payoff to Player 1 given by

$$\liminf_{N\to\infty} E^{z_1}_{\sigma,\tau}\big[\frac{1}{N}\sum_{j=1}^{N} g_j\big]$$

(which, by Fatou's lemma, is at least as much as the liminf-payoff) and ask what Player 1 can guarantee in such a game when subject to large information lags. This question is still open.

Remark II. The results presented here still do not pertain to certain lag functions. For instance, if a player in a stochastic game has an information lag given by $f(n) \sim \frac{n}{\log n}$, can he guarantee the same as he could have in the game with perfect monitoring? Perhaps he can; however, such a lag function does not satisfy the condition of Theorem 4.1.2:

Suppose $f(n) = \pi(n) = \#\{1 < p \leq n \mid p \textit{ is prime}\}$ is the prime-counting function. Then the prime number theorem states that $\pi(n) \sim \frac{n}{\log n}$ [7, Chapter 18]. Let $p(n)$ denote the n-th prime. Note that $\pi(p(n)) = n$ for all n.

Suppose $g(n)$ is related to $f(n)$ as in Theorem 4.1.2, and suppose λ satisfies the conditions there. Then we would have to have for any $M > 0$, and almost all n,

$$\frac{1}{\lambda(\pi(n))\cdot n} \leq \frac{1}{\lambda(M+g(n))\cdot n} \to 0$$

which tells us that

$$\frac{1}{\lambda(\pi(n))} = o(n)$$

and therefore,

$$\frac{1}{\lambda(n)} = o(p(n))$$

and therefore, for almost all n,

$$\frac{1}{p(n)} \leq \lambda(n)$$

However, the harmonic series of primes $\sum_{n=1}^{\infty} \frac{1}{p(n)}$ diverges [7, Chapter 7], a contradiction to condition (2.a) on λ of Theorem 4.1.2.

Remark 4.3.1. The same contradiction would arise for any non-decreasing $f(n) \sim \frac{n}{\log(n)}$ by taking the counting function of the set $\{k \in \mathbb{N} | f(k) \neq f(k+1)\}$; it's easier, however, to rely on the known result on the divergence of the harmonic series of primes.

Remark III. The condition $\limsup_{n\to\infty} \frac{f(n)}{n} > 0$ in itself is a weak condition. For any $\varepsilon \in (0,1)$, there is a lag function so that $\limsup_{n\to\infty} \frac{f(n)}{n} = \varepsilon$ and nevertheless Player 1 subject to this lag in the Big Match can guarantee $-\varepsilon$ in the undiscounted game. This occurs if the structure of the lag is such that there are large periods of perfect monitoring with short periods of imperfect monitoring in between.

Choose $\eta > 0$, and let σ be an η-optimal strategy in the perfect monitoring game for Player 1. Define sequences $(n_j), (m_j)$ of positive integers such that:
(i) $n_1 < m_1 < n_2 < m_2 < n_3 < \ldots$
(ii) For all j, $\frac{m_j}{n_{j+1}} \underset{j\to\infty}{\longrightarrow} 0$.
(iii) $\lim_{j\to\infty} \frac{m_j - n_j}{m_j} = \varepsilon$.

Denote $m_0 = 1$. Define the lag function so that for all j and $m_j \leq k < n_{j+1}$, $f(k) = 0$, and for all j and $n_j \leq k < m_j$, $f(k) = k - n_j + 1$. (That is, Player 1 views the entire history of the game after each stage k such that $m_j \leq k < n_{j+1}$, and gets no new information if $n_j \leq k < m_j$).

A strategy for Player 1 in the Big Match depends, at any given time, only on the previous actions of Player 2, so, for brevity, we list only those in the history of the game. Define a strategy σ' for Player 1 by:

$$\sigma'(b_1,\dots,b_k)=\begin{cases} T & \text{if } n_j<k\le m_j\\ \sigma(b_1,\dots,b_{n_1},b_{m_1+1},\dots,b_{n_2},b_{m_2+1},\dots, & \\ \quad b_{n_3},b_{m_4+1},\dots,\dots,b_{n_j},b_{m_j+1},\dots,b_k) & \text{if } m_j<k\le n_{j+1}\\ \sigma(\emptyset) & \text{if } k=0\end{cases}$$

where $(b_1,\dots,b_k)$ are the previous actions of Player 2. That is, σ' plays T when $n_j<k\le m_j$ ("playing it safe"), and otherwise plays as σ would have played against a history in which the actions at stages j for which there is i such that $n_i<j\le m_i$ have been removed. Such a strategy σ' is easily shown to yield a payoff of at least $-2\eta-\varepsilon$ in the liminf game and in long enough games of finite length (with average payoff).

Remark IV. Suppose that in the Big Match, with liminf (or limsup) payoff, Player 1 is subject to an information lag f so that there is a sequence of integers $n_1<n_2<n_3<\dots$ so that if $n_k\le n_k+j<n_{k+1}$, then $f(n_k+j)=j$. That is, Player 1 views the entire history the game after each of the stages $n_1,n_2,n_3,\dots$, and none of Player 2's actions in between. Denote $n_0=0$, and denote $m_k=n_k-n_{k-1}$. Assume Player 2 has perfect monitoring. We define an auxiliary game (denoted G) as follows:

The (countably infinite) state space S of G consists of states $(s_i)_{i\in\mathbb{N}}$. In state s_k, Player 2's action space is the set of pure behavioral strategies for Player 2 in the Big Match of finite length m_k; Player 1's action space is the set of pure behavioral strategies for Player 1 that are independent of Player 2's actions in this Big Match of finite length. (These are both finite action spaces. Note that these are only the pure strategies.) State s_k can then be viewed as a play of length m_k in the Big Match in which Player 2 has perfect monitoring and Player 1 observes none of Player 2's actions. The transition probabilities are defined so that s_i is followed by state s_{i+1}.

The Borel payoff in G is defined to be the payoff that would result in the Big Match with liminf (or limsup) payoff if the players played from stages $n_{k-1}+1$ through n_k as they did in stage k of G, assuming absorption had not occurred yet.

G is therefore a stochastic game with countable state space and finite action spaces, a bounded Borel payoff, and perfect monitoring. By Theorem 3.3.7, it has a value; therefore, in the case that Player 1 has this particular type of information lag, the Big Match has a liminf value and a limsup value.

A conjecture, based on indirect contact with E. Shmaya, is that the liminf and limsup values in any stochastic game with finite state and action spaces exist when the players are subject to arbitrary information lags in which all actions are eventually observed.

Remark V. In [21], the concepts of minmax and maxmin for multi-player undiscounted stochastic games are introduced and the existence of the minmax is demonstrated. An example provided by A. Neyman (private communication) demonstrates that if a player is subject to an information lag of even a single move, the minmax may shift. Observe, first, the following three-player stochastic game with a single state whose actions and payoffs to Player 1 are given by:

W

	L	R
T	-1	0
B	0	0

E

	L	R
T	0	0
B	0	-1

Player 1 is the box player (E, W), Player 2 is the row player (T, B), and Player 3 is the column player (L,R). With perfect monitoring, the undiscounted minmax is seen to be $-\frac{1}{4}$: A minmaxing strategy for Players 2,3 would be for each to play $(0.5, 0.5)$ at each stage. For a strategy profile (σ^2, σ^3) of Players 2,3, a (σ^2, σ^3)-maximizing strategy σ^1 for Player 1 is to observe the history h and play W if $(\sigma^2, \sigma^3)(h)(T, L) < (\sigma^2, \sigma^3)(h)(R, B)$, and E otherwise; since $\min[(\sigma^2, \sigma^3)(h)(T, L), (\sigma^2, \sigma^3)(h)(R, B)] \leq \frac{1}{4}$, this demonstrates that the undiscounted minmax is $-\frac{1}{4}$.

Now, observe the following variation of the above game:

W

	L_1	L_2	R_1	R_2
T	-1	-1	0	0
B	0	0	0	0

E

	L_1	L_2	R_1	R_2
T	0	0	0	0
B	0	0	-1	-1

If the game had perfect monitoring, then the minmax would be $-\frac{1}{4}$, as above. Now impose a one-move information lag on Player 1; that is, the lag function satisfies $f(k) = 1$ for all $k \geq 1$. I contend that the undiscounted minmax is now $-\frac{1}{2}$. For any strategy profile (σ^2, σ^3) of Players 2,3, Player 1 can guarantee $-\frac{1}{2}$ by playing $(0.5, 0.5)$. A minmaxing strategy profile (σ^2, σ^3) of Players $2, 3$ is the following: Play arbitrarily at the first stage. Thereafter, their actions are determined by what they had played at the previous stage. If R_1 or L_1 was played, Player 3 plays T and Player 2 plays $(0.5, 0.5, 0, 0)$; if R_2 or L_2 was played, Player 3 plays B and Player 2 plays $(0, 0, 0.5, 0.5)$.

In this manner, the actions of Player 2 and Player 3 are correlated. Player 1, however, doesn't see the actions played at the previous stage and so, from his point of view, (T, L_1) or (T, L_2) occurs with probability $\frac{1}{2}$, and (B, R_1) or (B, R_2) occurs with probability $\frac{1}{2}$.

It's not clear whether the minmax must exists when a bounded information lag is imposed on one player in a multi-player game.

4.4 Failure of a "Direct Approach" in Games with Information Lag

In this section I bring an example to demonstrate why near-optimal strategies in a stochastic game with perfect monitoring do not translate immediately into near-optimal strategies in games with information lag.

Let us take, for example, the Big Match, presented in Section 3.2. A near-optimal strategy for Player 1 – similar to the one derived from [16] – is given by

$$\sigma(h_n)(B) = \frac{1}{[(M + k_n + \varepsilon n)^+]^2 + 1}$$

where $M > 0$, $\varepsilon < 1$, k_n denotes the excess of L's over R's of Player 2 in the history h_n of length n in which B has not been played, and $a^+ = \max(a, 0)$. As ε is taken smaller, and M is taken larger, the strategy becomes more optimal.

Now, suppose the lag function $f(n) = \lfloor n^{0.8} \rfloor$ is imposed on Player 1. Denote $q(n) = n - f(n)$. Suppose Player 1 tries to play a strategy

$$\sigma'(h_n)(B) = \frac{1}{[(M + k_{q(n)} + \varepsilon n)^+]^2 + 1}$$

That is, suppose Player 1 modifies his strategy to accommodate the information lag (the rationale being that $f(n) = o(\varepsilon n)$). We shall now demonstrate how Player 2 can assure that with high probability that (B, L) is eventually played.

Choose $0 < \delta$ (which will be taken to be small), $0 < P$ (which will be taken to be large), and $C > 5$. Choose $M' > 0$ so that $M \leq M'$ and M' is large enough so that

$$\sum_{n=1}^{\infty} \frac{1}{(M' + n^{0.75})^2 + 1} < \delta \tag{4.4.1}$$

and choose $S \in \mathbb{N}$ large enough so that

$$\frac{S}{(M')^2 + 1} > P \tag{4.4.2}$$

Let $U \in \mathbb{N}$ be the smallest integer so that

$$M + q(U) + \varepsilon U \geq M' + C \tag{4.4.3}$$

We take integers W, K such that $K \leq f(U + W + K)$ (and other properties to be specified later), and describe τ:

Step 1: τ plays L for the first U stages.

Step 2: For the next W stages, τ plays a sequence of L's and R's such that

$$M' + (n - U)^{0.75} \leq M + k_{q(n)} + \varepsilon n$$

and so that $n \leq W + K$ large enough,

$$M + k_{q(n)} + \varepsilon n \leq M' + (n - U)^{0.75} + 2C$$

Since it never happens that $q(n) = q(n+1)$ for two consecutive integers ≥ 2, this step in the description of τ can be defined.

Step 3: For the next K stages, τ plays R.

Step 4: Thereafter, τ plays L.

We say that Player 1 plays, at stage $n + 1$, against Step j if $q(n) + 1$ is a stage at which Step j was played.

Since the growth $n^{0.8}$ of the information lag is greater than the growth of $(n - U)^{0.75}$, it's not hard to see that W, K can be chosen large enough so that Player 2 plays Step 4 while Player 1 plays against Step 2 (i.e., $K \leq f(U + W + K)$, and so Step 2 is well defined), and so that there are at least S integers n so that, at stage n, Player 1 plays against Step 3, and so that $M + k_{q(n)} + \varepsilon n \leq M'$.

Let A_i (resp., B_i) denote the set of infinite histories in which B (resp. (B, L)) is played while Player 1 is plays against Step i. Note that playing B occurs at most once, as the game moves to an absorbing state when this happens, so the (A_i) (resp., (B_i)), are mutually exclusive. Observe that $A_1 = B_1$ and $A_3 = B_3$, since Player 1 plays against Step 3 after Player 2 begins Step 4. Let $\mathbb{B} = \bigcup_{i=1}^{4} B_4$.

If there is a sequence of Bernoulli trials, with probabilities of success $p_1, \ldots, p_n$, E_i denoting the event that the first success occurs on the i-th trial, and $\mathbb{E} = \bigcup_{j=1}^{n} E_i$, then

$$P(\mathbb{E}) = \sum_{k=1}^{n} (1 - P(\bigcup_{i<k} E_i))p_i \geq \sum_{k=1}^{n} (1 - P(\mathbb{E}))p_i$$

and therefore

$$P(\mathbb{E}) \geq \frac{\sum_{k=1}^{n} p_i}{1 + \sum_{k=1}^{n} p_i}$$

Inequality (4.4.1) tells us that $P(A_2|A_1^c) < \delta$. Since there are at least S stages which occur while Step 4 is being played for which $M + k_{q(n)} + \varepsilon n \leq M'$, using inequality (4.4.2) and the above observation on sequences of trials, we derive $P(B_3|(A_1 \bigcup A_2)^c) \geq \frac{P}{1+P}$.

Therefore,

$$\begin{aligned}
P(\mathbb{B}) &\geq P(B_1) + (1 - P(A_1) - P((A_1)^c)P(A_2|(A_1)^c))P(B_3|(A_1 \bigcup A_2)^c) \\
&\geq P(B_1) + (1 - P(B_1) - P((A_1)^c)\delta)\frac{P}{1+P}
\end{aligned}$$

which tends to 1 as $P \to \infty,\ \delta \to 0$.

5 Proofs

Sections 5.1 and 5.2 pertain to the Big Match. We use the following notations and conventions:

- z_0 denotes the non-absorbing state of the Big Match.
- g_k denotes the payment of Player 2 to Player 1 at stage k.
- $Y = \liminf_{n\to\infty} \frac{g_1+\ldots+g_n}{n}$.
- $\mathbb{A} \subseteq H_\infty$ denotes the subset of infinite histories of the Big Match, beginning in state z_0, in which absorption occurs.
- $\mathbb{E} = \mathbb{A}^c$; that is, $\mathbb{E} \subseteq H_\infty$ is the subset of infinite histories of the Big Match, beginning in state z_0, in which absorption does not occur.

5.1 Proof of Proposition 4.2.1

Choose

$$\varepsilon < \frac{1}{4} \cdot \limsup_{n\to\infty} \frac{f(n)}{n}$$

Choose a countable partition of $\mathbb{N}$ (that is, a countable partition of the stages of the game), denoted $\{R, B_1, B_2, B_3, \ldots\}$, such that:

• Each B_k is a non-empty finite set of consecutive integers (that is, a set of consecutive stages).

• If $j < k$, $\max B_j < \min B_k$.

• If n_k denotes the largest element of B_k, then $\frac{f(n_k)}{n_k} \geq \frac{|B_k|}{n_k} \geq 4\varepsilon$.

The definition of ε implies that we can choose such a partition.

Let σ be any strategy for Player 1, subject to the information lag f. Let s_1, s_2 denote one-shot strategies for Player 2: $s_1 = (0.5+\varepsilon, 0.5-\varepsilon)$, $s_2 = (0,1)$. Note that the expected payoff to Player 1 at a stage in which the strategy profile (T, s_1) is played is 2ε.

Let $0 < \eta$, and take any sequence of positive real numbers $(\alpha_i)_{i=0}^{\infty}$ such that $\sum_{i=0}^{\infty} \alpha_i \leq \eta$. Define τ such:

• At each stage of R, τ plays s_1.

• For a block B_n, τ will play s_2, for each stage in that block, if τ has played s_2 in precisely k of the previous blocks and the probability that σ will play B during this block (based on the history preceding this block) is $\leq \alpha_k$; otherwise, τ plays s_1 for every stage in the block.

Note that this is well defined because anything σ plays during a block B_k is independent of the actions of Player 2 within the block.

Let $\mathbb{A}_1 \subseteq \mathbb{A}$ (resp. $\mathbb{A}_2 \subseteq \mathbb{A}$) denote the plays where absorption occurs when s_1 (resp. s_2) is played. Then $\mathbb{A} = \mathbb{A}_1 \bigcup \mathbb{A}_2$.

$E^{z_1}_{\sigma,\tau}(Y|\mathbb{A}_1) = -2\varepsilon,\ E^{z_1}_{\sigma,\tau}(Y|\mathbb{A}_2) = 1$. Furthermore, $P(\mathbb{A}_2) \leq \eta$. Therefore, since $\varepsilon \leq \frac{1}{2}$,

$$\begin{aligned} P(\mathbb{A})E^{z_1}_{\sigma,\tau}(Y|\mathbb{A}) &= P(\mathbb{A}_1)E^{z_1}_{\sigma,\tau}(Y|\mathbb{A}_1) + P(\mathbb{A}_2)E^{z_1}_{\sigma,\tau}(Y|\mathbb{A}_2) = -2\varepsilon P(\mathbb{A}_1) + P(\mathbb{A}_2) \\ &= -2\varepsilon(P(\mathbb{A}) - P(\mathbb{A}_2)) + P(\mathbb{A}_2) = -2\varepsilon P(\mathbb{A}) + P(\mathbb{A}_2)(1+2\varepsilon) \\ &\leq -2\varepsilon P(\mathbb{A}) + \eta(1+2\varepsilon) < -2\varepsilon P(\mathbb{A}) + 2\eta \qquad (5.1.1) \end{aligned}$$

(The inequality between the first and final terms still holds if $P(\mathbb{A}_1)$ or $P(\mathbb{A}_2)$ are 0; if $P(\mathbb{A}) = 0$, the calculation becomes irrelevant.)

Proposition 5.1.1. *Against almost every play in $\mathbb{E}$, τ plays s_2 infinitely many times.*

Proof. Let $\Sigma \subseteq 2^{\mathbb{N}}$ denote the collection of finite sets of positive integers. For every $I \in \Sigma$, let $\mathbb{E}_I$ denote the subset of $\mathbb{E}$ of plays against which τ plays s_2 only in blocks $(B_i)_{i\in I}$. Since Σ is countable, it's enough to verify that for all $I \in \Sigma$, $P(\mathbb{E}_I) = 0$.

Given such I, denote $m = maxI$, and $L = |I|$. Then, for every $k > m$, σ plays B with probability greater than α_L in the block B_k against any history in $\mathbb{E}_I$. Absorption then occurs almost surely, but absorption does not occur in $\mathbb{E}_I \subseteq \mathbb{E}$; therefore, $P(\mathbb{E}_I) = 0$.

□

Suppose that in some play, up to the beginning of a block B_k, the average payoff has been ρ, and suppose s_2 is played during B_k and absorption does not occur. Then, the average payoff at the end of the block is

$$\frac{\rho(n_k - |B_k|) - |B_k|}{n_k} = \rho(1 - \frac{|B_k|}{n_k}) - \frac{|B_k|}{n_k}$$

By the definition of the blocks, the strong law of large numbers, and Proposition 5.1.1,

$$\begin{aligned} E^{z_1}_{\sigma,\tau}(Y|\mathbb{E}) &\leq (1-\liminf_{k\to\infty}\frac{|B_k|}{n_k})2\varepsilon - \liminf_{k\to\infty}\frac{|B_k|}{n_k} \\ &\leq 2\varepsilon - 4\varepsilon = -2\varepsilon \end{aligned}$$

and therefore

$$\begin{aligned} E^{z_1}_{\sigma,\tau}(Y) &= P(\mathbb{E})E^{z_1}_{\sigma,\tau}(Y|\mathbb{E}) + P(\mathbb{A})E^{z_1}_{\sigma,\tau}(Y|\mathbb{A}) \leq -2\varepsilon(1-P(\mathbb{A})) - 2\varepsilon P(\mathbb{A}) + 2\eta \\ &= -2\varepsilon + 2\eta \qquad (5.1.2) \end{aligned}$$

and we could have chosen η arbitrarily small and ε arbitrarily close to $\frac{1}{4} \cdot \limsup_{n\to\infty} \frac{f(n)}{n}$.

5.2 Proof of Propositions 4.2.2, 4.2.4, and 4.2.5

The proofs of Propositions 4.2.2, 4.2.4, and 4.2.5 all involve a similar technique. The game will be partitioned into blocks $(B_k)_{k=1}^{\infty}$ comprising consecutive stages, such that Player 1 sees none of Player 2's actions within B_k while that block is being played. Player 2's strategies will be to choose $\varepsilon > 0$ small and, in each block, play only R with probability ε, and play only L with probability $1-\varepsilon$. Since Player 1 observes none of Player 2's actions in the current block, the smaller ε is, the more Player 1 is discouraged from playing B.

These proofs make use of the conventions introduced at the beginning of Section 5.

Proof of Proposition 4.2.2. Let (n_k) be a sequence as in Proposition 4.2.2, and denote $n_0 = 0$. Choose $\alpha > 0$ such that for all k,

$$\frac{n_{k+1} - n_k}{n_{k+1}} \geq \alpha$$

(Note that $\alpha < 1$.) This can be rewritten as

$$n_{k+1} \geq \frac{1}{1-\alpha} n_k \qquad (5.2.1)$$

Let B_k denote the set of stages $n_{k-1}+1$ through n_k. Choose now $\varepsilon > 0$. (5.2.1) implies that there is K so that for all m,

$$\frac{\sum_{k=1}^{m} |B_k|}{\sum_{k=1}^{m+K} |B_k|} < \varepsilon \tag{5.2.2}$$

I.e., there is K so that at any point in the game, the most recent K blocks will contain a fraction of at least $1-\varepsilon$ of the stages in the game up to that point.

Player 2's strategy τ is to play, in block B_k, R for the entire block with probability ε and L for the entire block with probability $1-\varepsilon$, independent of anything that occurred before block B_k.

If absorption does not occur, it almost surely occurs infinitely often that only R is played in K consecutive blocks. By (5.2.2),

$$E^{z_1}_{\sigma,\tau}[Y|\mathbb{E}] \leq -1(1-\varepsilon) + 1 \cdot \varepsilon = -1 + 2\varepsilon$$

since the average payoff to Player 1 in the blocks preceding any such K blocks is at most 1.

On the other hand, Player 1 does not observe whether Player 2 is currently playing L or R in the current block and Player 2's choices are independent of anything that occurred previously. As such, given that absorption occurs, the probability of absorption occurring when R is played is ε. That is,

$$E^{z_1}_{\sigma,\tau}[Y|\mathbb{A}] = -1(1-\varepsilon) + 1 \cdot \varepsilon = -1 + 2\varepsilon$$

Therefore,

$$E^{z_1}_{\sigma,\tau}[Y] \leq -1 + 2\varepsilon$$

and ε can be chosen arbitrarily small. □

Proof of Proposition 4.2.4. We will construct a sequence $n_1 < n_2 < n_3 < \cdots$, and define the information lag to be such that if $n_k \leq n_k + j < n_{k+1}$, then $f(n_k + j) = j$. For convenience, denote $n_0 = 0$; then the k-th block, B_k, comprises the stages from $n_{k-1}+1$ through stages n_k.

Once the sequence (n_k), and hence the blocks (B_k), are defined, Player 2's strategy τ will be to play, for $\varepsilon > 0$ chosen arbitrarily small, R for the entire k-th block with probability ε, and L for the entire k-th block otherwise, independent of the history preceding the k-th block.

The idea will be to mimic the exponential growth of the size of the blocks used in the proof of Proposition 4.2.2, but to carefully decrease the base of the growth.

For each $m \geq 1$, take K_m to be such that if a sequence in $\mathbb{N}$, (p_k), satisfies for all k that

$$\frac{p_{k+1} - p_k}{p_{k+1}} \geq \frac{1}{2+m} \tag{5.2.3}$$

then for all k,

$$\frac{p_k}{p_{k+K_m}} < \varepsilon \tag{5.2.4}$$

(The choice of K_m can be done in a fashion similar to the choice of K in the proof of Proposition 4.2.2). We will define a sequence $k_1 < k_2 < k_3 < \dots$ such that

$$\sum_{j=1}^{m} k_j < k \leq \sum_{j=1}^{m+1} k_j \Longrightarrow \frac{n_k - n_{k-1}}{n_k} > \frac{1}{2+m} \tag{5.2.5}$$

holds. (5.2.5) will recursively determine the (n_k): Choose n_1 arbitrarily, and given n_{k-1}, choose n_k to be the smallest integer satisfing (5.2.5). Choose k_1 arbitrarily, and choose k_{m+1} (recursively) to be even and large enough, according to the following criteria:

(i) k_{m+1} is chosen so that, if (5.2.5) holds, then the number of stages in the first $\sum_{j=1}^{m} k_j$ blocks is at most an $\frac{1}{m}$-fraction of the stages in the first $\sum_{j=1}^{m} k_j + \frac{k_{m+1}}{2}$ blocks.

(ii) The probability that R is played for K_m consecutive blocks, sometime between block $\sum_{j=1}^{m} k_j + \frac{k_{m+1}}{2} + 1$ and block $\sum_{j=1}^{m+1} k_j$, is at least $\frac{1}{2}$.

Then, given that absorption does not occur, it a.s. happens for infinitely many m that R is played for K_m consecutive blocks between block $\sum_{j=1}^{m} k_j + \frac{k_{m+1}}{2} + 1$ and block $\sum_{j=1}^{m+1} k_j$. Therefore, that this strategy yields a liminf payoff of less than $-1 + 2\varepsilon$ is shown using an argument very similar to the one towards the end of the proof of Proposition 4.2.2, together with condition (i) on the sequence (k_j). □

Proof of Proposition 4.2.5. Assume that $f(n)$ is non-decreasing (if it is not, replace it with the function $g(n) = \min_{k \geq n} f(k)$, which also satisfies the condition of the proposition).

Choose $0 < \varepsilon < \frac{1}{2}$, and $0 < \delta < (-\ln(\varepsilon))^{-1}$. Choose $C > 0$ so that $e^{C\cdot\delta} > \frac{1}{\varepsilon}$. Therefore, for m large enough,

$$(1 + \frac{C}{m})^{\delta\cdot m} > \frac{1}{\varepsilon} \tag{5.2.6}$$

Define $n_0 = 0$, and take n_1 large enough so that:
(i) $m \geq \log(\log(n_1))$ satisfies (5.2.6).
(ii) $\frac{1}{\log(\log(x))}$ is defined in $[n_1, \infty)$.
(iii) For $n \geq n_1$, $f(n) \geq \frac{C\cdot n}{\log(\log(n))}$.

Define recursively

$$n_{k+1} = f(n_k) + n_k \tag{5.2.7}$$

Note that since f is non-decreasing, if $n_k \leq n_k + j < n_{k+1}$

$$j \leq n_{k+1} - n_k = f(n_k) \leq f(n_k + j) \tag{5.2.8}$$

The k-th block, B_k, comprises the stages from $n_{k-1} + 1$ through stage n_k. Inequality (5.2.8) implies that Player 1 sees none of Player 2's actions in the block B_k while B_k is being played.

Define $p_1 = 1$, and define recursively p_k to be the smallest integer so that

$$n_{p_k} \geq \frac{1}{\varepsilon} n_{p_{k-1}} \tag{5.2.9}$$

Note that $n_k \leq 2 \cdot n_{k-1}$ for all $k > 1$ (since f is a lag function) and therefore, since p_k is the smallest integer to satisfy (5.2.9),

$$n_{p_k} \leq \frac{2}{\varepsilon} n_{p_{k-1}}$$

In particular,

$$n_{p_k} \leq (\frac{2}{\varepsilon})^{k-1} n_1$$

Player 2's strategy τ in block B_k is to play R for the entire block with probability ε, and L for the entire block with probability $1 - \varepsilon$, independent of the history preceding the k-th block. Let $Y_k \subseteq \mathbb{E}$ denote the event that R is played in the block B_k and that absorption never occurs, and let

$X_k = \bigcap_{j=p_k}^{p_{k+1}-1} Y_j$.

The events $\{X_k\}$ are independent, given that absorption never occurs. Therefore, if for any strategy σ of Player 1,

$$\sum_{k=1}^{\infty} P_{\sigma,\tau}^{z_1}(X_k|\mathbb{E}) = \infty$$

it will result from the Borel-Cantelli lemma that, given that absorption does not occur, a.s. infinitely many of the X_k occur. Hence, that this strategy yields a liminf payoff of $< -1 + 3\varepsilon$ is shown using an argument very similar to the one towards the end of the proof of Proposition 4.2.2, together with inequality (5.2.9).

Observe that, since $\varepsilon < \frac{1}{2}$ and $n_{k+1} \leq 2n_k$, from (5.2.9) we deduce that $p_{k+1} \geq p_k + 2$, and therefore,

$$\begin{aligned}
(1+\frac{C}{\log(\log(n_{p_{k+1}}))})^{\delta\cdot\log(\log(n_{p_{k+1}}))} &> \frac{1}{\varepsilon} > \frac{n_{p_{k+1}-1}}{n_{p_k}} \\
&= \prod_{j=p_k}^{p_{k+1}-2} \frac{n_{j+1}}{n_j} = \prod_{j=p_k}^{p_{k+1}-2} (1 + \frac{f(n_j)}{n_j}) \geq \prod_{j=p_k}^{p_{k+1}-2} (1 + \frac{C}{\log(\log(n_j))}) \\
&\geq \prod_{j=p_k}^{p_{k+1}-2} (1 + \frac{C}{\log(\log(n_{p_{k+1}}))}) = (1 + \frac{C}{\log(\log(n_{p_{k+1}}))})^{p_{k+1}-p_k-1} \qquad (5.2.10)
\end{aligned}$$

Therefore

$$\delta \cdot \log(\log(n_{p_{k+1}})) \geq p_{k+1} - p_k - 1$$

Therefore,

$$\begin{aligned}
P_{\sigma,\tau}^{z_1}(X_k|\mathbb{E}) &= \varepsilon^{p_{k+1}-p_k} \geq \varepsilon \cdot \varepsilon^{\delta\cdot\log(\log(n_{p_{k+1}}))} = \varepsilon \cdot e^{-\log(\frac{1}{\varepsilon})\delta\cdot\log(\log(n_{p_{k+1}}))} \\
&\geq \varepsilon \cdot e^{-\log(\frac{1}{\varepsilon})\delta\cdot\log(\log((\frac{2}{\varepsilon})^k n_1))} \sim \varepsilon \cdot e^{-\log(\frac{1}{\varepsilon})\delta\cdot\log(k)} \\
&= \varepsilon \cdot k^{-\delta\log(\frac{1}{\varepsilon})} \geq \varepsilon k^{-1} \qquad (5.2.11)
\end{aligned}$$

which completes the proof.

□

5.3 Proofs of Theorem 4.1.2

Without loss of generality, we will assume that Player 1 is subject to the information lag $f(n)$.

Choose an initial state z_1. Let $g(\cdot,\cdot,\cdot)$ be the payoff function for the game, and let $A = ||g(\cdot,\cdot,\cdot)||$. Let $(g_n), (z_n)$ denote the sequences of payoffs and states, respectively. z_n and g_n are measurable with respect to $\mathscr{H}_n$.

Choose $0 < \varepsilon < \frac{1}{6}$.

Theorem 4.1.2 refers to two types of games, called Case 1 and Case 2 for short:

Case 1: Games with finite space and actions spaces.

Case 2: Absorbing games with compact metric action spaces.

In Case 1, denote

$$v(z, \lambda) = v(\lambda)(z)$$

and

$$v(z, 0) = \lim_{\lambda \to 0} v(\lambda)(z),$$

which exists by Theorem 3.3.2.

In Case 2, we take $v(z, \lambda)$ to be independent of λ in $[0, 1)$ and define it by

$$v(z, \lambda) = \begin{cases} \lim_{\lambda \to 0} v(\lambda)(z) - \varepsilon & \text{if } z = z_0 \\ val_{A_z \times B_z}(g(z, \cdot, \cdot)) & \text{if } z \neq z_0 \end{cases}$$

which exists by Theorem 3.4.5, where z_0 denotes the non-absorbing state, and A_z, B_z denote the action spaces in state z.

We need to show that Player 1 can guarantee $v(z_1, 0)$.

Using Corollary 3.4.6 for the Case 2-type games, together with with the characterization of the λ-discounted value given in Theorem 3.3.2 for Case 1-type games, we derive that for λ small enough,

$$\Phi(\lambda, v(z, \lambda)) \geq v(z, \lambda) \tag{5.3.1}$$

In each case, choose a measurable assignment that assigns to each $\lambda < \lambda_0$, for appropriately small λ_0, a stationary strategy $(x_z)_{z\in S}$ for Player 1 such that for every stationary strategy $(y_z)_{z\in S}$ of Player 2 and every state $z \in S$

$$E_{x_z,y_z,q}[\lambda g(z,a,b) + (1-\lambda)v(z',\lambda)] \geq v(z,\lambda) \tag{5.3.2}$$

In both cases, this can be done by inequality (5.3.1): in Case 1, this can be done in a measurable fashion (for small enough λ)[12] by Theorem 3.3.2; in Case 2, this can be done in a measurable fashion by Corollary 3.4.8 and Remark 3.4.9.

Have λ_0 be taken small enough to also assure that $v(z,\lambda)$ is monotonic, for each z, in $(0,\lambda_0)$, and so that for $\lambda < \lambda_0$ and all z,

$$|v(z,0) - v(z,\lambda)| \leq \frac{\varepsilon}{12} \tag{5.3.3}$$

(In Case 1, such a choice of λ_0 can be made by Theorem 3.3.2; in Case 2 it follows from the independence of $v(z,\lambda)$ on λ.)

Suppose $(\mu_n)_{n=1}^{\infty}$ is a sequence of measurable functions such that $\mu_n : (H_n, \mathscr{H}_n) \to (0,\lambda_0)$. We can define a behavioral strategy σ for Player 1 that at the n-th stage plays a mixed action x_z so that for every mixed action y_z, (5.3.2) holds with the discount factor μ_n. In other words, for every behavioral strategy τ of Player 2,

$$E^{z_1}_{\sigma,\tau}[v(z_{n+1},\mu_n) - v(z_n,\mu_n) + \lambda_n(g_n - v(z_{n+1},\mu_n))|\mathscr{H}_n] \geq 0 \tag{5.3.4}$$

In the next two sections we demonstrate ways to define $(\mu_n)_{n=1}^{\infty}$ so that σ is subject to the information lag f, and prove that these choices generate the required near-optimal strategies.

5.4 First Proof of Theorem 4.1.2

This proof is based on [16, Section 2], and assumes that the conditions of Theorem 4.1.2 with the pair (c.I) and (d.I) hold.

Assume without loss of generality that $A \leq \frac{1}{6}$. Choose g, λ satisfying the conditions of Theorem 4.1.2, together with the pair of conditions (c.I) and

[12] It can actually be done in a measurable (semi-algebraic) fashion in all of $(0,1)$; see [20, Theorem 4].

(d.I). Recall that $0 < \varepsilon < \frac{1}{6}$.

Choose $M > 1$ large enough so that the following assumptions on λ hold:

i) For $s \geq M - 1$, $\lambda(s) < \lambda_0$.

ii)

$$\int_M^\infty \lambda(s)ds < \varepsilon \tag{5.4.1}$$

(This can be done by condition (2.a).)

iii) For $s \geq M + g(k)$ and $|\theta| \leq 1$,

$$\left|\frac{\lambda(s + \theta \cdot \max(1, f(k)))}{\lambda(s)} - 1\right| \leq \varepsilon. \tag{5.4.2}$$

(This can be done by condition (2.c.I) and the monoticity of λ.)

iv) In Case 1, for all $s > M$ and $|\theta| \leq 1$,

$$||v(\lambda(s + \theta)) - v(\lambda(s))|| \leq \varepsilon \cdot \lambda(s) \tag{5.4.3}$$

(This can be done by condition (2.d.I), the monoticity of $v(z, \lambda)$ in $(0, \lambda_0)$, and by the fact that, from (i), $\lambda(s \pm \theta) < \lambda_0$.)

Define functions $h(k) = M + g(k)$, $q(k) = k - f(k)$.

Define $s_1 = h(1) \geq M$, $\lambda_k = \lambda(s_k)$, and

$$s_{k+1} = Max[h(k + 1), s_k + g_k - v(z_{k+1}, \lambda_{q(k)}) + 4\varepsilon] \tag{5.4.4}$$

σ denotes the behavioral strategy for Player 1 defined in Section 5.3 with respect to the sequence $(\mu_n) = (\lambda_{q(n)})$; that is, for any strategy τ of Player 2,

$$E^{z_1}_{\sigma,\tau}[v(z_{n+1}, \lambda_{q(n)}) - v(z_n, \lambda_{q(n)}) + \lambda_{q(n)}(g_n - v(z_{n+1}, \lambda_{q(n)}))|\mathscr{H}_n] \geq 0 \tag{5.4.5}$$

Note that as Player 1 is about to play stage $k + 1$, he knows $g_1, \ldots, g_{q(k)}$ and $z_1, \ldots, z_{k+1}$, and therefore knows $s_1, \ldots, s_{q(k)+1}$, and in particular, knows $s_{q(k+1)}$ (since $q(k) \leq q(k + 1) \leq q(k) + 1$). Therefore, this strategy is subject to the information lag $f(n)$.

Note that

$$|s_{k+1} - s_k| \leq 1 \tag{5.4.6}$$

since g, being a lag function, satisfies $g(y) - g(x) \leq y - x$ for $y > x$, and therefore

$$|s_{q(k)} - s_k| \leq f(k) \tag{5.4.7}$$

Also note that

$$\begin{aligned} g_k - v(z_{k+1}, \lambda_{q(k)}) + 4\varepsilon &\leq s_{k+1} - s_k \\ &\leq g_k - v(z_{k+1}, \lambda_{q(k)}) + 4\varepsilon + 2I(s_{k+1} = h(k+1)) \end{aligned} \tag{5.4.8}$$

Observe that by (5.4.2) and (5.4.6),

$$|\lambda_{k+1} - \lambda_k| \leq \varepsilon\lambda_k$$

and, by (5.4.7) and 5.4.2, for the appropriate $|\theta| \leq 1$,

$$|\frac{\lambda_{q(k)}}{\lambda_k} - 1| = |\frac{\lambda(s_{q(k)})}{\lambda(s_k)} - 1| = |\frac{\lambda(s_k + \theta f(k)))}{\lambda(s_k)} - 1| \leq \varepsilon \tag{5.4.9}$$

as $s_k \geq h(k)$. Also, in Case 1, by (5.4.3) and (5.4.9),

$$||v(\lambda_{q(k+1)}) - v(\lambda_{q(k)})|| \leq \varepsilon\lambda_{q(k)} \leq \frac{3}{2}\varepsilon\lambda_k \tag{5.4.10}$$

Therefore, by (5.4.8),(5.4.5) and (5.4.10), together with the independence of $v(z, \lambda)$ on λ in Case 2,

$$E^{z_1}_{\sigma,\tau}(v(z_{k+1}, \lambda_{q(k+1)}) - v(z_k, \lambda_{q(k)}) + \lambda_{q(k)}(s_{k+1} - s_k)|\mathscr{H}_k) \geq \frac{5}{2}\varepsilon\lambda_k \tag{5.4.11}$$

Define $t(y) = \int_y^\infty \lambda(s)ds$, $t_k = t(s_k)$. By (5.4.1), $t(y)$ is well defined and $t_k \leq \varepsilon$. By (5.4.8), (5.4.9) and (5.4.6), together with the monoticity of λ, we get

$$t_k - t_{k+1} \geq \lambda_k(s_{k+1} - s_k) - \varepsilon\lambda_k \geq \lambda_{q(k)}(s_{k+1} - s_k) - 2\varepsilon\lambda_k$$

Define $Y_k = v(z_k, \lambda_{q(k)}) - t_k$. (5.4.11) then implies

$$E^{z_1}_{\sigma,\tau}[Y_{k+1} - Y_k | \mathscr{H}_k] \geq \frac{1}{2}\varepsilon\lambda_k \tag{5.4.12}$$

Therefore, $(Y_k)_{k=1}^{\infty}$ is a submartingale with respect to $(\mathscr{H}_k)_{k=1}^{\infty}$, and is bounded by $\frac{1}{6} + \varepsilon$, and therefore converges a.s. to a limit, denoted Y_∞, such that (by (5.3.3))

$$E^{z_1}_{\sigma,\tau}[Y_\infty] \geq E^{z_1}_{\sigma,\tau}[Y_1] \geq v(z_1, \lambda_1) - t_1 \geq v(z_1, 0) - 2\varepsilon \tag{5.4.13}$$

Furthermore, for all k,

$$0 \leq \frac{1}{2}\varepsilon E^{z_1}_{\sigma,\tau}[\sum_{j=1}^{k} \lambda_j] \leq E^{z_1}_{\sigma,\tau}[Y_k - Y_1] \leq 2t(M) + 2\frac{1}{6} \leq 2\varepsilon + \frac{1}{3} \leq 1 \tag{5.4.14}$$

and so, by Lebesgue's monotone convergence theorem,

$$E^{z_1}_{\sigma,\tau}(\sum_{k=1}^{\infty} \lambda_k) \leq 2\varepsilon^{-1} \tag{5.4.15}$$

In particular, $\lambda_k \to 0$ a.s., and therefore $s_k \to \infty$ a.s., and therefore $t_k \to 0$ a.s.. Therefore, $v(z_k, \lambda_{q(k)}) \to Y_\infty$ a.s., and by using (5.4.10),

$$v(z_{k+1}, \lambda_{q(k)}) \overset{a.s.}{\to} Y_\infty \tag{5.4.16}$$

Also observe that by (5.4.14) and (5.4.1), for every k

$$E^{z_1}_{\sigma,\tau}[v(z_{k+1}, \lambda_{q(k+1)})] \geq E^{z_1}_{\sigma,\tau}[t_{k+1}] + v(z_1, \lambda_1) - t_1 \geq v(z_1, \lambda_1) - \varepsilon$$

and using (5.3.3) and (5.4.10),

$$E^{z_1}_{\sigma,\tau}(v(z_{k+1}, \lambda_{q(k)})) \geq v(z_1, 0) - 3\varepsilon \tag{5.4.17}$$

Denoting $P = \{k | s_{k+1} = h(k+1)\}$ and using (5.4.8),

$$\sum_{k \leq n} g_k \geq \sum_{k \leq n} v(z_{k+1}, \lambda_{q(k)}) + s_{n+1} - s_1 - 4n\varepsilon - 2\sum_{k \leq n} I(k \in P) \tag{5.4.18}$$

Since $\frac{1}{\lambda}$ and h are non-decreasing functions, and if $k \in P$ then $\frac{1}{\lambda_{k+1}} = \frac{1}{\lambda(h(k+1))}$,

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n} I(k\in P) &= \frac{1}{n}\sum_{k=1}^{n} I(k\in P)\lambda_{k+1}\frac{1}{\lambda_{k+1}}\\ &\le \frac{1}{n}\sum_{k=1}^{n} I(k\in P)\lambda_{k+1}\frac{1}{\lambda(h(k+1))}\\ &\le \frac{1}{n\lambda(h(n+1))}\sum_{k=1}^{n} I(k\in P)\lambda_{k+1} \qquad (5.4.19)\end{aligned}$$

We deduce from (5.4.14) that for all n,

$$E^{z_1}_{\sigma,\tau}(\sum_{k=1}^{n} I(k\in P)\lambda_{k+1}) \le 2\varepsilon^{-1} \qquad (5.4.20)$$

and therefore, combining (5.4.19) and (5.4.20),

$$E^{z_1}_{\sigma,\tau}\Big[\frac{1}{n}\sum_{k=1}^{n} I(k\in P)\Big] \le \frac{2\varepsilon^{-1}}{n\lambda(M+g(n+1))}$$

Therefore, (5.4.18) implies that

$$E^{z_1}_{\sigma,\tau}[\frac{1}{n}\sum_{k\le n} g_k] \ge E^{z_1}_{\sigma,\tau}[\frac{1}{n}\sum_{k\le n} v(z_{k+1},\lambda_{q(k)})] - \frac{1}{n}s_1 - 4\varepsilon - \frac{4\varepsilon^{-1}}{n\lambda(M+g(n+1))}$$

Applying (5.4.17) yields, finally,

$$E^{z_1}_{\sigma,\tau}[\frac{1}{n}\sum_{k\le n} g_k] \ge v(z_1,0) - 3\varepsilon - \frac{1}{n}s_1 - 4\varepsilon - \frac{4\varepsilon^{-1}}{n\lambda(M+g(n+1))}$$

And using condition (2.b) of Theorem 4.1.2 gives the required near optimality in long enough games with average payoff (note that the right-hand side of this last inequality is independent of the strategy τ of Player 2.)

Since (5.4.15) implies that a.s. $\sum_{k=1}^{\infty}\lambda_k < \infty$, using (5.4.19) and condition (2.b) of the theorem, we obtain,

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n} I(k\in P) &\le \frac{1}{n\lambda(h(n+1))}\sum_{k=1}^{n} I(k\in P)\lambda_{k+1}\\ &\le \frac{1}{n\lambda(M+g(n+1))}\sum_{k=1}^{\infty}\lambda_k \overset{a.s.}{\to} 0\end{aligned}$$

and therefore, (5.4.18) implies that

$$E^{z_1}_{\sigma,\tau}[\liminf_{n\to\infty}\frac{1}{n}\sum_{k\le n} g_k] \ge E^{z_1}_{\sigma,\tau}[\liminf_{n\to\infty}\frac{1}{n}\sum_{k\le n} v(z_{k+1},\lambda_{q(k)})] - 4\varepsilon$$

Applying (5.4.13) and (5.4.16) gives

$$E^{z_1}_{\sigma,\tau}[\liminf_{n\to\infty}\frac{1}{n}\sum_{k<n} g_k] \ge v(z_1,0) - 6\varepsilon$$

and the proof is complete.

5.5 Second Proof of Theorem 4.1.2

This proof is based on [16, Section 3], and assumes that the conditions of Theorem 4.1.2 with the pair (c.II) and (d.II) hold.

.

Assume without loss of generality that $A \le \frac{1}{4}$. Choose g, λ satisfying the conditions of Theorem 4.1.2, together with the pair of conditions (c.II),(d.II). Recall that $0 < \varepsilon < \frac{1}{6}$.

Set $\delta = \frac{\varepsilon}{12}$. Choose $0 < D < \frac{1}{4}$ small enough, and $M > \frac{8}{D}$ large enough, so that the following assumptions on λ hold:

i) For $s \ge \frac{M}{1-D}$, $\lambda(s) < \lambda_0$.

ii)

$$\int_M^\infty \lambda(s)ds < \delta \tag{5.5.1}$$

(This can be done by (2.a).)

iii) For $s \ge M$ and $|\theta| \le 1$

$$\Big|\frac{\lambda(s(1+\theta D))}{\lambda(s)} - 1\Big| \le \delta. \tag{5.5.2}$$

(This can be done by condition (2.c.II) and the monoticity of λ.)

iv) In Case 1, for all $s > M$ and $|\theta| \le 1$,

$$||v(\lambda(s(1+\theta D))) - v(\lambda(s))|| \le \delta \cdot \lambda(s) \tag{5.5.3}$$

(This can be done by condition (2.d.II), the monoticity $v(z, \lambda)$ in $(0, \lambda_0)$, and by the fact that, from (i), $\lambda(s(1+\theta D)) < \lambda_0$.)

v) For $s \geq M$,

$$D \cdot s\lambda(s) < \delta \tag{5.5.4}$$

(This results from condition (2.a) and the monoticity of λ.)

(We will enlarge M beyond these requirements.)

Define functions $h(k) = M + g(k)$, $L(s) = \lfloor Ds \rfloor$.

Define $s_0 = h(1) \geq M$, $B_0 = 1$, $L_k = L(s_k)$, $B_{k+1} = B_k + L_k$, $\lambda_k = \lambda(s_k)$, and

$$s_{k+1} = Max[h(B_{k+1}), s_k + \sum_{B_k \leq i < B_{k+1}} (g_i - v(z_{B_{k+1}}, 0) + \varepsilon)] \tag{5.5.5}$$

Note that

$$|s_{k+1} - s_k| \leq \max(L_k, g(B_{k+1}) - g(B_k)) = L_k \tag{5.5.6}$$

since g, being a lag function, satisfies $g(y) - g(x) \leq y - x$ for $y > x$ and therefore $|g(B_{k+1}) - g(B_k)| \leq L_k$.

Denote $\mathscr{G}_k = \mathscr{H}_{B_k}$ and define the $(\mathscr{G}_k)_{k=0}^{\infty}$ stopping times by

$$k(i) = \inf\{k | B_k > i\} \tag{5.5.7}$$

Using the facts that $M > \frac{8}{D}$ (and therefore $L_k \geq 8$) and $D \leq \frac{1}{4}$, we see that for all $n \geq 1$

$$\begin{aligned} |\frac{L_n}{L_{n-1}} - 1| &\leq \frac{1}{L_{n-1}} |L_n - L_{n-1}| \leq \frac{1}{L_{n-1}} (D|s_n - s_{n-1}| + 2) \\ &\leq \frac{1}{L_{n-1}} (\frac{1}{4} L_{n-1} + 2) \leq \frac{1}{L_{n-1}} (\frac{1}{2} L_{n-1}) = \frac{1}{2} \end{aligned} \tag{5.5.8}$$

Observe that

$$Dg(B_n) \leq Ds_n \leq L_n + 1 \leq 2L_n \leq 4L_{n-1}$$

and M can be taken large enough so that for $k \geq B_1$,

$$f(k) \leq \frac{1}{32} Dg(k)$$

and using g's monoticity we deduce that for $k \geq 1$ and $B_k \leq n < B_{k+1}$, $f(n) \leq \frac{1}{32} Dg(B_{k+1})$ and therefore,

$$f(n) \leq \frac{1}{4} L_{k-1},\ f(n) \leq \frac{1}{4} L_k \tag{5.5.9}$$

Accordingly, at stage B_k, Player 1 will have observed all the stages except at most $\frac{1}{4}$ of the previous block.

Denote $s'_0 = s_0$ and for $k \geq 1$,

$$s'_k = Max[h(B_k), s_{k-1} + \sum_{B_{k-1} \leq i < B_k - f(B_k)} (g_i - v(z_{B_{k+1}}, 0) + \varepsilon)] \tag{5.5.10}$$

And therefore, from (5.5.9),

$$|s'_k - s_k| \leq \frac{1}{4} L_k \leq L_k \leq D \cdot s_k \tag{5.5.11}$$

and therefore, taking $\lambda'_k = \lambda(s'_k)$ and using (5.5.2) yields

$$|\lambda'_k - \lambda_k| \leq \delta \lambda_k \tag{5.5.12}$$

Define $\hat{\lambda}_i = \lambda'_{k(i)-1}$, and let σ be the strategy as defined in Section 5.3 with respect to the sequence $(\mu_i)_{i=1}^{\infty} = (\hat{\lambda}_i)_{i=1}^{\infty}$. Explicitly, σ is such that for any k and $0 \leq i < L_k$ and any strategy τ for Player 2,

$$E^{z_1}_{\sigma,\tau}(\hat{\lambda}_j g_j + (1 - \hat{\lambda}_j) v(z_{j+1}, \hat{\lambda}_j) | \mathscr{H}_j) \geq v(z_j, \hat{\lambda}_j) \tag{5.5.13}$$

In other words, the strategy σ plays, from B_k through $B_{k+1} - 1$, is a strategy that is optimal in the λ'_k-discounted game.

We note that by (5.5.9), σ is subject to the information lag $f(n)$. At move B_k, Player 1 is aware of the value of L_{k-1} and is aware of s'_k, and knows that $L_k \geq \lfloor \frac{1}{2} L_k \rfloor$; by move $B_k + \frac{1}{2} L_k$, he is already aware of the value of s_k and therefore of the value of B_{k+1}.

From (5.5.13), taking the expected value with respect to $\mathscr{G}_k$ for $B_k \leq j = B_k + i < B_{k+1}$ for $0 \leq i < L_k$, we get

$$E^{z_1}_{\sigma,\tau}(\lambda'_k g_{B_k+i} + (1 - \lambda'_k) v(z_{B_k+i+1}, \lambda'_k) | \mathscr{G}_k) \geq E[v(z_{B_k+i}, \lambda'_k) | \mathscr{G}_k]$$

Multiplying by $(1-\lambda'_k)^i$ and summing gives

$$E^{z_1}_{\sigma,\tau}(\lambda'_k \sum_{0\le i<L_k} (1-\lambda'_k)^i g_{B_k+i} + (1-\lambda'_k)^{L_k} v(z_{B_{k+1}},\lambda'_k) - v(z_{B_k},\lambda'_k)|\mathscr{G}_k) \ge 0 \tag{5.5.14}$$

Now, since (5.5.12) and $L_k\lambda_k < \delta$ (condition (v)),

$$\begin{aligned}
&\left|\lambda_k \sum_{0\le i<L_k} (1-\lambda_k)^i g_{B_k+i} - \lambda'_k \sum_{0\le i<L_k} (1-\lambda'_k)^i g_{B_k+i}\right| \\
&= \left|\lambda_k \sum_{0\le i<L_k} [(1-\lambda_k)^i - (1-\lambda'_k)^i] g_{B_k+i} + (\lambda_k-\lambda'_k) \sum_{0\le i<L_k} (1-\lambda'_k)^i g_{B_k+i}\right| \\
&\le \lambda_k \sum_{0\le i<L_k} |\lambda_k-\lambda'_k| L_k + |\lambda_k-\lambda'_k| L_k \le \lambda_k(\lambda_k L_k)L_k + \delta\lambda_k L_k \le 2\delta\lambda_k L_k
\end{aligned} \tag{5.5.15}$$

Also observe that

$$|v(z,\lambda'_k)-v(z,\lambda_k)| = |v(z,\lambda(s_k+\frac{s'_k-s_k}{L_k}L_k))-v(z,\lambda(s_k))| \le \delta L_k\lambda_k \tag{5.5.16}$$

(In Case 1, this results from (5.5.3); in Case 2, it is immediate.) Therefore,

$$\begin{aligned}
&|(1-\lambda_k)^{L_k} v(z,\lambda_k) - (1-\lambda'_k)^{L_k} v(z,\lambda'_k)| \\
&= \left|(1-\lambda_k)^{L_k}(v(z,\lambda_k)-v(z,\lambda'_k)) + v(z,\lambda'_k)((1-\lambda_k)^{L_k} - (1-\lambda'_k)^{L_k})\right| \\
&\le \delta L_k\lambda_k + L_k|\lambda_k-\lambda'_k| \le 2\delta\lambda_k L_k
\end{aligned} \tag{5.5.17}$$

Using (5.5.15), (5.5.16) and (5.5.17) on (5.5.14), we derive

$$E^{z_1}_{\sigma,\tau}(\lambda_k \sum_{0\le i<L_k} (1-\lambda_k)^i g_{B_k+i} + (1-\lambda_k)^{L_k} v(z_{B_{k+1}},\lambda_k) - v(z_{B_k},\lambda_k)|\mathscr{G}_k) \ge -5\delta L_k\lambda_k \tag{5.5.18}$$

Denote $l_k = v(z_{B_k},\lambda_k)$. By (5.4.3),

$$|v(z_{B_{k+1}},\lambda_k) - l_{k+1}| \le \delta L_k\lambda_k \tag{5.5.19}$$

Lemma 5.5.1. $E^{z_1}_{\sigma,\tau}(l_{k+1} - l_k + \lambda_k(s_{k+1}-s_k)|\mathscr{G}_k) \ge 3\delta L_k\lambda_k$

Proof. In (5.5.18), we apply

$$1 - \lambda \sum_{0 \le i < L} (1-\lambda)^i = (1-\lambda)^L$$

to derive

$$E^{z_1}_{\sigma,\tau}(v(z_{B_{k+1}}, \lambda_k) - l_k + \lambda_k \sum_{0 \le i < L_k} (1-\lambda_k)^i (g_{B_k+i} - v(z_{B_{k+1}}, \lambda_k)) | \mathscr{G}_k) \ge -5\delta L_k \lambda_k$$

Since $1 - \lambda_k L_k \le (1-\lambda_k)^i \le 1$ and $L_k \lambda_k < \delta$,

$$|\lambda_k \sum_{0 \le i < L_k} (1-\lambda_k)^i (g_{B_k+i} - v(z_{B_{k+1}}, \lambda_k)) - \lambda_k \sum_{0 \le i < L_k} (g_{B_k+i} - v(z_{B_{k+1}}, \lambda_k))|$$
$$\le 2\lambda_k \cdot | \sum_{0 \le i < L_k} ((1-\lambda_k)^i - 1)| \le 2\lambda_k L_k (\lambda_k L_k) \le 2\delta \lambda_k L_k$$

Putting these last two inequalities together with (5.5.19) and (5.3.3), and recalling that $\delta = \frac{\varepsilon}{12}$, we obtain

$$E^{z_1}_{\sigma,\tau}(l_{k+1} - l_k + \lambda_k \sum_{0 \le i < L_k} (g_{B_k+i} - v(z_{B_{k+1}}, 0)) | \mathscr{G}_k) \ge -5\delta L_k \lambda_k - 4\delta L_k \lambda_k$$

and by $s_{k+1} - s_k \ge \sum_{0 \le i < L_k} (g_{B_k+i} - v(z_{B_{k+1}}, 0) + 12\delta)$, we get the required. □

Define

$$t(s) = \int_s^\infty \lambda(x) dx$$

and $t_k = t(s)$. By (5.5.1), $t_k \le \delta$.

Lemma 5.5.2. $E^{z_1}_{\sigma,\tau}(l_{k+1} - t_{k+1} - (l_k - t_k) | \mathscr{G}_k) \ge 2\delta L_k \lambda_k$

Proof. Using inequalities (5.5.6) and (5.5.2) together with the monoticity of λ we get

$$t_{k+1} - t_k = \int_{s_{k+1}}^{s_k} \lambda(s) ds \le \lambda_k (s_k - s_{k+1}) + \delta L_k \lambda_k$$

and using Lemma 5.5.1 completes the proof. □

Denote $\overline{\lambda}_i = \lambda_{k(i)-1}$ and $\overline{l_i} = v(z_{B_{k(i)}}, 0)$.

Proposition 5.5.3. *a) l_k converges almost surely (denote the limit l_∞) and s_k converges almost surely to ∞.*

b) For any $(\mathscr{G}_k)_{k=0}^\infty$ stopping time T,

$$E_{\sigma,\tau}^{z_1}[l_T|\mathscr{G}_0] \geq l_0 - t_0 \geq l_0 - \delta$$

c) $E_{\sigma,\tau}^{z_1}[\sum_{i=0}^\infty (L_i\lambda_i)] = E_{\sigma,\tau}^{z_1}[\sum_{i=0}^\infty \overline{\lambda}_i] < \delta^{-1}$
d) $\overline{l}_i$ converges to $\overline{l}_\infty = l_\infty$.

e) For all $i = 1, 2, \ldots, \infty$, $E_{\sigma,\tau}^{z_1}[\overline{l_i}|G_0] \geq \overline{l_0} - 3\delta = v_\infty(z_1) - 3\delta$

Proof. Denote $Y_k = l_k - t_k$; then Lemma 5.5.2 implies that $(Y_k)_{k=0}^\infty$ is a submartingale with respect to $(\mathscr{G}_k)_{k=0}^\infty$, and therefore converges a.s. to Y_∞. Since $|Y_k| \leq \frac{1}{4} + \varepsilon \leq \frac{1}{2}$,

$$0 \leq 2\delta E_{\sigma,\tau}^{z_1}[\sum_{k<B_n} \overline{\lambda}_k] = 2\delta E_{\sigma,\tau}^{z_1}[\sum_{k<n} \lambda_k L_k] \leq E_{\sigma,\tau}^{z_1}[Y_n - Y_1] \leq 1 \qquad (5.5.20)$$

(c) results from (5.5.20), using the monotone convergence theorem.

Therefore, $\lambda_k \to 0$ a.s. and $t_k \to 0$; therefore $l_k \to Y_\infty$ a.s. and $l_\infty = Y_\infty$, so (a) is proved.

(b) is an immediate corollary of the stopping theorem for bounded submartingales:

$$E_{\sigma,\tau}^{z_1}[l_T|\mathscr{G}_0] = E_{\sigma,\tau}^{z_1}[Y_T + t_T|\mathscr{G}_0] \geq E_{\sigma,\tau}^{z_1}[Y_T|\mathscr{G}_0] \geq Y_0 = l_0 - t_0$$

(d) results from the uniform convergence of $v(z,\lambda) \to v(z,0)$, and (e) results from (b) and (5.3.3).

□

Denote $P = \{k | s_k = h(B_k)\}$.

Lemma 5.5.4.

$$E_{\sigma,\tau}^{z_1}[\frac{1}{N}\sum_{k=1}^{k(N)-1} I(k \in P)L_k] \leq \delta^{-1}\frac{1}{\lambda(h(N))N} \underset{N\to\infty}{\to} 0 \qquad (5.5.21)$$

and almost surely

$$\frac{1}{N}\sum_{k=1}^{k(N)-1} I(k\in P)L_k \underset{N\to\infty}{\to} 0 \tag{5.5.22}$$

Proof.

$$\begin{aligned}\frac{1}{N}\sum_{k=1}^{k(N)-1} I(k\in P)L_k &= \frac{1}{N}\sum_{k=1}^{k(N)-1} I(k\in P)L_k\lambda_k\frac{1}{\lambda_k}\\ &= \frac{1}{N}\sum_{k=1}^{k(N)-1} I(k\in P)L_k\lambda_k\frac{1}{\lambda(h(B_k))}\\ &\le \frac{1}{\lambda(h(N))}\frac{1}{N}\sum_{k=1}^{k(N)-1} I(k\in P)L_k\lambda_k \end{aligned}\tag{5.5.23}$$

since $\frac{1}{\lambda}$ and h are non-decreasing, and $B_{k(N)-1} \le N$. From Proposition (5.5.3),

$$E^{z_1}_{\sigma,\tau}[\sum_{k=1}^{\infty} I(k\in P)L_k\lambda_k] \le E^{z_1}_{\sigma,\tau}[\sum_{k=1}^{\infty} L_k\lambda_k] \le \delta^{-1} \tag{5.5.24}$$

Therefore, using condition (2.b) of the theorem and (5.5.23),

$$E^{z_1}_{\sigma,\tau}[\frac{1}{N}\sum_{k=1}^{k(N)-1} I(k\in P)L_k] \le \frac{1}{\lambda(h(N))\cdot N}\delta^{-1} \underset{N\to\infty}{\to} 0$$

That gives (5.5.21). Similarly, since we get from inequality (5.5.24) that a.s.

$$\sum_{k=1}^{\infty} I(k\in P)L_k\lambda_k < \infty$$

we again use (5.5.23) to derive that a.s.

$$\begin{aligned}\frac{1}{N}\sum_{k=1}^{k(N)-1} I(k\in P)L_k &\underset{N\to\infty}{\le} \frac{1}{\lambda(h(N))}\frac{1}{N}\sum_{k=1}^{k(N)-1} I(k\in P)L_k\lambda_k\\ &\le \frac{1}{\lambda(h(N))}\frac{1}{N}\sum_{k=1}^{\infty} I(k\in P)L_k\lambda_k \underset{N\to\infty}{\to} 0\end{aligned}\tag{5.5.25}$$

□

Lemma 5.5.5.

$$\sum_{1}^{n} g_i \geq \sum_{1}^{n} \bar{l}_i - 2s_0 - 8\delta n - 4\delta \sum_{k=0}^{k(n)} I(k+1 \in P)L_{k+1}$$

Proof. Using (5.5.5) and (5.5.8), we get

$$\begin{aligned} s_{k+1} - s_k &\leq \sum_{B_k \leq i < B_{k+1}} (g_i - \bar{l}_i) + 6\delta L_k + I(k+1 \in P)2L_k \\ &\leq \sum_{B_k \leq i < B_{k+1}} (g_i - \bar{l}_i) + 6\delta L_k + I(k+1 \in P)4L_{k+1} \end{aligned}$$

and summing, we get

$$s_k - s_0 \leq \sum_{i < B_k} (g_i - \bar{l}_i) + 6\delta B_k + 4\delta \sum_{j=0}^{k} I(j+1 \in P)L_{j+1}$$

Therefore,

$$\begin{aligned} \sum_{i \leq n} g_i &\geq \sum_{i \leq n} \bar{l}_i - 6\delta B_{k(n)} + s_{k(n)} - s_0 - (B_{k(n)} - n) - 4\delta \sum_{j=0}^{k(n)} I(j+1 \in P)L_{j+1} \\ &\geq \sum_{i \leq n} \bar{l}_i - s_0 - 6\delta n - 2(B_{k(n)} - n) - 4\delta \sum_{j=0}^{k(n)} I(j+1 \in P)L_{j+1} \end{aligned} \tag{5.5.26}$$

Therefore

$$B_{k(n)} - n \leq L(s_{k(n)-1}) \leq \delta s_{k(n)-1} \leq \delta(s_0 + n) \tag{5.5.27}$$

completes the proof of the lemma.

□

Therefore, since

$$\sum_{j=0}^{k(n)} I(j+1 \in P)L_{j+1} \leq \sum_{j=0}^{k(n)-1} I(j \in P)L_j + L_{k(n)+1} \leq \sum_{j=0}^{k(n)-1} I(j \in P)L_j + 3L_{k(n)}$$

and

$$\frac{L_{k(n)}}{n} \leq \frac{L(s_{k(n)})}{n} \leq \frac{L(s_0 + n)}{n} \leq \delta \frac{s_0 + n}{n}$$

we get that bounding the terms of Lemma 5.5.5 with the results of Proposition 5.5.3 and of Lemma 5.5.4 gives the required near-optimality of σ, similar to the proof in Section 5.5.

5.6 Proof of Theorem 4.1.1

Choose $\beta > 1$ as in Theorem 4.1.1, and choose $\gamma \in (1, \beta)$. Choose λ so that $\lambda(s) = \frac{1}{s(\ln(s))^\gamma}$ for large s.[13]

Observe that $f(n)(\ln(f(n)))^\gamma = o(n)$. Choose $g(n)$ to be a non-decreasing lag function so that $f(n) << g(n)$, $g(n)(\ln(g(n)))^\gamma = o(n)$ (for example, denote $\hat{f}(k) = \max_{j \leq k} f(j)$, and choose $g(n)$ so that $g(n)(\ln(g(n)))^\gamma \sim \sqrt{\max\{\hat{f}(n)(\ln(\hat{f}(n)))^\gamma, 1\} \cdot n}$); therefore, for any $M > 0$,

$$\lambda(M + g(n))n \sim \lambda(g(n))n \sim \frac{n}{g(n)(\ln(g(n)))^\gamma} \to \infty$$

Accordingly, conditions (1), (2.a), and (2.b) of Theorem 4.1.2 hold. (2.c.I) and (2.c.II) are easily seen to hold. I will demonstrate that in the case of games with finite state and action spaces, (2.d.I) and (2.d.II) hold.

From Theorem 3.3.2, there is $r \in [0, 1)$ such that for small enough λ,

$$|\frac{dv(z, \lambda)}{d\lambda}| \leq B\lambda^{-r} \tag{5.6.1}$$

It is easily verified that for any constant $\eta > 0$, if s is large enough,

$$|\frac{d\lambda}{ds}(s)| \leq \eta \cdot (\lambda(s))^{1+r}$$

Therefore, using (2.c.I),

[13] Or $\lambda(s) = \frac{1}{s \ln(s)(\ln(\ln(s)))^\gamma}$, or $\lambda(s) = \frac{1}{s \ln(s) \ln(\ln(s))(\ln(\ln(\ln(s))))^\gamma}$, etc.; I will proceed to the proof for the first case; the others following similarly.

$$
\begin{aligned}
||v(\lambda(s\pm 1))-v(\lambda(s))|| &\leq B\cdot \sup_{0\leq\pm(\zeta-s)\leq 1}(\lambda(\zeta))^{-r}\sup_{0\leq\pm(\xi-s)\leq 1}|\frac{d\lambda}{ds}(\xi)| \\
&\leq B(1-\varepsilon)^{-r}\sup_{0\leq\pm(\xi-s)\leq 1}(\lambda(\xi))^{-r}|\frac{d\lambda}{ds}(\xi)| \\
&\leq B(1-\varepsilon)^{-r}\eta\sup_{0\leq\pm(\xi-s)\leq 1}\lambda(\xi) \\
&\leq B(1-\varepsilon)^{-r}(1+\varepsilon)\eta\lambda(s) \qquad (5.6.2)
\end{aligned}
$$

Taking η small enough and s large enough shows that (2.d.I) holds.

Similarly, using (2.c.II) we get

$$
\begin{aligned}
||v(\lambda(s\pm Ds))-v(\lambda(s))|| &\leq B(Ds)\cdot \sup_{0\leq\pm(\zeta-s)\leq Ds}(\lambda(\zeta))^{-r}\sup_{0\leq\pm(\xi-s)\leq Ds}|\frac{d\lambda}{ds}(\xi)| \\
&\leq B(Ds)(1-\varepsilon)^{-r}\sup_{0\leq\pm(\xi-s)\leq Ds}(\lambda(\xi))^{-r}|\frac{d\lambda}{ds}(\xi)| \\
&\leq B(Ds)(1-\varepsilon)^{-r}\eta\sup_{0\leq\pm(\xi-s)\leq Ds}\lambda(\xi) \\
&\leq B(Ds)(1-\varepsilon)^{-r}(1+\varepsilon)\eta\lambda(s) \qquad (5.6.3)
\end{aligned}
$$

Taking η small enough and s large enough shows that (2.d.II) holds.

References

[1] Aumann, R.J. (1964), Mixed and Behavior Strategies in Infinite Extensive Games, in M. Dresher, L.S. Shapley and A.W. Tucker (eds.), *Advances in Game Theory, Annals of Mathematics Studies* 52, Princeton University Press, Princeton, NJ, pp. 627-650.

[2] Bewley, T. and Kohlberg, E. (1976), The Asymptotic Theory of Stochastic Games, *Math. Oper. Res.* 1, 1976, 197–208

[3] Bewley, T. and Kohlberg, E. (1978), On Stochastic Games with Stationary Optimal Strategies, *Math. Oper. Res.* 3, 1978, 104–125

[4] Blackwell, D. and Ferguson, T.S. (1968), The Big Match, *Annals of Mathematical Statistics* 39, 159-163.

[5] Brown, L.D. and Purves, R. (1973), Measurable Selection of Extrema, *The Annals of Statistics*, 1, 902–912

[6] Coulomb, J.M. (2003), Stochastic Games Without Perfect Monitoring, *International Journal of Game Theory*, 32, 73–96

[7] Davenport, H. (1967), *Multiplicative Number Theory*, Markham Publishing Company, Chicago, Illinois.

[8] Dubins, L.E. (1957), A Discrete Invasion Game, A.W. Tucker, M. Dresher, P. Wolfe (eds.), *Contributions to the Theory of Games*, Vol. 3, Princeton University Press, Princeton, NJ., pp. 231-255

[9] Gillette, D. (1957), Stochastic Games With Zero Stop Probabilities, A.W. Tucker, M. Dresher, P. Wolfe (eds.), *Contributions to the Theory of Games*, Vol. 3 , Princeton University Press, Princeton, NJ., pp. 179–188

[10] Himmelberg, C.J., Parthasarathy, T., and Van Vleck, F.S. (1976), Optimal Plans for Dynamic Programming Problems, *Mathematics of Operations Research*, 1, 390–394

[11] Klein, E. and Thompson, A. (1984), *Theory of Correspondences (Canadian Mathematical Society Series of Monographs & Advanced Texts)*, John Wiley & Sons Inc., Hoboken, NJ.

[12] Kuratowski, K. and Ryll-Nardzewski, C. (1965), A General Theorem on Selectors, *Bulletin of the Polish Academy of Sciences (Ser. Mathematics)* 13, 379-403

[13] Maitra, A. and Sudderth, W. (2003), Stochastic Games With Borel Payoffs, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Kluwer Academic Publishers, Dordrecht, Chapter 24, pp. 367-373.

[14] Martin, D. (1998), The Determinacy of Blackwell Games, *The Journal of Symbolic Logic*, 63, 1565–1581.

[15] Mertens, J.F. (1986), The Minmax Theorem for U.S.C.–L.S.C. Payoff Functions, *International Journal of Game Theory*, 15, 237–250.

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

[17] Mertens, J.F., Neyman, A., and Rosenberg, D. (2007), Absorbing Games with Compact Action Spaces, Discussion Paper #456, Center for the Study of Rationality, Hebrew University, Jerusalem.

[18] Mertens, J.F., Neyman, A., and Rosenberg, D. (2008), Absorbing Games with Compact Action Spaces, *Mathematics of Operations Research*, forthcoming.

[19] Mertens, J.F., Sorin, S., and Zamir, S. (1994), Repeated Games, Core Discussion Papers 9420–9421–9422, Université Catholique de Louvain la Neuve, Belgium.

[20] Neyman, A. (2003), Real Algebraic Tools in Stochastic Games, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Kluwer Academic Publishers, Dordrecht, Chapter 6, pp. 57-75.

[21] Neyman, A. (2003), Stochastic Games: Existence of the Minmax, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Kluwer Academic Publishers, Dordrecht, Chapter 11, pp. 173–193.

[22] Nowak, A. (1985), Universally Measurable Strategies in Zero-Sum Stochastic Games, *Annals of Probability* 13, 269-287.

[23] Rosenberg, D., Solan, E., and Vieille, N. (2003), The MaxMin Value of Stochastic Games with Imperfect Monitoring, *International Journal of Game Theory* , 32, 133–150.

[24] Rosenberg, D., Solan, E., and Vieille, N. (2008), Protocols with No Acknowledgement, *Operations Research*, forthcoming.

[25] Rosenberg, D. and Sorin, S. (2001), An Operator Approach to Zero-Sum Repeated Games, *Israel Journal of Mathematics* 121, 221–246.

[26] Rudin, W. (1973), *Functional Analysis*, McGraw-Hill Inc., New York, NY.

[27] Scarf, H.E. and Shapley, L.S. (1957), Games With Partial Information, A.W. Tucker, M. Dresher, P. Wolfe (eds.), *Contributions to the Theory of Games*, Vol. 3 , Princeton University Press, Princeton, NJ., (1957) pp. 213-229

[28] Shapley, L.S., Stochastic Games, *Proceedings of the National Academy of Sciences*, U.S.A. 39, 1095-1100.

[29] Sorin, S. (2003), Discounted Stochastic Games: The Finite Case, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Kluwer Academic Publishers, Dordrecht, Chapter 5, pp. 51-55.

[30] Sorin, S. (2003), Stochastic Games with Incomplete Information, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Kluwer Academic Publishers, Dordrecht, Chapter 25, pp. 375-395.

[31] Thuijsman, F. (2003), The Big Match and the Paris Match, in A. Neyman and S. Sorin (eds.), *Stochastic Games and Applications*, NATO Science Series C, Mathematical and Physical Sciences, Vol. 570, Kluwer Academic Publishers, Dordrecht, Chapter 12, pp. 195–204.