SIAM J. CONTROL OPTIM.
Vol. 43, No. 1, pp. 86–110

STOCHASTIC GAMES WITH A SINGLE CONTROLLER AND INCOMPLETE INFORMATION*

DINAH ROSENBERG†, EILON SOLAN‡, AND NICOLAS VIEILLE§

Abstract. We study stochastic games with incomplete information on one side, in which the transition is controlled by one of the players.

We prove that if the informed player also controls the transitions, the game has a value, whereas if the uninformed player controls the transitions, the max-min value as well as the min-max value exist, but they may differ.

We discuss the structure of the optimal strategies, and provide extensions to the case of incomplete information on both sides.

Key words. stochastic games, incomplete information, single controller

AMS subject classifications. 91A05, 91A15

DOI. 10.1137/S0363012902407107

1. Introduction.

In a seminal work, Aumann and Maschler [1, 2] introduced infinitely repeated two-player zero-sum games with incomplete information on one side. Those are repeated games where the payoff matrix is known to one player, say player 1, but is not known to the other player—all player 2 knows is that the payoff matrix was drawn according to some known probability distribution from a finite set of possible matrices. Aumann and Maschler proved that those games have a value.

The issue faced by player 1 is the optimal use of information. On the one hand, player 1 needs to reveal his information (at least partially) in order to make use of it. On the other hand, any piece of information that is revealed to player 2 can later be exploited against player 1.

In the optimal strategies devised by Aumann and Maschler, player 1 reveals part of his information at the first stage, but no further information is revealed during the game. Player 2, on the other hand, has to play optimally whatever the actual payoff matrix may be. Aumann and Maschler achieved this by using Blackwell's approachability strategies.

When the underlying game is a stochastic game rather than a repeated one, the difficulties the players face are more serious.

Is it optimal for player 1 to reveal information only once in every state, or will he reveal information several times in each state? In repeated games, it does not help to dilute the revelation of information over time, since player 2 would wait until player 1 has revealed all the information he will ever reveal, and since interim payoffs are irrelevant in the long run. In stochastic games, by contrast, the game can move to a different state that can be more or less favorable to the informed player. By

*Received by the editors May 6, 2002; accepted for publication (in revised form) August 9, 2003; published electronically May 25, 2004. The authors acknowledge the financial support of the Arc-en-Ciel/Keshet program for 2001/2002. The research of the second author was supported by the Israel Science Foundation (grant 69/01-1).

http://www.siam.org/journals/sicon/43-1/40710.html

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

‡MEDS Department, Kellogg School of Management, Northwestern University and the School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel (eilons@post.tau.ac.il).

§Département Finance et Economie, HEC, 1 rue de la Libération, 78 Jouy-en-Josas, France (vieille@hec.fr).

giving away some information about the true game at the initial stage, player 1 might induce player 2 to adapt in an adverse way, while postponing this disclosure might allow player 1 to escape from specific states. This is a crude explanation for why it may help player 1 to conceal his information for a while.

For player 2 the issue is to devise the analog of Blackwell's approachability strategies for stochastic games.

Sorin [20, 21] and Sorin and Zamir [23] studied classes of stochastic games with incomplete information on one side that have a single nonabsorbing state, and proved that these games have a min-max value, a max-min value, and that the values of the n-stage (resp., λ-discounted) games converge as n goes to infinity (resp., as λ goes to 0) to the max-min value. Rosenberg and Vieille [17] studied recursive games with incomplete information on one side, and proved that the max-min value exists and is equal to the limit of the values of n-stage games (resp., λ-discounted games) as n goes to infinity (resp., as λ goes to 0).

In the present paper we study stochastic games in which one player controls the transitions; that is, the evolution of the stochastic state depends on the actions of one player but is independent of the actions of his opponent.

We show that if player 1 (who is the informed player) controls the transitions, then the game admits a value. We also propose a specific optimal strategy for player 1 and explain the way this strategy uses the additional information he possesses. Roughly speaking, the state space is partitioned into disjoint sets, which are called communicating sets. Whenever the play enters a communicating set, player 1 chooses a stationary nonrevealing strategy, and he plays this strategy until a new communicating set is visited. The random choice of the stationary strategy itself may be revealing, in that the distribution used at stage n to select a stationary strategy depends on the actual payoff function.

If player 2 controls the transitions, then the game admits a min-max value and a max-min value. We use an example to show that the two values may differ.

The techniques and the characterizations we provide extend the ideas of Aumann and Maschler for incomplete information games to our framework.

In the last section of the paper we extend the existence results to the case of stochastic games with a single controller and incomplete information on both sides; that is, to the case when each of the players has some partial private information about the true stochastic game being played.

2. The model and the main results.

2.1. The model. A *two-player zero-sum stochastic game* G is described by (i) a finite set Ω of states, and an initial state $\omega \in \Omega$; (ii) finite action sets I and J for the two players; (iii) a transition rule $q : \Omega \times I \times J \to \Delta(\Omega)$, where $\Delta(\Omega)$ is the simplex of probability distributions over Ω; and (iv) a reward function $g : \Omega \times I \times J \to \mathbf{R}$.

A *two-player zero-sum stochastic game with incomplete information* is described by a finite collection $(G_k)_{k\in K}$ of stochastic games, together with a distribution $p \in \Delta(K)$ over K. We assume that the games G_k differ only through their reward functions g^k, but they all have the same sets of states and actions, and the same transition rule. We denote the common transition rule by q.

The game is played in stages. An element $k \in K$ is chosen according to p. Player 1 is informed of k, while player 2 is not. At every stage $n \in \mathbf{N}$, the two players choose simultaneously actions $i_n \in I$ and $j_n \in J$, and ω_{n+1} is drawn according to $q(\cdot \mid \omega_n, i_n, j_n)$. Both players are informed of (i_n, j_n, ω_{n+1}). We stress that the actual reward $g^k(\omega_n, i_n, j_n)$ is not told to player 2 (but is known to player 1).

We parametrize the game by the initial distribution p and by the initial state ω, and denote it by $\Gamma(p,\omega)$. We write Γ for $(\Gamma(p,\omega))_{(p,\omega)\in\Delta(K)\times\Omega}$.

A few remarks are in order. This model is an extension of the classical model of zero-sum stochastic games. It is also an extension of Aumann and Maschler's model of repeated games with incomplete information, where a zero-sum matrix game is first drawn using p, then played repeatedly over time. Here, nature chooses a stochastic game that is then played over time.

We assume without loss of generality (w.l.o.g.) that $0 \le g^k \le 1$ for every $k \in K$, and we identify each $k \in K$ with the probability measure over K that gives weight 1 to k.

2.2. Strategies and values. Players may base their choices on the stochastic states the play has visited so far, as well as on past choices of actions (of the two players). Player 1 can base his choices also on the state of the world k.

The space of histories of length n is $H_n = (\Omega \times I \times J)^{n-1} \times \Omega$, the space of finite histories is $H = \cup_{n\in\mathbf{N}} H_n$, and the space of plays (infinite histories) is $H_\infty = (\Omega \times I \times J)^\infty$. H_n defines naturally a finite algebra $\mathcal{H}_n$ over H_∞. We equip H_∞ with the σ-algebra $\vee_{n\in\mathbf{N}}\mathcal{H}_n$ spanned by all cylinder sets. A (behavioral) *strategy* of player 1 is a function $\sigma : K \times H \to \Delta(I)$. A strategy for player 2 is a function $\tau : H \to \Delta(J)$. A strategy $\sigma = (\sigma_k)_{k\in K}$ of player 1 is *nonrevealing* if σ_k is independent of $k \in K$.[1]

A strategy σ is stationary if the mixed action played at every stage depends only on the current state. We identify each vector $x = (x_\omega)_{\omega\in\Omega} \in (\Delta(I))^\Omega$ with the stationary strategy that plays the mixed action x_ω whenever the game visits ω. Stationary strategies of player 2 are defined analogously.

Every distribution p, every initial stochastic state ω, and every pair of strategies (σ,τ) induce a probability measure $\mathbf{P}_{p,\omega,\sigma,\tau}$ over $K \times H_\infty$ (equipped with the product σ-algebra). We denote by $\mathbf{E}_{p,\omega,\sigma,\tau}$ the corresponding expectation operator.

We let k, ω_n, i_n and j_n denote, respectively, the actual game being played, the current state at stage n, and the actions played at stage n. These are random variables.

Define the expected average payoff up to stage N by

$$\gamma_N(p,\omega,\sigma,\tau) = \mathbf{E}_{p,\omega,\sigma,\tau}\left[\overline{g}_N\right],$$

where $\overline{g}_N = \frac{1}{N}\sum_{n=1}^N g^k(\omega_n, i_n, j_n)$. For fixed strategies σ, τ, $\gamma_N(p,\omega,\sigma,\tau)$ is linear in p and 1-Lipshitz.

We recall the definitions of the max-min value, the min-max value, and the (uniform) value.

Definition 1. *Player* 1 *can* guarantee $\phi \in \mathbf{R}$ *in the game* $\Gamma(p,\omega)$ *if, for every* $\epsilon > 0$*, there exists a strategy* σ *of player* 1 *and* $N \in \mathbf{N}$ *such that*

$$\forall \tau, \forall n \ge N, \quad \gamma_n(p,\omega,\sigma,\tau) \ge \phi - \epsilon.$$

We then say that the strategy σ guarantees $\phi - \epsilon$ *in* $\Gamma(p,\omega)$.

Player 1 *can guarantee a function* $\phi : \Delta(K) \times \Omega \to \mathbf{R}$ *if player* 1 *can guarantee* $\phi(p,\omega)$ *in the game* $\Gamma(p,\omega)$*, for every* $(p,\omega) \in \Delta(K) \times \Omega$.

Note that, due to the Lipshitz property on payoffs and the compactness of $\Delta(K)$, the integer N in Definition 1 can be chosen to be independent of (p,ω). The definition

[1] The strategy is nonrevealing in the sense that knowledge of the strategy σ and of past play does not enable player 2 to gain information on k. This property relies on the fact that transitions are independent of k.

of a function that is guaranteed by player 2 is similar, with the roles of the two players exchanged.

DEFINITION 2. *Player 2 can* defend $\phi \in \mathbf{R}$ *in the game* $\Gamma(p, \omega)$ *if, for every* $\epsilon > 0$ *and every strategy* σ *of player 1, there exists a strategy* τ *of player 2 and* $N \in \mathbf{N}$ *such that*

$$\forall n \geq N, \quad \gamma_n(p, \omega, \sigma, \tau) \leq \phi + \epsilon. \tag{1}$$

We say that such a strategy τ defends $\phi + \epsilon$ *against* σ *in* $\Gamma(p, \omega)$.

Player 2 can defend a function $\phi : \Delta(K) \times \Omega \to \mathbf{R}$ *if player 2 can defend* $\phi(p, \omega)$ *in the game* $\Gamma(p, \omega)$*, for every* $(p, \omega) \in \Delta(K) \times \Omega$.

The definition of a function that is defended by player 1 is similar, with the roles of the two players exchanged. The following lemma follows from the definitions.

LEMMA 3. *Player 1 can guarantee (resp., defend)* $\max\{\phi, \phi'\}$ *as soon as he can guarantee (resp., defend) both* ϕ *and* ϕ'*. Player 2 can guarantee (resp., defend)* $\min\{\phi, \phi'\}$ *as soon as he can guarantee (resp., defend) both* ϕ *and* ϕ'.

DEFINITION 4. *A function* $\phi : \Delta(K) \times \Omega \to \mathbf{R}$ *is*

- *the* (uniform) value *of* Γ *if both players can guarantee* ϕ*;*
- *the* max-min value *of* Γ *if player 1 can guarantee* ϕ *and player 2 can defend* ϕ*;*
- *the* min-max value *of* Γ *if player 1 can defend* ϕ *and player 2 can guarantee* ϕ.

Note that the value exists if and only if the max-min value and min-max value exist and coincide.

The value (resp., max-min value, min-max value) is denoted by v (resp., $\underline{v}, \bar{v}$) when it exists. Observe that $\underline{v} \leq \bar{v}$ whenever the two exist. Note that each of the functions $\underline{v}$ and $\bar{v}$ is 1-Lipshitz in p as soon as it exists. When the value v exists, any strategy that guarantees v up to ε is ε*-optimal*. Strategies that are ε-optimal for each $\varepsilon > 0$ are also termed *optimal*.

2.3. Related literature. Most of the literature deals with the polar cases where either Ω or K is a singleton. In the former case, the game is a repeated game with incomplete information. Such games have a value; see Aumann and Maschler [2]. Moreover, an explicit formula for the value exists. Letting $u^*(p)$ be the value of the matrix game with payoff function $\sum_k p_k g^k(\cdot, \cdot)$, the value of the repeated game with incomplete information is the concavification cav u^* of u^* (see section 3.1 for definitions).

When K is a singleton the game is a standard stochastic game. Such games have a value; see Mertens and Neyman [9].

For general stochastic games with incomplete information, little is known, but some classes were studied in the literature. For "Big Match" games Sorin [20, 21] and Sorin and Zamir [23] proved the existence of the max-min value and of the min-max value. These values may differ.

For recursive games, Rosenberg and Vieille [17] proved that the max-min value exists and provided an example where the value does not exist.

Parthasarathy and Raghavan [14] were the first to study the class of stochastic games in which one player controls the transitions. They proved that in this class the value exists, and both players have optimal stationary strategies. They also studied the two-player non-zero-sum game. Filar [5] studied the situation in which states are partitioned into two subsets, and each player controls the transitions from states in his subset of the partition. Several finite-stage algorithms that calculate the value and optimal stationary strategies were proposed in the literature (see the survey by Raghavan and Filar [15] and the references therein).

Recently Renault [16] studied games where transitions do not depend on the actions chosen by the players and only player 1 observes the current state of the world. All that player 2 observes are the actions of player 1.

2.4. Statements of the results. In the present paper we consider games where a single player controls the transitions.

DEFINITION 5. *Player* 1 controls the transitions *if, for every* $\omega \in \Omega$ *and every* $i \in I$, *the transition* $q(\cdot \mid \omega, i, j)$ *does not depend on* j. *Player* 2 *controls the transitions if the symmetric property holds. We then simply write* $q(\cdot \mid \omega, i)$ *or* $q(\cdot \mid \omega, j)$ *depending on who controls the transitions.*

We prove the following two results.

THEOREM 6. *If player* 1 *controls the transitions, the value exists.*

THEOREM 7. *If player* 2 *controls the transitions, both the min-max value and max-min value exist.*

We provide an example of a game where player 2 controls the transitions and $\overline{v} \neq \underline{v}$. We also provide a characterization of $\bar{v}$ and $\underline{v}$ as a unique solution to a functional equation, and we study the structure of simple optimal strategies of player 1.

We prove no result on the existence of the limit of the values of the finitely repeated games. In the games analyzed so far (see section 2.3), this limit is known to exist and coincides with $\underline{v}$. This property is conjectured to hold in general by Mertens [8].

3. Various tools. This section gathers a few results that we use in subsequent sections. The first three subsections introduce a few extensions of tools used in the analysis of games with incomplete information.

For three vectors $a, b, c \in \mathbf{R}^K$, $c = a + b$ if and only if $c_k = a_k + b_k$ for every $k \in K$, $c = \max\{a, b\}$ if and only if $c_k = \max\{a_k, b_k\}$ for every $k = 1, \ldots, K$, and $a \geq b$ if and only if $a_k \geq b_k$ for every $k = 1, \ldots, K$. For a scalar $r \in \mathbf{R}$, $c = a + r$ if and only if $c_k = a_k + r$ for every $k = 1, \ldots, K$, and $c = ra$ if and only if $c_k = ra_k$ for every $k = 1, \ldots, K$. Finally, the L_1-norm and L_∞-norm will be denoted by $\|\cdot\|_1$ and $\|\cdot\|_\infty$, respectively.

3.1. Concavification. Given a continuous function $u : \Delta(K) \to \mathbf{R}$, we denote by cav u its concavification, namely, the least concave function v defined over $\Delta(K)$, such that $v \geq u$. It is the function whose hypograph is the convex hull of the hypograph of u. Similarly, we denote by vex u its convexification, namely, the largest convex function v such that $v \leq u$. Both cav u and vex u are well defined. Thus, cav and vex are functional operators that act on real-valued functions defined on $\Delta(K)$.

LEMMA 8 (see, e.g., Laraki [7]). *When* $\Delta(K)$ *is endowed with the* L_1*-norm, the two operators* cav *and* vex *map* C*-Lipshitz functions into* C*-Lipshitz functions.*

LEMMA 9. *When the set of functions* $u : \Delta(K) \to \mathbf{R}$ *is endowed with the* L_∞*-norm, the two operators* cav *and* vex *are nonexpansive.*

Proof. For any two real-valued continuous functions over $\Delta(K)$, u, and v, one has

$$\|u^{**} - v^{**}\|_\infty \leq \|u^* - v^*\|_\infty \leq \|u - v\|_\infty,$$

where $u^*(x) = \inf\{\langle y, x\rangle - u(y), y \in \mathbf{R}^K\}$ is the conjugate of u. Since $u^{**} = \text{cav } u$, the result follows.

The argument for the operator vex is analogous. $\square$

The following lemma is classical (see, e.g., Mertens, Sorin, and Zamir [10, Corollary V.1.3], or the discussion in Zamir [24, p. 118]).

LEMMA 10. *Assume that player* 1 *can guarantee* u. *Then player* 1 *can guarantee* cav u.

Proof. We briefly recall the main ideas of the proof. Prior to the first stage, player 1 performs a state-dependent lottery, designed as follows. By the Carathéodory theorem there exist $p_e \in \Delta(K), \alpha_e \in [0,1]$, for $e = 1, \ldots, |K|+1$, such that $\sum_e \alpha_e = 1$, $\sum_e \alpha_e p_e = p$, and

$$\text{cav } u(p) \leq \sum_e \alpha_e u(p_e) + \varepsilon. \tag{2}$$

If u is continuous, ε may be set to zero in (2). To guarantee cav $u(p)$ in $\Gamma(p,\omega)$, player 1 chooses a fictitious distribution p_e, and he plays optimally in $\Gamma(p_e,\omega)$. The distributions (p_e) must satisfy that their average is p. We now provide one mechanism player 1 can employ.

For each e set $\mu^k(e) = \alpha_e p_e^k / p^k$ if $p^k > 0$, and we let μ^k be arbitrary if $p^k = 0$. Observe that $\sum_e p^k \mu^k(e) = \sum_e \alpha_e p_e^k = p^k$. The following strategy of player 1 guarantees cav $u(p) - 2\varepsilon$: given k, choose e according to μ^k, and play a strategy σ_e that guarantees $u(p_e) - \varepsilon$. ∎

The following result will be useful later.

LEMMA 11. *Let* $(\mathcal{A}_i)_{i \in I}$ *be a finite collection of convex closed upwards comprehensive sets, and let* $\mathcal{A}$ *be the set* $\left\{a \in R^K : a = \max_{i \in I} a_i, a_i \in \mathcal{A}_i\right\}$. *Then*

$$f_{\mathcal{A}}(p) = (\text{cav } \max_{i \in I} f_{\mathcal{A}_i})(p),$$

where, for any convex upwards comprehensive set $\mathcal{B}$, $f_{\mathcal{B}}(p) = \inf_{a \in \mathcal{B}} \langle a, p \rangle$.

Proof. Since each $\mathcal{A}_i$ is upwards comprehensive, $\mathcal{A}$ coincides with $\cap_i \mathcal{A}_i$. Therefore $f_{\mathcal{A}} \geq f_{\mathcal{A}_i}$ for each i. In particular $f_{\mathcal{A}} \geq \max_{i \in I} f_{\mathcal{A}_i}$. Since $f_{\mathcal{A}}$ is concave, $f_{\mathcal{A}} \geq$ cav $\max_{i \in I} f_{\mathcal{A}_i}$.

To prove the opposite inequality, we first observe that if $\mathcal{B}$ is convex, closed, and upwards comprehensive, one has

$$\mathcal{B} = \left\{a \in \mathbf{R}^K : \langle a, p \rangle \geq f_{\mathcal{B}}(p) \text{ for each } p \in \Delta(K)\right\}. \tag{3}$$

Set $g = $ cav $\max_{i \in I} f_{\mathcal{A}_i}$, and

$$\mathcal{D} = \left\{a \in \mathbf{R}^K : \langle a, p \rangle \geq g(p) \text{ for each } p \in \Delta(K)\right\}.$$

Since $g \geq f_{\mathcal{A}_i}$ for each $i \in I$, and using (3) with $\mathcal{B} = \mathcal{A}_i$, one has $\mathcal{D} \subseteq \mathcal{A}_i$. Therefore, $\mathcal{D} \subseteq \mathcal{A}$, which readily implies $g \geq f_{\mathcal{A}}$. ∎

3.2. Approachability. We present here the basic approachability result of Blackwell [4], in the framework of stochastic games. Let G be a stochastic game with payoffs in $\mathbf{R}^K$. The description of such a game is the same as that of a two-player zero-sum stochastic game given in section 2.1, except that the reward function g now takes values in $\mathbf{R}^K$. The definition of strategies in this framework is similar to that given in section 2.2.

We denote $\bar{g}_N = \frac{1}{N} \sum_{n=1}^N g(\omega_n, i_n, j_n) \in \mathbf{R}^K$, the average vector payoff in the first N stages.

DEFINITION 12. *A vector $a \in \mathbf{R}^K$ is* approachable *by player 2 at ω if, for every $\varepsilon > 0$, there is a strategy τ of player 2 and $N \in \mathbf{N}$ such that*[2]

$$\forall \sigma, \mathbf{E}_{\omega,\sigma,\tau}\left[\sup_{n \geq N} (\bar{g}_n - a)^+\right] \leq \varepsilon.$$

We say that such a strategy τ approaches *$a + \varepsilon$ at ω.*

In other words, for every ε player 2 has a strategy such that the average payoff vector will eventually not exceed $a + \varepsilon$. Note that a is approachable if and only if $a + \varepsilon$ is approachable for every $\varepsilon > 0$, so that the set of approachable vectors is closed and upwards comprehensive.

Our definition differs slightly from that of Blackwell [4], where the strategy τ is required to be independent of ε (i.e., the original definition of Blackwell reads as $\exists \tau, \forall \varepsilon > 0$, etc.). Any vector a that is approachable in Blackwell's sense is also approachable in our sense. The two definitions are not equivalent. However, it is easily checked that, if a is approachable (in our sense) at each state, it is also approachable in Blackwell's sense.

Every stochastic game with incomplete information $\Gamma(p, \omega)$ induces a stochastic game with vector payoffs $\Gamma^V(\omega)$, in which the payoff coordinates are given by the reward functions of the component games (G_k) of $\Gamma(p, \omega)$. The next lemma relates the two games. Its proof is straightforward.

LEMMA 13. *If $a \in \mathbf{R}^K$ is approachable at ω in the game Γ^V, then player 2 can guarantee $\langle a, p\rangle$ in $\Gamma(p, \omega)$ for every $p \in \Delta(K)$.*

We now state Blackwell's sufficient condition for approachability in this context. Denote by $u_\infty(p, \omega)$ the uniform value of the two-player zero-sum stochastic game with reward function $\sum_{k \in K} p_k g^k(\omega, \cdot, \cdot)$. The existence of u_∞ follows, by Mertens and Neyman [9] or by Parthasarathy and Raghavan [14]. We also denote by $u_n(p, \omega)$ the value of the n-stage version of that game (thus, $\lim_{n\to\infty} u_n = u_\infty$, and the limit is uniform in p).

PROPOSITION 14. *If* cav $u_\infty(p, \omega) \leq \langle a, p\rangle$ *for every $(p, \omega) \in \Delta(K) \times \Omega$, then a is approachable in Γ^V by player 2 at ω, for each $\omega \in \Omega$.*

In this statement (and in later ones), cav u_∞ is the concavification of u_∞ with respect to the first variable, p: cav $u_\infty(p, \omega) = (\text{cav } u_\infty(\cdot, \omega))(p)$.

Sketch of the proof. Let $\varepsilon > 0$, and choose N such that $\|u_N - u_\infty\| \leq \varepsilon$, so that cav $u_N(p, \omega) \leq \langle a + \varepsilon, p\rangle$. We define an auxiliary game with vector payoffs, where each stage corresponds to N stages in the original game. We apply Blackwell's result to the auxiliary game, noting that Blackwell's proof still holds when the stage game changes from stage to stage, with payoffs remaining bounded. □

A more precise result was proved by Milman [13, Theorem 2.1.1]. For results with similar flavor, see Shimkin and Shwartz [19].

3.3. Information revelation. Let σ be a given strategy of player 1. For $n \in \mathbf{N}$, we denote by p_n the conditional distribution over K given $\mathcal{H}_n$: it is the belief held by player 2 about the true game being played.[3] The difference $\|p_n - p_{n+1}\|_1$ may be interpreted as the amount of information that is revealed at stage n.

It is well known (see, e.g., Sorin [22, Lemma 3.4] or Mertens, Sorin, and Zamir

[2] For every real $c \in \mathbf{R}$, $c^+ = \max\{c, 0\}$.

[3] The value of p_n at a specific atom of $\mathcal{H}_n$ depends on σ but not on τ. Since the distribution on $\mathcal{H}_n$ depends on τ, the law of p_n depends on both σ and τ.

[24, Lemma IV.2.1]) that, for each τ,

$$\mathbf{E}_{p,\omega,\sigma,\tau}\left[\sum_{n=1}^{\infty}\|p_n-p_{n+1}\|_2^2\right]\leq |K|\,. \tag{4}$$

Given $p\in\Delta(K)$, we denote by σ^p the average nonrevealing strategy defined by $\sigma^p(h)=\sum_{k\in K}p(k)\sigma(k,h)$ for each finite history h. It is convenient to relate the benefit derived by player 1 from using his information at a given stage to the amount of information revealed at that stage. Let $n\in\mathbf{N}$ be given. The expected payoff at stage n, conditional on past play, is

$$\mathbf{E}_{p,\omega,\sigma,\tau}\left[g_n|\mathcal{H}_n\right]=\sum_{k\in K}p_n(k)g^k(\omega_n,\sigma(k,h_n),\tau(h_n)),$$

where $\sigma(k,h_n)$ and $\tau(h_n)$ are the mixed moves used by the two players at that stage.[4] By Proposition 3.2 and Lemma 3.13 in Sorin [22],

$$\left|\mathbf{E}_{p,\omega,\sigma,\tau}\left[g_n|\mathcal{H}_n\right]-\langle p_n,g(\omega_n,\sigma^{p_n}(h_n),\tau(h_n))\rangle\right|\leq\mathbf{E}\left[\|p_n-p_{n+1}\|_1|\mathcal{H}_n\right]. \tag{5}$$

Definition 15. *Let $\widetilde{\mathcal{T}}$ be a set of strategies of player 2. Let $\varepsilon>0$ and σ be given. The strategy $\widetilde{\tau}\in\widetilde{\mathcal{T}}$ is ε-*exhausting information *given (p,ω) and σ if $\widetilde{\tau}$ maximizes $\mathbf{E}_{p,\omega,\sigma,\tau}[\sum_{n=1}^{\infty}\|p_n-p_{n+1}\|_2^2]$ up to ε over $\widetilde{\mathcal{T}}$.*

This notion is relative to the class $\widetilde{\mathcal{T}}$. Which class of strategies is meant will always be clear from the context.

Lemma 16. *Let $\widetilde{\mathcal{T}},\varepsilon,\sigma,(p,\omega)$ as in Definition 15. Let $\widetilde{\tau}\in\widetilde{\mathcal{T}}$ be an ε-exhausting strategy given (p,ω) and σ, and let $N\in\mathbf{N}$ be such that $\mathbf{E}_{p,\omega,\sigma,\widetilde{\tau}}[\sum_{n=N}^{\infty}\|p_n-p_{n+1}\|_2^2]\leq\varepsilon$. Then for each strategy $\tau\in\widetilde{\mathcal{T}}$ that coincides with $\widetilde{\tau}$ until stage N, one has*

$$\mathbf{E}_{p,\omega,\sigma,\tau}\left[\sum_{n=N}^{\infty}\|p_n-p_{n+1}\|_1\right]\leq\sqrt{2\varepsilon},\ \textit{and}\ \mathbf{E}_{p,\omega,\sigma,\tau}\left[\|p_l-p_N\|_2\right]\leq\sqrt{2\varepsilon}\ \textit{for each}\ l\geq N.$$

Proof. By Jensen's inequality and since (p_n) is a martingale, for every $l\geq N$ one has

$$\left(\mathbf{E}_{p,\omega,\sigma,\tau}\left[\|p_l-p_N\|_2\right]\right)^2\leq\mathbf{E}_{p,\omega,\sigma,\tau}\left[\|p_l-p_N\|_2^2\right]=\mathbf{E}_{p,\omega,\sigma,\tau}\left[\sum_{n=N}^{l-1}\|p_n-p_{n+1}\|_2^2\right]. \tag{6}$$

The equality in (6) is a standard result for martingales; see, e.g., Karatzas and Shreve [6, p. 32]. The second inequality follows. The first inequality follows using Jensen's inequality (applied to each stage independently) and since $\|\cdot\|_1\leq\|\cdot\|_2$. ◻

The next lemma is specific to stochastic games with incomplete information. In effect, it proves that the amount of information revealed by player 1 up to stage $l\in\mathbf{N}$ is an upper bound on the excess gain from the private information.

Lemma 17. *Let (σ,τ) be given. For every $p\in\Delta(K)$, every $\omega\in\Omega$, and every $l\in\mathbf{N}$, one has*

$$\left|\mathbf{E}_{p,\omega,\sigma,\tau}\left[\overline{g}_l\right]-\mathbf{E}_{p,\omega,\sigma^p,\tau}\left[\overline{g}_l\right]\right|\leq 4\mathbf{E}_{p,\omega,\sigma,\tau}\left[\sum_{m=1}^{l}\|p_m-p_{m+1}\|_1\right].$$

[4] There is a small notational inconsistency here, since the right-hand side is the value of the left-hand side on a typical atom of $\mathcal{H}_n$.

Proof. To distinguish between $\mathbf{E}_{p,\omega,\sigma,\tau}$ and $\mathbf{E}_{p,\omega,\sigma^p,\tau}$, we denote the latter by $\widetilde{\mathbf{E}}_{p,\omega,\sigma^p,\tau}$. Let $n \leq l$ be given. Since σ^p is nonrevealing, and by the Lipshitz property,

$$\begin{aligned}
&\left|\langle p_n, g(\omega_n, \sigma^{p_n}(h_n), \tau(h_n))\rangle - \widetilde{\mathbf{E}}_{p,\omega,\sigma^p,\tau}\left[g_n|\mathcal{H}_n\right]\right| \\
&\qquad = \left|\langle p_n, g(\omega_n, \sigma^{p_n}(h_n), \tau(h_n))\rangle - \langle p, g(\omega_n, \sigma^{p}(h_n), \tau(h_n))\rangle\right| \\
&\qquad \leq 2\left\|p_n - p\right\|_1 . \qquad (7)
\end{aligned}$$

By (5), it follows that

$$\left|\mathbf{E}_{p,\omega,\sigma,\tau}\left[g_n|\mathcal{H}_n\right] - \widetilde{\mathbf{E}}_{p,\omega,\sigma^p,\tau}\left[g_n|\mathcal{H}_n\right]\right| \leq 2\left\|p_n - p\right\|_1 + \mathbf{E}_{p,\omega,\sigma,\tau}\left[\left\|p_n - p_{n+1}\right\|_1|\mathcal{H}_n\right]. \qquad (8)$$

On the other hand, it is easily checked that the probabilities $\mathbf{P}^n_{p,\omega,\sigma,\tau}$ and $\widetilde{\mathbf{P}}^n_{p,\omega,\sigma^p,\tau}$ induced by $\mathbf{P}$ and $\mathbf{P}_{p,\omega,\sigma^p,\tau}$ on $\mathcal{H}_n$ satisfy

$$\left\|\mathbf{P}^n_{p,\omega,\sigma,\tau} - \widetilde{\mathbf{P}}^n_{p,\omega,\sigma^p,\tau}\right\|_1 \leq \mathbf{E}_{p,\omega,\sigma,\tau}\left[\sum_{m=1}^{n}\left\|p_m - p_{m+1}\right\|_1\right]. \qquad (9)$$

By (8) and (9),

$$\left|\mathbf{E}_{p,\omega,\sigma,\tau}\left[g_n\right] - \widetilde{\mathbf{E}}_{p,\omega,\sigma^p,\tau}\left[g_n\right]\right| \leq 4\mathbf{E}_{p,\omega,\sigma,\tau}\left[\sum_{m=1}^{n}\left\|p_m - p_{m+1}\right\|_1\right],$$

which implies the result. □

3.4. A partition of states. In this section we define a partition of the set of states that will be extensively used in what follows. It hinges on the fact that a single player controls the transitions, but it does not matter who is the controller. The partition is similar to the one defined by Ross and Varadarajan [18] for Markov decision processes, who also provide an algorithm to calculate it.

We assume that player 1 controls the transitions. The partition when player 2 controls the transitions is defined analogously. Since transitions are independent of player 2's actions, we here omit player 2's strategy from the notations.

Given $\omega \in \Omega$, we denote by

$$r_\omega = \min\left\{n \in \mathbf{N}, \omega_n = \omega\right\}$$

the stage of the first visit to ω. By convention, the minimum over an empty set is $+\infty$.

DEFINITION 18. *Let $\omega_1, \omega_2 \in \Omega$. We say that ω_1* leads to *ω_2 if $\omega_1 = \omega_2$, or if $\mathbf{P}_{\omega_1,\sigma}(r_{\omega_2} < +\infty) = 1$ for some strategy σ of player 1.*

Note that the relation *leads to* is reflexive and transitive.

We define an equivalence relation over Ω by

$$\omega \leftrightarrow \omega' \text{ if and only if } \omega \text{ leads to } \omega' \text{ and } \omega' \text{ leads to } \omega.$$

The equivalence classes of this relation are called *communicating sets*. Given $\omega \in \Omega$, we let C_ω denote the communicating set that contains ω, and we define

$$I_\omega = \{i \in I \colon q(C_\omega \mid \omega, i) = 1\}.$$

Thus, whenever C_ω contains at least two elements, by properly selecting actions in $(I_{\omega'})_{\omega' \in C_\omega}$ player 1 can ensure that the play reaches any state in C_ω infinitely often, provided the play starts in C_ω.

The set I_ω may (but does not have to) be empty only if $|C_\omega| = 1$. Actions in I_ω are called *stay* actions, and any state ω such that $I_\omega = \emptyset$ is a *null* state. The set of nonnull states is denoted by Ω_c. Note that $C_\omega \subseteq \Omega_c$ whenever $\omega \in \Omega_c$.

LEMMA 19. *$\omega \in \Omega_c$ if and only if there is a stationary strategy x_{C_ω} such that C_ω is a recurrent set for x.*

Thus, a state is null if it is visited only finitely many times, whatever player 1 plays: $I_\omega = \emptyset$ if and only if ω is transient for every stationary strategy x.

Proof. We start with the direct implication. Let $\omega \in \Omega_c$. For $\omega' \in C_\omega$, define $x_{\omega'} \in \Delta(A)$ by

$$x_{\omega'}[i] = \begin{cases} 0, & i \notin I_{\omega'}, \\ 1/|I_{\omega'}|, & i \in I_{\omega'}, \end{cases}$$

and let x be any stationary strategy that coincides with $x_{\omega'}$ in each state $\omega' \in C_\omega$. It is easy to show that C_ω is recurrent under x.

The reverse implication is straightforward. $\square$

Some communicating sets are absorbing, in the sense that once entered, the play remains there forever. We now single them out. Let x^* be a fully mixed stationary strategy, i.e., $x^*_\omega[i] > 0$ for every $\omega \in \Omega$ and every $i \in I$. If $R \subseteq \Omega$ is a recurrent set for x^*, then R is a communicating set, and $I_\omega = I$ for every $\omega \in R$.

We denote by Ω_0 the union of these sets:

$$\Omega_0 = \cup\{R : R \text{ recurrent for } x^*\} = \{\omega \in \Omega : I_{\omega'} = I \text{ for every } \omega' \in C_\omega\}.$$

The following lemma implies that the max-min value and the min-max value are constant over C_ω for every $\omega \in \Omega_0$, provided they exist.

LEMMA 20. *Assume player* 1 *controls transitions. Let $\omega \in \Omega$ and $\omega' \in C_\omega$. If one of the players can guarantee ϕ in $\Gamma(p, \omega)$, he can also guarantee ϕ in $\Gamma(p, \omega')$.*

Proof. Assume first that player 1 can guarantee ϕ in $\Gamma(p, \omega)$. Let σ be a strategy that guarantees $\phi - \varepsilon$ in $\Gamma(p, \omega)$, and let σ^* be the strategy that plays x_{C_ω} until r_ω, then switches to σ. In the game $\Gamma(p, \omega')$, the strategy σ^* guarantees $\phi - \varepsilon'$ for each $\varepsilon' > \varepsilon$.

Assume now that player 2 can guarantee ϕ in $\Gamma(p, \omega)$, but assume to the contrary that he *cannot* guarantee ϕ in $\Gamma(p, \omega')$ for some $\omega' \in C_\omega$. We argue that player 2 cannot guarantee ϕ in $\Gamma(p, \omega)$, a contradiction. Since player 2 cannot guarantee ϕ in $\Gamma(p, \omega')$, there is $\varepsilon > 0$ such that for every strategy τ of player 2 and every $N \in \mathbf{N}$ there is a strategy $\sigma_{\tau,N}$ of player 1 and an integer $n_{\tau,N} \geq N$ such that $\gamma_{n_{\tau,N}}(p, \omega', \sigma_{\tau,N}, \tau) > \phi + \varepsilon$. Let τ and N be given. Let σ^* be the strategy of player 1 defined as follows. Play x_{C_ω} until stage $r_{\omega'}$, then switch to $\sigma_{\tau_\nu, M}$, where τ_ν is the strategy induced by τ after stage ν, and M is sufficiently large so that $\mathbf{P}_{\omega, x_{C_\omega}}(r_{\omega'} < M) > 1 - \frac{\varepsilon}{2}$. One can verify that there is $n' \geq N$ such that $\gamma_{n'}(p, \omega, \sigma^*, \tau) > \phi + \varepsilon/2$, a contradiction. $\square$

When player 2 controls the transitions, we denote by J_ω the set of stay actions at ω:

$$J_\omega = \{j \in J : q(C_\omega \mid \omega, j) = 1\}.$$

3.5. Auxiliary games. As for the analysis of zero-sum repeated games with incomplete information on one side, it is convenient to introduce an average game in which no player is informed of the realization of k.

For notational ease, assume that player 1 is the controller. For every $p \in \Delta(K)$ and every nonnull state $\omega \in \Omega$, we denote by $\widetilde{\Gamma}_R(p,\omega)$ the zero-sum stochastic game with (i) initial state ω, (ii) state space C_ω, (iii) reward function $\sum_k p_k g^k$, (iv) action sets $I_{\omega'}$ and J at each state $\omega' \in C_\omega$, and (v) transition function induced by q.

In the case where player 2 is the controller, the game $\widetilde{\Gamma}_R(p,\omega)$ is defined by restricting player 2's action set to $J_{\omega'}$ in each state $\omega' \in C_\omega$.

Thus, $\widetilde{\Gamma}_R(p,\omega)$ is the stochastic game in which player 1 is not informed of the realization of k (or does not use his information), and the controller is restricted to stay actions. In particular, the game remains in C_ω forever. The letter R is a symbol for *restricted*, while the symbol $\widetilde{}$ stands for average.

Note that $\widetilde{\Gamma}_R(p,\omega)$ is a single controller game, so that both players have optimal stationary strategies. Denote by $\widetilde{u}(p,\omega)$ its value. Note that $\widetilde{u}(p,\omega) = u_\infty(p,\omega)$ for each $\omega \in \Omega_0$.

By convention, if ω is a null state, we set $\widetilde{u}(p,\omega) = -\infty$ if player 1 controls the transitions, and $\widetilde{u}(p,\omega) = +\infty$ if player 2 controls the transitions. By Lemma 20, for every communicating set C, $\widetilde{u}(p,\omega)$ is independent of $\omega \in C$.

PROPOSITION 21. *For every $\omega \in \Omega_0$ and every $p \in \Delta(K)$ the value $v(p,\omega)$ of $\Gamma(p,\omega)$ exists and is equal to* cav $\widetilde{u}(p,\omega)(=$ cav $u_\infty(p,\omega))$.

Thus, restricted to Ω_0, the game is similar to a standard repeated game with incomplete information.

Proof. The proof of this lemma is similar to the proof for repeated games with incomplete information on one side. Let $p \in \Delta(K)$ and $\omega \in \Omega_0$ be given. Clearly player 1, by not using his information, can guarantee $\widetilde{u}(p,\omega)$. By Lemma 10, player 1 can guarantee cav $\widetilde{u}(p,\omega)$.

The proof that player 2 can guarantee cav $\widetilde{u}$ is based on approachability results, and closely follows classical lines. Let $a \in \mathbf{R}^K$ be such that

$$\begin{aligned}
&\langle a,p\rangle = \text{cav } \widetilde{u}(p,\omega),\\
&\langle a,q\rangle \geq \text{cav } \widetilde{u}(q,\omega) \text{ for every } q \in \Delta(K).
\end{aligned}$$

If cav $\widetilde{u}(\cdot,\omega)$ is differentiable at p, then a is defined by the hyperplane tangent to cav $\widetilde{u}(\cdot,\omega)$ at p. By Proposition 14, a is approachable. By Lemma 13, player 2 can guarantee cav $\widetilde{u}(p,\omega)$. ◻

Let $\Gamma_R(p,\omega)$ be a game similar to $\widetilde{\Gamma}_R(p,\omega)$, but in which player 1 is informed of k. Thus, $\Gamma_R(p,\omega)$ differs from $\Gamma(p,\omega)$ only in that actions of the controller are restricted.

Since in Γ_R, for each nonnull state ω the game cannot leave C_ω, Proposition 21 yields the following.

LEMMA 22. *Let ω be a nonnull state. Then $\Gamma_R(p,\omega)$ has a value, which is* cav $\widetilde{u}(p,\omega)$.

We denote by Γ_R^V the stochastic game with vector payoffs in which the controller is restricted to stay actions.

3.6. Functional equations. Let $\mathcal{B}$ denote the set of functions $\phi\colon \Delta(K)\times\Omega \to [0,1]$ that are 1-Lipshitz with respect to p, when $\Delta(K)$ is endowed with the L_1-norm. We here define three operators on $\mathcal{B}$ that will be used to characterize the solutions of the game.

When transitions are controlled by player 1, we define the operator T_1 by

$$T_1\phi(p,\omega) = \text{cav } \max\left\{ \widetilde{u}, \max_{\omega' \in C_\omega, i \notin I_{\omega'}} \mathbf{E}\left[\phi \mid \omega', i\right]\right\}(p,\omega). \tag{10}$$

By convention, a maximum over an empty set is $-\infty$. In this expression, $\mathbf{E}\left[\phi \mid \omega', i\right]$ stands for the expectation of ϕ under $q(\cdot \mid \omega', i)$.

Note that $T_1\phi(p,\omega)$ is equal to $\text{cav } \max\left\{\text{cav } \widetilde{u}, \max_{\omega' \in C_\omega, i \notin I_{\omega'}} \mathbf{E}\left[\phi \mid \omega', i\right]\right\}(p,\omega)$ as well.

When transitions are controlled by player 2, we define the operators T_2 and T_3 by

$$T_2\phi(p,\omega) = \text{cav} \min\left\{\widetilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E}\left[\phi \mid \omega', j\right]\right\}(p,\omega),$$

$$T_3\phi(p,\omega) = \min\left\{\text{cav } \widetilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E}\left[\phi \mid \omega', j\right]\right\}(p,\omega).$$

Since the maximum (or minimum) of a finite number of elements of $\mathcal{B}$ belongs to $\mathcal{B}$, and since by Lemma 8 concavification preserves Lipshitz properties when $\Delta(K)$ is endowed with the L_1-norm, all three operators T_1, T_2, and T_3 map $\mathcal{B}$ into $\mathcal{B}$. Note that for each $i = 1, 2, 3$ the operator T_i is monotonic: $\phi_1 \leq \phi_2$ implies $T_i\phi_1 \leq T_i\phi_2$. Moreover, for every $\phi \in \mathcal{B}$, $T_i\phi$ is constant over C_ω, for each $\omega \in \Omega$.

We now assume that player 1 controls transitions, and prove a few results on T_1. When transitions are controlled by player 2, identical results hold for both T_2 and T_3. Since the proofs are similar, they are omitted.

Proposition 23.

1. *T_1 has a unique fixed point ϕ.*
2. *The sequences (ϕ_n^0) and (ϕ_n^1) defined by $\phi_0^j = j$, $\phi_{n+1}^j = T_1\phi_n^j$ for $j = 0, 1$, are monotonic and converge uniformly to ϕ.*
3. *ϕ coincides with* cav $\widetilde{u}$ *on Ω_0.*
4. *If $f \in \mathcal{B}$ satisfies $f \leq T_1 f$ (resp., $f \geq T_1 f$), then $f \leq \phi$ (resp., $f \geq \phi$).*

Since $T_1\phi$ and $T_2\phi$ are concave for every $\phi \in \mathcal{B}$, the fixed points of those operators are concave functions. Since 0 is concave, and since T_3 maps concave functions to concave functions, the analog of Proposition 23 for T_3 implies that the fixed point of T_3 is concave as well.

Proof. By monotonicity of T_1, item 2 follows from item 1. Since cav $\widetilde{u}(p,\omega)$ is constant on every communicating set, so is $T_1\phi(p,\omega)$ for every $\phi \in \mathcal{B}$. Since $I_\omega = I$ for every $\omega \in \Omega_0$, $T_1\phi(p,\omega) = \text{cav } \widetilde{u}(p,\omega)$ for every $\phi \in \mathcal{B}$, every $\omega \in \Omega_0$, and every $p \in \Delta(K)$. Thus, item 3 will follow from item 1. We now prove item 1. By Ascoli's characterization, $\mathcal{B}$ is a compact metric space when endowed with the L_∞-norm. By Lemma 9, T_1 is nonexpansive, so that it is continuous on $\mathcal{B}$. Hence T_1 has a fixed point.

We prove uniqueness by contradiction. Let ϕ_1 and ϕ_2 be two distinct fixed points of T_1, and assume w.l.o.g. that $\delta := \max_{(p,\omega)\in\Delta(K)\times\Omega}(\phi_1(p,\omega) - \phi_2(p,\omega)) > 0$. Let

$$D = \{\omega \in \Omega, \phi_1(p,\omega) - \phi_2(p,\omega) = \delta \text{ for some } p \in \Delta(K)\}$$

contain those states where the difference is maximal. Since both $\phi_1(p,\cdot)$ and $\phi_2(p,\cdot)$ are constant on each communicating set, $C_\omega \subseteq D$ whenever $\omega \in D$.

Since $\phi_1 = \phi_2$ on Ω_0, $D \subseteq \Omega \setminus \Omega_0$. Let $\omega \in D$ be given, and let $p_0 \in \Delta(K)$ be an extreme point of the convex hull of the set $\{p \in \Delta(K) : \phi_1(p,\omega) - \phi_2(p,\omega) = \delta\}$.

Thus, $\phi_1(p_0,\omega) - \phi_2(p_0,\omega) = \delta > 0$. Since $\phi_1(\cdot,\omega)$ and $\phi_2(\cdot,\omega)$ are concave, it also follows that $(p_0, \phi_1(p_0,\omega))$ is an extreme point of the hypograph of the concave function $\phi_1(\cdot,\omega)$. This implies

$$\phi_1(p_0,\omega) = \max\left\{\text{cav } \widetilde{u}, \max_{\omega' \in C_\omega, i \notin I_\omega} \mathbf{E}\left[\phi_1 | \omega', i\right]\right\}(p_0,\omega).$$

Since $\phi_1(p_0,\omega) > \phi_2(p_0,\omega) \geq \text{cav } \ \widetilde{u}(p_0,\omega)$, one has $\phi_1(p_0,\omega) = \mathbf{E}\left[\phi_1(p_0,\cdot) \mid \omega', i\right]$ for some $\omega' \in C_\omega$ and $i \notin I_{\omega'}$. Since $T_1\phi_2 = \phi_2$, $\phi_2(p_0,\omega) \geq \mathbf{E}\left[\phi_2(p_0,\cdot) \mid \omega', i\right]$, and therefore

$$\delta = \phi_1(p_0,\omega) - \phi_2(p_0,\omega) \leq \mathbf{E}\left[\phi_1(p_0,\cdot) - \phi_2(p_0,\cdot) \mid \omega', i\right].$$

By the definition of D, this implies that $q(D \mid \omega', i) = 1$.

Thus, for every $\omega \in D$ there exists $\omega' \in C_\omega$ and $i \notin I_{\omega'}$ that satisfy $q(D \mid \omega', i) = 1$. This implies the existence of $\omega_1, \omega_2 \in D$ such that $C_{\omega_1} \neq C_{\omega_2}$ and $\omega_1 \leftrightarrow \omega_2$, a contradiction. This proves 1.

To prove 4, we assume that $\delta = \max_{(p,\omega)\in\Delta(K)\times\Omega}\left(f(p,\omega) - \phi(p,\omega)\right) > 0$, and repeat the second part of the proof of 1 to obtain a contradiction. □

4. Incomplete information on one side.

4.1. Preliminaries. We here single out a useful lemma. The lemma concerns a standard two-player zero-sum stochastic game G and its version G_R in which player 1 is restricted to stay actions. Thus, K is a singleton.

LEMMA 24. *Let G be a two-player zero-sum stochastic game with transitions controlled by player 1, and let $\omega \in \Omega$. If player 2 can guarantee that $\alpha \in \mathbf{R}$ in $G_R(\omega)$ and he can guarantee that $\phi : \Omega \to \mathbf{R}$ in G, then he can also guarantee* $\max\{\alpha, \max_{\omega' \in C_\omega, i \notin I_{\omega'}} \mathbf{E}[\phi|\omega', i]\}$ *in $G(\omega)$.*

Proof. By Lemma 20 player 2 can guarantee α in $G_R(\omega')$ for every $\omega' \in C_\omega$. Let τ_1 be a strategy that guarantees $\alpha + \varepsilon$ in $G_R(\omega')$ for every $\omega' \in C_\omega$, and let τ_2 be a strategy that guarantees $\phi + \varepsilon$ in G. Let $N \in \mathbf{N}$ be such that for every $n \geq N$, every $\omega' \in C_\omega$, and every σ in $G_R(\omega)$, $\gamma_n(\omega', \sigma, \tau_1) \leq \alpha + \varepsilon$, and for every σ in G, $\gamma_n(\omega', \sigma, \tau_2) \leq \phi(\omega') + \varepsilon$.

Define $\nu = 1 + \inf\{n \geq 1, i_n \notin I_{\omega_n}\}$. Define a strategy τ for player 2 as follows.

- Until stage ν, τ plays in blocks of size N (the last block may be shorter). In block $l \geq 0$, where $lN < \nu$, τ forgets past play and follows $\tau_1(\omega_{lN+1})$ for N stages.
- At stage ν, τ forgets past play and starts following τ_2.

Let σ be an arbitrary pure strategy. We will compute an upper bound on $\mathbf{E}_{\omega,\sigma,\tau}\left[\overline{g}_n\right]$ for n sufficiently large. Set $L_* = \lceil \frac{\ln \varepsilon}{\ln(1-\varepsilon)} \rceil$[5] and take $n \geq N_1 := \lceil L_* N/\varepsilon^2 \rceil$. Denote by $\overline{g}_{m_1 \to m_2}$ the average payoff from stage m_1 to stage m_2. With $\theta^* := N \times \lceil \frac{\nu}{N} \rceil$, since payoffs are nonnegative one has

$$\overline{g}_n \leq \frac{\theta^*}{n}\overline{g}_{\theta^*} + \frac{n+1-\nu}{n}\overline{g}_{\nu \to n}. \tag{11}$$

On the event $\nu \leq n - N$, one has

$$\mathbf{E}_{\omega,\sigma,\tau}\left[\overline{g}_{\nu \to n} | \mathcal{H}_\nu\right] = \mathbf{E}_{\omega_\nu,\sigma_\nu,\tau_2}\left[\overline{g}_{n-\nu+1}\right] \leq \phi(\omega_\nu) + \varepsilon, \tag{12}$$

[5]For every real c, $\lceil c \rceil$ is the smallest integer larger than or equal to c.

where σ_ν is the strategy induced by σ after ν. Since σ is pure, $\nu - 1$ is a stopping time and, using (12),

$$\begin{aligned}\mathbf{E}_{\omega,\sigma,\tau}\left[\overline{g}_{\nu\to n}|\mathcal{H}_{\nu-1}\right] &\le \mathbf{E}\left[\phi|\omega_{\nu-1}, i_{\nu-1}\right] + \varepsilon \\ &\le \max\nolimits_{\omega'\in C_\omega, i\notin I_{\omega'}} \mathbf{E}\left[\phi|\omega', i\right] + \varepsilon.\end{aligned} \tag{13}$$

On the other hand, on the event $\nu > n - N$,

$$\frac{n+1-\nu}{n} \le \varepsilon. \tag{14}$$

We now proceed to the first term in the decomposition (11) of $\overline{g}_n$. For each l, we let $\pi_l = \mathbf{P}_{\omega,\sigma,\tau}\left(\nu \le (l+1)N \mid \mathcal{H}_{lN+1}\right)$. By the choice of N,

$$\mathbf{E}_{\omega,\sigma,\tau}\left[\overline{g}_{lN+1\to(l+1)N}|\mathcal{H}_{lN+1}\right] \le \alpha + \varepsilon \ + \mathbf{P}_{\omega,\sigma,\tau}\left(\nu \le (l+1)N \mid \mathcal{H}_{lN+1}\right)$$

on the event $lN + 1 < \nu$. By taking expectations, this yields

$$\mathbf{E}_{\omega,\sigma,\tau}\left[\overline{g}_{lN+1\to(l+1)N}\mathbf{1}_{lN+1<\nu}\right] \le (\alpha+\varepsilon)\mathbf{P}_{\omega,\sigma,\tau}(lN+1<\nu) + \mathbf{P}_{\omega,\sigma,\tau}((l+1)N \ge \nu).$$

By summation over l, and using the definition of θ^*, this yields

$$\mathbf{E}_{\omega,\sigma,\tau}\left[\sum_{l=0}^{\theta^*-1}\overline{g}_{lN+1\to(l+1)N}\right] \le (\alpha+\varepsilon)\mathbf{E}_{\omega,\sigma,\tau}\left[\theta^*\right] + 1,$$

hence

$$\mathbf{E}_{\omega,\sigma,\tau}\left[\frac{N\theta^*}{n}\overline{g}_{N\theta^*}\right] \le (\alpha+\varepsilon)\mathbf{E}_{\omega,\sigma,\tau}\left[\frac{N\theta^*}{n}\right] + \frac{N}{n}. \tag{15}$$

The result follows by (11), (13), (14), and (15). $\square$

We shall need a variant of the previous result whose proof is identical to the previous proof. Consider the stochastic game with incomplete information $\Gamma(p,\omega)$ where ω is a nonnull state and assume that transitions are controlled by player 2. Assume that player 1 can guarantee a function ϕ. Then player 1 can also guarantee $\min\left\{\widetilde{u}, \min_{\omega'\in C_\omega, j\notin J_{\omega'}} \mathbf{E}\left[\phi|\omega', j\right]\right\}(p,\omega)$ in $\Gamma(p,\omega)$.

4.2. Transitions controlled by player 1. In this section we assume that transitions are controlled by player 1.

4.2.1. Existence of the value.

PROPOSITION 25. *The unique fixed point of* T_1 *is the value of* Γ.

Proof. Let ϕ be the unique fixed point of T_1, and fix $\epsilon > 0$ once and for all.

Step 1. Player 1 can guarantee ϕ in Γ. By Lemma 22 player 1 can guarantee cav $\widetilde{u}$. Set $\phi_0^0 = 0$, and, for $n \ge 0$, define $\phi_{n+1}^0 = T_1\phi_n^0$. Assume that player 1 can guarantee ϕ_n^0 for some $n \in \mathbf{N}$. Let $p \in \Delta(K)$ and $\omega \in \Omega$ be given. Plainly, for every $\omega' \in C_\omega$ and every $i \notin I_\omega$, player 1 can guarantee $\mathbf{E}\left[\phi_n^0 \mid \omega', i\right](p,\omega)$ in $\Gamma(p,\omega')$; first he plays the action i at ω', and then a strategy that guarantees $\phi_n^0(p,\cdot)$ (up to ε). By Lemma 20, he can guarantee $\mathbf{E}\left[\phi_n^0 \mid \omega', i\right](p,\omega)$ in $\Gamma(p,\omega)$. By Lemmas 3 and 10 he can guarantee $T_1\phi_n^0 = \phi_{n+1}^0$ in Γ. Since player 1 can guarantee $\phi_0^0 = 0$, and since $\lim_{n\to\infty}\phi_n^0 = \phi$, the result follows.

We now prove that player 2 can guarantee ϕ.

Step 2. Definition of approachable sets. For $\omega \in \Omega$, let $\mathcal{B}_\omega$ be the set of vectors approachable in Γ^V by player 2 at ω. We also define

$$\mathcal{A}_\omega = \left\{a \in \mathbf{R}^K : \langle a, p\rangle \geq \text{cav } \widetilde{u}(p,\omega) \text{ for every } p\right\}.$$

By Proposition 14 and Lemma 22, $\mathcal{A}_\omega$ is the set of vectors approachable by player 2 at ω in the stochastic game with vector payoffs Γ_R^V. Both sets $\mathcal{A}_\omega$ and $\mathcal{B}_\omega$ are nonempty, closed, convex, and upwards comprehensive.

For every $\omega \in \Omega$ define

$$\mathcal{D}_\omega = \left\{d = \max\left\{a, \max_{\omega' \in C_\omega, i \notin I_{\omega'}} \mathbf{E}\left[b(\cdot) \mid \omega', i\right]\right\} : a \in \mathcal{A}_\omega, b(\omega'') \in \mathcal{B}_{\omega''} \text{ for every } \omega'' \in \Omega\right\}.$$

Step 3. $\mathcal{D}_\omega \subseteq \mathcal{B}_\omega$. Fix $d \in \mathcal{D}_\omega$. Let τ_1 be a strategy that approaches $a + \varepsilon$ at ω, and let τ_2 be a strategy that approaches $b(\omega'') + \varepsilon$ at each state ω''. For each k the strategy τ_1 guarantees $a^k + \varepsilon$ in the game $\Gamma(k,\omega)$, and τ_2 has a similar property. By Lemma 24, applied independently to each G^k, the strategy obtained by concatenation of τ_1 and τ_2 guarantees $\max\left\{a^k, \max_{\omega' \in C_\omega, i \notin I_{\omega'}} \mathbf{E}\left[b^k(\cdot) \mid \omega', i\right]\right\} + 3\varepsilon = d^k + 3\varepsilon$ in G^k. Lemma 13 implies that $d \in \mathcal{B}_\omega$.

Step 4. Player 2 can guarantee ϕ. Let $f(p,\omega) = \inf_{a \in \mathcal{B}_\omega} \langle a, p\rangle$ and $h(p,\omega) = \inf_{a \in \mathcal{D}_\omega} \langle a, p\rangle$, so that by Step 3 $f \leq h$. By Lemma 13 player 2 can guarantee $\langle a, p\rangle$ in $\Gamma(p,\omega)$ for every $a \in \mathcal{B}_\omega$. Therefore he can guarantee $f(p,\omega)$ as well. By Lemma 11, the definition of $\mathcal{D}_\omega$ may be rephrased as

$$h = \text{cav } \max\left\{\text{cav } \widetilde{u}, \max_{\omega' \in C_\omega, i \notin I_{\omega'}} \mathbf{E}\left[f \mid \omega', i\right]\right\} = T_1 f.$$

Thus, $f \leq T_1 f$. By item 4 in Proposition 23, $f \leq \phi$. Therefore, player 2 can guarantee ϕ. □

4.2.2. Optimal strategies. The proof of Proposition 25 yields no information on ε-optimal strategies for player 1. We argue here that player 1 has an optimal strategy σ of a simple type. In effect, σ has the following structure. Whenever the play enters a communicating set, say at stage $n \geq 0$, σ randomly selects a nonrevealing stationary strategy that is used until the play moves to a new communicating set, if ever. The random choice of the stationary strategy may itself be revealing, in that the distribution used at stage n to select a stationary strategy depends both on p_n and on k. We describe below such a strategy in more detail.

Let v be the value of the game. Let $(p,\omega) \in \Delta(K) \times \Omega$ be given. Upon entering a communicating set at stage $n \geq 0$, player 1 computes $v(p_n) = \text{cav } \max\{\widetilde{u}, \max_{\omega' \in C_{\omega_n}, i \notin I_{\omega'}} \mathbf{E}[v \mid \omega', i]\}(p_n, \omega_n)$, and performs a state-dependent lottery, as described in the proof of Lemma 10. To be specific, one determines $\widetilde{p}_e \in \Delta(K), \alpha_e \in [0,1]$, for $e = 1, \ldots, |K|+1$, such that $\sum_e \alpha_e = 1$, $\sum_e \alpha_e \widetilde{p}_e = p_n$, and

$$v(p_n) = \sum_e \alpha_e \max\left\{\widetilde{u}, \max_{\omega' \in C_{\omega_n}, i \notin I_{\omega'}} \mathbf{E}\left[v \mid \omega', i\right]\right\}(\widetilde{p}_e, \omega_n).$$

If G^k is the actual game that is played, player 1 chooses e according to a state-dependent lottery μ^k, where $\mu^k(e) = \alpha_e \widetilde{p}_e^k / p_n^k$.

If $\max\left\{\widetilde{u}, \max_{\omega' \in C_{\omega_n}, i \notin I_{\omega'}} \mathbf{E}\left[v \mid \omega', i\right]\right\}(p^e, \omega_n) = \widetilde{u}(p^e, \omega_n)$, player 1 plays a stationary (nonrevealing) strategy which guarantees $\widetilde{u}(p^e, \omega_n)$ in the restricted game $\widetilde{\Gamma}_R(p^e, \omega_n)$. Recall that there are finite-stage algorithms that compute this strategy.

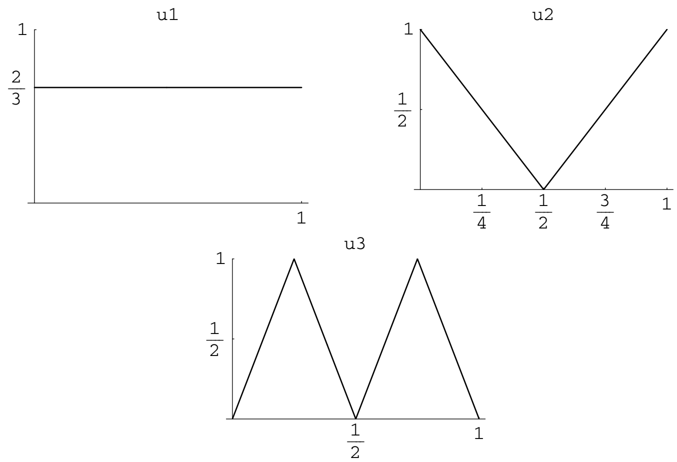

FIG. 1. *The value of the restricted games* $\widetilde{\Gamma}_R(p, \omega_i)$.

If, on the other hand, $\max\left\{\widetilde{u}, \max_{\omega' \in C_{\omega_n}, i \notin I_{\omega'}} \mathbf{E}\left[v \mid \omega', i\right]\right\}(p^e, \omega_n) = \mathbf{E}[v(p^e, \cdot) \mid \omega', i]$ for some $\omega' \in C_{\omega_n}$ and $i \notin I_{\omega'}$, player 1 plays the stationary strategy $x_{C_{\omega_n}}$ until the play reaches ω', and at ω' he plays the action i. He then recursively switches to a strategy that guarantees $v(p^e, \cdot)$.

Under σ, player 1 will end up in finite time playing an optimal stationary strategy in some restricted game $\widetilde{\Gamma}_R(p', \omega')$, with $p' \in \Delta(K)$ and $\omega' \in \Omega$. It can be checked that σ guarantees $v(p, \omega) - \varepsilon$ for every $\varepsilon > 0$. In that sense, σ is optimal. The proof is standard and therefore omitted.

4.2.3. An example. We here provide a simple example that illustrates the basic issues of splitting and information revelation. In particular, in this example the informed player will perform two state-dependent lotteries and therefore reveal information in two different stages of the game, unlike what happens in standard repeated games with incomplete information. The game has three states $\Omega = \{\omega_1, \omega_2, \omega_3\}$, where ω_2 and ω_3 are absorbing. There are two possible payoff functions, so that $K = \{1, 2\}$. A distribution over K is identified with the probability $p \in [0, 1]$ assigned to $k = 2$.

We first describe the main features of the example before providing the payoff and transition matrices of the game.

All actions of player 1 at state ω_1 are stay actions, except one, which leads to either state ω_2 and ω_3 with equal probability.

The value $u_i(p)$ of the restricted game $\tilde{\Gamma}_R(p, \omega_i)$ is given by (see Figure 1)

$$u_1(p) = 2/3,$$

$$u_2(p) = \max\{1 - 2p, 2p - 1\},$$

$$u_3(p) = \begin{cases} 4p & 0 \le p \le 1/4, \\ 2 - 4p & 1/4 \le p \le 1/2, \\ 4p - 2 & 1/2 \le p \le 3/4, \\ 4 - 4p & 3/4 \le p \le 1. \end{cases}$$

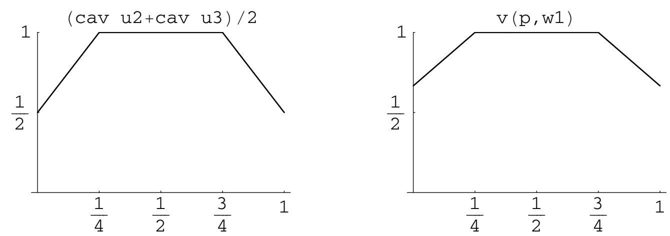

FIG. 2.

Note that $1/2(\text{cav } u_2 + \text{cav } u_3)$ is given by (see Figure 2)

$$1/2(\text{cav } u_2 + \text{cav } u_3)(p) = \begin{cases} 2p + 1/2 & 0 \le p \le 1/4, \\ 1 & 1/4 \le p \le 3/4, \\ 5/2 - 2p & 3/4 \le p \le 1. \end{cases}$$

It is the payoff that is guaranteed by player 1, when starting at (p, ω_1), and exiting from ω_1 without revealing any information. Indeed, once in ω_2 or in ω_3, the game will stay there and therefore the value is given by cav u_2 and cav u_3, respectively.

By section 4.2, the value $v(p, \omega_1) = \text{cav } \max\left\{\frac{2}{3}, \frac{1}{2}(\text{cav } u_2 + \text{cav } u_3)\right\}$ is given by (see Figure 2)

$$v(p, \omega_1) = \begin{cases} (2 + 4p)/3 & 0 \le p \le 1/4, \\ 1 & 1/4 \le p \le 3/4, \\ 2 - 4p/3 & 3/4 \le p \le 1. \end{cases}$$

Assume that the game starts in state ω_1, with $p = 1/8$. Note that $v(1/8, \omega_1) = \frac{1}{2}u_1(0) + \frac{1}{2}(\text{cav } u_2(1/4) + \text{cav } u_3(1/4))$, and that cav $u_3(1/4) = \frac{3}{4}u_3(0) + \frac{1}{4}u_3(1)$, while cav $u_2(1/4) = u_2(1/4)$.

The optimal strategy described in section 4.2.2 is as follows. Player 1 starts by tossing a state-dependent coin. If the coin comes up heads, player 1 plays forever an optimal stationary strategy in the game $k = 1$. If the coin comes up tails, player 1 first plays the nonstay action. The game then moves with equal probability to states ω_2 and ω_3. In the former case, player 1 continues with an optimal nonrevealing stationary strategy in the average game $\tilde{\Gamma}_R(1/4, \omega_2)$. In the latter case, player 2 again tosses a (degenerate) state-dependent coin. If $k = 1$ (resp., $k = 2$), player 1 continues with an optimal strategy in the game $\tilde{\Gamma}_R(0, \omega_3)$ (resp., $\tilde{\Gamma}_R(1, \omega_3)$).

Note that the amount of information revealed by player 1 depends on the actual play.

To complete the example, we provide in Figure 3 payoff matrices that satisfy the required specifications. The vertical arrows that appear in the bottom row of the two top matrices stand for the random transition to either ω_2 or ω_3.

In both states ω_1 and ω_2, player 2 is a dummy. The incomplete information game that corresponds to the game of state ω_3 coincides with example 1.3 in Zamir [24].

4.3. Transitions controlled by player 2. In this section we assume that transitions are controlled by player 2. We prove that both the min-max value and the max-min value exist, but that they may differ.

4.3.1. The max-min value.

LEMMA 26. *The unique fixed point of T_2 is the max-min value of Γ .*

	$k=1$			$k=2$			
	j_1	j_2	j_3	j_1	j_2	j_3	
T	2/3	2/3	2/3	2/3	2/3	2/3	ω_1
B	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	
	j_1	j_2	j_3	j_1	j_2	j_3	
T	1	1	1	-1	-1	-1	ω_2
B	-1	-1	-1	1	1	1	
	j_1	j_2	j_3	j_1	j_2	j_3	
T	4	0	2	0	4	-2	ω_3
B	4	0	-2	0	4	2	

FIG. 3. *The payoff matrices.*

Proof. Let ϕ be the unique fixed point of T_2, and fix $\varepsilon > 0$.

Step 1. Player 1 can guarantee ϕ. Set ϕ_0^0 and, for $m \geq 0$, set $\phi_{m+1}^0 = T_2\phi_m^0$. Assume that player 1 can guarantee ϕ_m^0 for some $m \in \mathbf{N}$. By the remark following Lemma 24, player 1 can guarantee $\min\left\{\widetilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E}\left[\phi_m^0 \mid \omega', j\right]\right\}$. Hence player 1 can also guarantee cav $\min\left\{\widetilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E}\left[\phi_m^0 \mid \omega', j\right]\right\} = \phi_{m+1}^0$. Since player 1 can guarantee $\phi_0^0 \equiv 0$, and since $\phi = \lim_{m\to\infty} \phi_m^0$, the result follows.

We now prove that player 2 can defend ϕ. Assume that player 2 can defend ϕ_m^1 for some $m \in \mathbf{N}$, and let σ be an arbitrary strategy of player 1. We prove in Steps 2 and 3 below that in this case player 2 can defend ϕ_{m+1}^1. Since $\phi = \lim_{m\to\infty} \phi_m^1$, and since player 2 can defend $\phi_0^1 \equiv 1$, he can defend ϕ as well.

Step 2. Definition of a reply. Given (p, ω), we let $\tau_1(p, \omega)$ be a (stationary) strategy that guarantees $\widetilde{u}(p, \omega) + \varepsilon$ in $\widetilde{\Gamma}_R(p, \omega)$. Choose $N_1 \in \mathbf{N}$ such that $\gamma_n(p, \omega, \widetilde{\sigma}, \tau_1(p, \omega)) \leq \widetilde{u}(p, \omega) + 2\varepsilon$ for every $n \geq N_1$ and every nonrevealing strategy $\widetilde{\sigma}$ of player 1.

By the remark that follows Definition 1, N_1 can be chosen independently of (p, ω). Let $\widetilde{\mathcal{T}}$ be the set of strategies of player 2 in $\widetilde{\Gamma}_R(p, \omega)$, and let $\widetilde{\tau} \in \widetilde{\mathcal{T}}$ be an $\varepsilon^2/32N_1^2$-exhausting information strategy given σ and (p, ω). Choose $N \in \mathbf{N}$ such that

$$\mathbf{E}_{p,\omega,\sigma,\widetilde{\tau}}\left[\sum_{n=N}^{+\infty} \|p_n - p_{n+1}\|_2^2\right] \leq \frac{\varepsilon^2}{32N_1^2}.$$

By Lemma 16,

$$\mathbf{E}_{p,\omega,\sigma,\widetilde{\tau}}\left[\sum_{n=N}^{+\infty} \|p_n - p_{n+1}\|_1\right] \leq \frac{\varepsilon}{\sqrt{32}N_1} \leq \frac{\varepsilon}{4}. \tag{16}$$

We define τ as follows.

- Play $\widetilde{\tau}$ up to stage N.

- At stage N compute $\beta_N := \min\left\{\widetilde{u}, \min_{\omega'\in C_\omega, j\notin J_{\omega'}} \mathbf{E}\left[\phi^1_m \mid \omega', j\right]\right\}(p_N, \omega_N)$.
  - If $\beta_N = \widetilde{u}(p_N, \omega_N)$, play by successive blocks of length N_1: in the $b+1$th block play the strategy $\tau_1(p_{N+bN_1}, \omega_{N+bN_1})$.
  - Otherwise, switch to a strategy that defends the quantity $\min_{\omega'\in C_\omega, j\notin J_{\omega'}} \mathbf{E}\left[\phi^1_m \mid \omega', j\right](p_N, \omega_N) + \varepsilon$ against σ_N, where σ_N is the strategy induced by σ after stage N.

Step 3. The computation. We here prove that τ defends $\phi^1_{m+1}(p, \omega) + 8\sqrt{\varepsilon}$ in $\Gamma(p, \omega)$. We abbreviate $\mathbf{E}_{p,\omega,\sigma,\tau}$ to $\mathbf{E}$. First, we provide an upper bound on the average payoff $\mathbf{E}\left[\overline{g}_{N\to N+n-1}|\mathcal{H}_N\right]$ between stages N and $N+n$ on the event

$$A := \{\beta_N = \widetilde{u}(p_N, \omega_N)\}\,. \tag{17}$$

First take $n = N_1$. By definition,

$$\mathbf{E}\left[\overline{g}_{N\to N+N_1-1}|\mathcal{H}_N\right] = \mathbf{E}_{p_N,\omega_N,\sigma_N,\tau_1(p_N,\omega_N)}\left[\overline{g}_{N_1}\right].$$

By the choice of N_1,

$$\mathbf{E}_{p_N,\omega_N,\sigma_N^{p_N},\tau_1(p_N,\omega_N)}\left[\overline{g}_{N_1}\right] \le \widetilde{u}(p_N, \omega_N) + 2\varepsilon. \tag{18}$$

On the other hand, by Lemma 17,

$$\begin{aligned}&\left|\mathbf{E}_{p_N,\omega_N,\sigma_N^{p_N},\tau_1(p_N,\omega_N)}\left[\overline{g}_{N_1}\right] - \mathbf{E}_{p_N,\omega_N,\sigma_N,\tau_1(p_N,\omega_N)}\left[\overline{g}_{N_1}\right]\right|\\ &\qquad \le 4\mathbf{E}_{p_N,\omega_N,\sigma_N,\tau_1(p_N,\omega_N)}\left[\sum_{m=1}^{N_1}\|p_m - p_{m+1}\|_1\right].\end{aligned}$$

Thus, using (18),

$$\mathbf{E}\left[\overline{g}_{N\to N+N_1-1}|\mathcal{H}_N\right] \le \widetilde{u}(p_N, \omega_N) + 2\varepsilon + 4\mathbf{E}\left[\sum_{m=N}^{N+N_1-1}\|p_m - p_{m+1}\|_1|\mathcal{H}_N\right].$$

The same computation applies to any block of N_1 stages. Specifically, for each $b \ge 0$,

$$\begin{aligned}\mathbf{E}\left[\overline{g}_{N+bN_1\to N+(b+1)N_1-1}|\mathcal{H}_{N+bN_1}\right] &\le \widetilde{u}(p_{N+bN_1}, \omega_{N+bN_1}) + 2\varepsilon\\ &\quad + 4\mathbf{E}\left[\sum_{m=N+bN_1}^{N+(b+1)N_1-1}\|p_m - p_{m+1}\|_1|\mathcal{H}_{N+bN_1}\right].\end{aligned}$$

Since $\widetilde{u}(p, \cdot)$ is constant on every communicating set, and since $\widetilde{u}(\cdot, \omega)$ is 1-Lipshitz, $\widetilde{u}(p_{N+bN_1}, \omega_{N+bN_1}) \le \widetilde{u}(p_N, \omega_N) + \|p_{N+bN_1} - p_N\|_1$. By taking expectations on the event A (defined by (17)), one gets, by Lemma 16, (16), and since $\|\cdot\|_1 \le \|\cdot\|_2$,

$$\begin{aligned}\mathbf{E}\left[\mathbf{1}_A\overline{g}_{N+bN_1\to N+(b+1)N_1-1}\right] &\le \mathbf{E}\left[\mathbf{1}_A\widetilde{u}(p_N, \omega_N)\right] + 2\varepsilon + \mathbf{E}\left[\mathbf{1}_A\|p_{N+bN_1} - p_N\|_2\right]\\ &\quad + 4\mathbf{E}\left[\mathbf{1}_A\sum_{m=N+bN_1}^{N+(b+1)N_1-1}\|p_m - p_{m+1}\|_2\right]\\ &\le \mathbf{E}\left[\mathbf{1}_A\widetilde{u}(p_N, \omega_N)\right] + 5\sqrt{\varepsilon}.\end{aligned}$$

By averaging over blocks, one obtains for every $n \geq \frac{2}{\varepsilon}(N+N_1)$

$$\mathbf{E}\left[\mathbf{1}_A \overline{g}_n\right] \leq \mathbf{E}\left[\mathbf{1}_A \widetilde{u}(p_N, \omega_N)\right] + 6\sqrt{\varepsilon}. \tag{19}$$

On the other hand, there is $N_2 \in \mathbf{N}$ such that for every $n \geq N_2$,

$$\mathbf{E}\left[\overline{g}_n | \mathcal{H}_N\right] \leq \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E}\left[\phi^1_m \mid \omega', j\right](p_N, \omega_N) + 2\varepsilon \text{ on the event } \overline{A}. \tag{20}$$

By taking expectations, (19) and (20) yield

$$\begin{aligned}\mathbf{E}\left[\overline{g}_n\right] &\leq \mathbf{E}\left[\min\left\{\widetilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E}\left[\phi^1_m \mid \omega', j\right]\right\}(p_N, \omega_N)\right] + 8\sqrt{\varepsilon} \\ &\leq \text{cav } \min\left\{\widetilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E}\left[\phi^1_m \mid \omega', j\right]\right\}(p, \omega) + 8\sqrt{\varepsilon}\end{aligned}$$

for every $n \geq \max\{N_2, \frac{2}{\varepsilon}(N+N_1)\}$. □

Let $\underline{v}$ denote the max-min of the game, and let (p, ω) be given. Similar to the discussion in section 4.2.2, it can be checked that there is a simple strategy for player 1 that guarantees $\underline{v}(p,\omega) - \varepsilon$ for each $\varepsilon > 0$. Under this strategy, player 1 chooses at random a nonrevealing stationary strategy whenever the play enters a communicating set, and uses it until the play moves to a new communicating set.

4.4. The min-max value.

LEMMA 27. *The unique fixed point of T_3 is the min-max value of Γ.*

Proof. Let ϕ be the unique fixed point of T_3, and fix $\varepsilon > 0$.

We first prove by induction that player 2 can guarantee ϕ. Set $\phi^1_0 \equiv 1$ and, for $m \geq 0$, set $\phi^1_{m+1} = T_3 \phi^1_m$. Assume that player 2 can guarantee ϕ^1_m for some $m \in \mathbf{N}$, and let (p, ω) be given. Plainly, for each $\omega' \in C_\omega, j \notin J_{\omega'}$, player 2 can guarantee $\mathbf{E}\left[\phi^1_m \mid \omega', j\right]$ in $\Gamma(p, \omega')$ by first playing j at ω', and then a strategy that guarantees ϕ^1_m (up to ε). By Lemma 20, he can guarantee $\mathbf{E}\left[\phi^1_m \mid \omega', j\right]$ in $\Gamma(p,\omega)$ as well. By Lemma 22, player 2 can guarantee cav $\widetilde{u}$. Thus, he can guarantee $T_3\phi^1_m = \phi^1_{m+1}$. Since he can guarantee ϕ^1_0, and since $\phi = \lim_{m\to\infty} \phi^1_m$, the result follows.

We now prove that player 1 can defend ϕ^0_m for each $m \in \mathbf{N}$. Clearly, player 1 can defend $\phi^0_0 \equiv 0$. Assume that player 1 can defend ϕ^0_m for some $m \in \mathbf{N}$. Let a strategy τ of player 2 and $(p, \omega) \in \Delta(K) \times \Omega$ be given. Set $\nu = 1 + \inf\{n \geq 1, j_n \notin J_{\omega_n}\}$. The supremum of $\mathbf{P}_{p,\omega,\sigma,\tau}(\nu < \infty)$ over all strategies σ coincides with the supremum over all nonrevealing strategies σ.[6] Denote by σ^* a *nonrevealing* strategy that achieves the supremum up to ε. We choose N such that $\mathbf{P}_{p,\omega,\sigma^*,\tau}(\nu > N) \leq \varepsilon$. The strategy σ^* thus exhausts the probability of leaving the initial communicating set. Denote by $\tau_{\min\{\nu,N\}}$ the strategy induced by τ after stage $\min\{\nu, N\}$.

On the event $\nu > N$, there is a strategy $\widetilde{\tau}$ in $\Gamma_R(p,\omega)$ such that $\|\mathbf{P}_{p,\omega_N,\sigma,\widetilde{\tau}} - \mathbf{P}_{p,\omega_N,\sigma,\tau_N}\| \leq \mathbf{P}_{p,\omega_N,\sigma,\tau_N}(\nu < +\infty)$ for every nonrevealing strategy σ in $\Gamma_R(p,\omega)$. This strategy depends on the history up to stage N.

We now define the reply σ of player 1 to τ as follows: play σ^* up to stage $\min\{\nu, N\}$.

- If $\nu > N$, switch to a strategy that defends cav $\widetilde{u}(p,\omega) + \varepsilon$ in $\Gamma_R(p, \omega_N)$ against $\widetilde{\tau}$.

[6] Indeed, for every strategy $\sigma = (\sigma^k)_k$, one has $\mathbf{P}_{p,\omega,(\sigma^k)_k,\tau}(\nu < +\infty) = \sum_k p_k \mathbf{P}_{k,\omega,\sigma^k,\tau}(\nu < +\infty) \leq \max_k \mathbf{P}_{k,\omega,\sigma^k,\tau}(\nu < +\infty)$. Let k_0 achieve the maximum, and let σ' be the nonrevealing strategy that plays σ_{k_0} irrespective of k. Since transitions are independent of k, one has $\mathbf{P}_{p,\omega,\sigma,\tau}(\nu < +\infty) \leq \mathbf{P}_{p,\omega,\sigma',\tau}(\nu < +\infty)$.

$k = 1$, ω_2:

	j_1	j_2	j_3	j_4	j_5
T	4	0	0	0	0
B	0	0	0	0	0

$k = 2$, ω_2:

	j_1	j_2	j_3	j_4	j_5
T	0	0	0	0	0
B	0	4	4	4	4

$k = 1$, ω_1:

	j_1	j_2	j_3	j_4	j_5
T	0	1	1	3	$\uparrow$
B	0	1	0	3	$\uparrow$

$k = 2$, ω_1:

	j_1	j_2	j_3	j_4	j_5
T	3	0	1	0	$\uparrow$
B	3	1	1	0	$\uparrow$

FIG. 4. *The payoff matrices.*

- If $\nu \leq N$, switch to a strategy that defends $\phi^0_m(p, \omega_\nu) + \varepsilon$ against τ_ν.

Since there are finitely many histories of length N, the set of strategies $(\tau_{\min\{\nu,N\}})$ is finite. It is straightforward to check that σ defends

$$\min\left\{\text{cav } \widetilde{u}, \min_{\omega' \in C_\omega, j \notin J_{\omega'}} \mathbf{E}\left[\phi^0_m \mid \omega', i\right]\right\}(p, \omega) + 2\varepsilon = \phi^0_{m+1}(p, \omega) + 2\varepsilon$$

against τ. □

4.5. An example. Here we provide an example where $\min\{\text{cav } f, g\}$ is strictly larger than $\text{cav } \min\{f, g\}$, so that the max-min value and the min-max value differ.

Consider the game depicted in Figure 4, where player 2 controls the transitions, and $|\Omega| = |K| = 2$, $|I| = 2$, and $|J| = 5$. The initial state is ω_1 (bottom two matrices). If in ω_1 player 2 chooses j_5, the game moves to ω_2, which is absorbing. If player 2 chooses another action in ω_1, the game remains in ω_1. Payoffs are as depicted in Figure 4 (the definition of $g^k(\omega_1, \cdot, j_5)$ is irrelevant).

Note that $I_{\omega_1} = \{j_1, j_2, j_3, j_4\}$, $\Omega_0 = \{\omega_2\}$, and $C_{\omega_1} = \{\omega_1\}$.

The game $\Gamma_R(p, \omega_1)$ is similar to Example 3.3 in Aumann and Maschler [2]. As calculated in Aumann and Maschler,

$$f(p) = \widetilde{u}(p, \omega_1) = \begin{cases} 3p & 0 \leq p \leq 2 - \sqrt{3}, \\ 1 - p(1-p) & 2 - \sqrt{3} \leq p \leq \sqrt{3} - 1, \\ 3(1-p) & \sqrt{3} - 1 \leq p \leq 1 \end{cases}$$

(see Figure 5). Note that $\text{cav } f \neq f$.

The game $\Gamma_R(p, \omega_2)$ is similar to the game presented in Aumann and Maschler [2, I.2], with all payoffs multiplied by 4.[7] As calculated in Aumann and Maschler,

$$g(p) = \widetilde{u}(p, \omega_2) = 4p(1-p).$$

As proved above, the max-min value when the initial state is ω_1 is $(\text{cav } \min\{f, g\})(p)$, while the min-max value is $\min\{\text{cav } f, g\}(p)$. A straightforward calculation shows that $\min\{\text{cav } f, g\}(1/2) = 3(2 - \sqrt{3})$ while $\text{cav } \min\{f, g\}(1/2) = 4/5$, so the two functions differ. The graphs of the two functions appear in Figure 6.

[7] We added the actions j_3, j_4, j_5, which do not change the calculation of the value. For our purposes, we could have multiplied all payoffs by any α, $3 < \alpha < 3/(\sqrt{3} - 1)$.

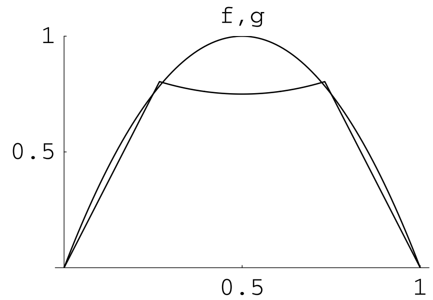

FIG. 5. *The value functions of the restricted games* $\widetilde{\Gamma}_R(p, \omega_i)$.

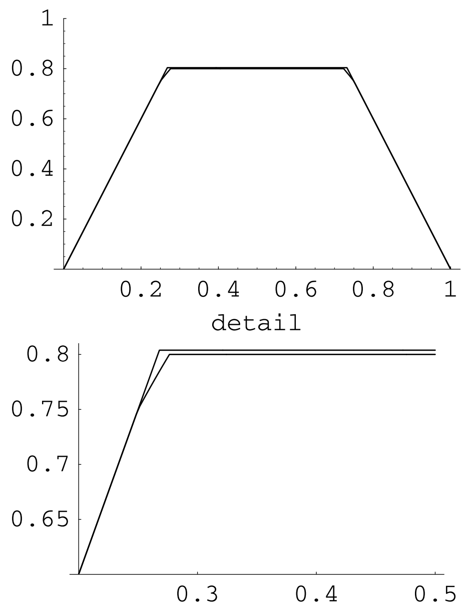

FIG. 6. *The functions* $\min\{\text{cav } f, g\}$ *and* $\text{cav } \min\{f, g\}$.

5. Incomplete information on both sides.

5.1. The model. We now extend our model to the case of incomplete information on both sides; that is, each player has some private information on the game that is to be played. Formally the model is extended as follows. For more details we refer to Mertens, Sorin, and Zamir [10] or Sorin [22].

A *two-player zero-sum stochastic game with incomplete information on both sides* is described by a finite collection $(G_{k,l})_{k \in K, l \in L}$ of stochastic games, together with a distribution $p \in \Delta(K)$ and a distribution $s \in \Delta(L)$. We assume that the games $G_{k,l}$ differ only through their reward functions $g^{k,l}$, but they all have the same sets of states Ω and actions I and J, and the same transition rule q.

The game is played in stages. At the outset of the game a pair $(k, l) \in K \times L$ is chosen according to $p \otimes s$. Player 1 is informed of k, and player 2 of l. At every stage n, the two players choose simultaneously actions $i_n \in I$ and $j_n \in J$, and ω_{n+1} is drawn according to $q(\cdot \mid \omega_n, i_n, j_n)$. Both players are informed of (i_n, j_n, ω_{n+1}).

W.l.o.g. we assume throughout this section that transitions are controlled by player 1. We will only sketch the proofs, since none of them involves any new idea.

5.2. Related literature. The main results in this framework are related to the case $|\Omega| = 1$ (repeated games with incomplete information) and are due to Aumann, Maschler, and Stearns [3] (see also Aumann and Maschler [2]) and Mertens and Zamir [11, 12]. As in the case of incomplete information on one side, we denote by $u(p, s)$ the value of the matrix game with action sets I and J and matrix payoff $(\sum_{k\in K,l\in L} p^k s^l g^{k,l}(i, j))_{i,j}$. Given $f : \Delta(K) \times \Delta(L) \to \mathbf{R}$, we let $\mathrm{cav}_p f$ denote the smallest function that is above f and concave in p, and $\mathrm{vex}_s f$ denotes the largest function that is below f and convex in s.

The min-max value of a repeated game with incomplete information exists and is equal to $\mathrm{vex}_s\mathrm{cav}_p u(p, s)$. The max-min value exists and is equal to $\mathrm{cav}_p\mathrm{vex}_s u(p, s)$.

5.3. Partitioning the states and the average restricted game. Since player 1 controls transitions, the partition defined in section 4 extends to this case, as well as the definition of the average restricted game $\widetilde{\Gamma}_R(p, s, \omega)$ in which none of the players has any information. Denote by $\widetilde{u}(p, s, \omega)$ the value of $\widetilde{\Gamma}_R(p, s, \omega)$. In addition, we define the average restricted game $\widetilde{\Gamma}^1_R(p, s, \omega)$ (resp., $\widetilde{\Gamma}^2_R(p, s, \omega)$) in which player 1 (resp., player 2) is informed of k (resp., l) while his opponent gets no information. Our first goal is to extend Proposition 21.

PROPOSITION 28. *For every* $(\omega, p, s) \in \Omega_0 \times \Delta(K) \times \Delta(L)$*, the min-max value of* $\Gamma(p, s, \omega)$ *exists and is equal to* $\mathrm{vex}_s\mathrm{cav}_p\widetilde{u}(p, s, \omega)$*. Similarly the max-min value exists and is equal to* $\mathrm{cav}_p\mathrm{vex}_s\widetilde{u}(p, s, \omega)$*.*

Proof. The proof follows the proof for repeated games with incomplete information, using the tools developed in the previous sections. We shall only sketch the arguments for the min-max value, and refer for details to Zamir [24].

First, we explain how player 2 can guarantee $\mathrm{vex}_s\mathrm{cav}_p\widetilde{u}(p, s, \omega)$. When player 2 ignores his information, he faces a game with incomplete information on one side with parameter set K and payoffs $\sum_{l\in L} s^l g^{k,l}$. By Proposition 21, player 2 can guarantee $\mathrm{cav}_p\widetilde{u}(p, s, \omega)$ in this game. Therefore by Lemma 10 (with the roles of the two players exchanged) he can also guarantee $\mathrm{vex}_s\mathrm{cav}_p\widetilde{u}(p, s, \omega)$.

To prove that player 1 can defend $\mathrm{vex}_s\mathrm{cav}_p\widetilde{u}(p, s, \omega)$, we adapt Zamir [24, Theorem 4.1]. Let τ be a given strategy of player 2. As in Step 2 of the proof of Lemma 26, we let player 1 play first an ε-exhausting strategy $\widetilde{\sigma}$ given τ. This strategy may be chosen to be nonrevealing (see, e.g., Sorin [22, Lemma IV.4.1]). Player 1 switches at some stage N to a strategy that defends $\mathrm{cav}_p\widetilde{u}(p, s_N, \omega_N)$ (up to ε) in $\Gamma(p, s_N, \omega_N)$ against the continuation strategy τ_N (see Step 3 of Lemma 26). Since $\widetilde{u}(\cdot, \cdot, \omega) = \widetilde{u}(\cdot, \cdot, \omega_N)$, $\mathrm{cav}_p\widetilde{u}(p, s_N, \omega_N) = \mathrm{cav}_p\widetilde{u}(p, s_N, \omega)$. Therefore player 1 defends $\mathbf{E}_{p,s,\omega,\widetilde{\sigma},\tau}[\mathrm{cav}_p\widetilde{u}(p, s_N, \omega)] \geq \mathrm{vex}_s\mathrm{cav}_p\widetilde{u}(p, s, \omega)$. □

5.4. The max-min value and the min-max value. Let $\mathcal{B}$ denote the set of all functions $\phi : \Delta(K) \times \Delta(L) \times \Omega \to [0, 1]$ that are 1-Lipshitz with respect to p and s. Denote by T_4 and T_5 the operators on $\mathcal{B}$ defined by

$$T_4\phi(p, s, \omega) = \mathrm{cav}_p \max\left\{\mathrm{cav}_p\mathrm{vex}_s\widetilde{u}, \max_{\omega'\in C_\omega, i\notin I_{\omega'}} \mathbf{E}[\phi \mid \omega', i]\right\}(p, s, \omega) \tag{21}$$

and

$$T_5\phi(p, s, \omega) = \mathrm{vex}_s\mathrm{cav}_p \max\left\{\mathrm{cav}_p\widetilde{u}, \max_{\omega'\in C_\omega, i\notin I_{\omega'}} \mathbf{E}[\phi \mid \omega', i]\right\}(p, s, \omega). \tag{22}$$

Our main result is the following.

THEOREM 29.

1. *The mappings T_4 and T_5 have unique fixed points, denoted respectively by $\underline{v}$ and $\overline{v}$.*
2. *The function $\underline{v}$ is the max-min value of the game.*
3. *The function $\overline{v}$ is the min-max value of the game.*

Note that if player 2 has no information, there is no vex operator in (21) and (22), and both T_4 and T_5 reduce to T_1. If player 1 has no information, there is no cav operator in (21) and (22), and T_4 and T_5 reduce respectively to T_3 and T_2 with the roles of the players reversed.

Proof. The first assertion follows the same lines as the proof of Proposition 23.

We now prove the second assertion. For $j = 0, 1$, we define the sequence $(\phi_n^j)_{n\geq 0}$ by $\phi_0^j = j$ and $\phi_{n+1}^j = T_4\phi_n^j$. We follow the inductive proof of Proposition 25, Step 1, or the first part of Lemma 27.

The sequence (ϕ_n^0) is increasing and converges uniformly to $\underline{v}$. It is clear that player 1 can guarantee ϕ_0^0. Assuming player 1 can guarantee ϕ_n^0, we prove that he can guarantee ϕ_{n+1}^0. By Lemma 10 it is sufficient to show that he can guarantee both $\mathrm{cav}_p\mathrm{vex}_s\widetilde{u}(p,s,\omega)$ and $\max_{\omega'\in C_\omega, i\notin I_{\omega'}}\mathbf{E}[\phi_n^0 \mid \omega', i]$, which is true by Proposition 28 and by Step 1 of Proposition 25.

To prove that player 2 can defend $\underline{v}$, we combine several ideas from the preceding sections. Let σ be given, and let $\widetilde{T}$ be the set of nonrevealing strategies of player 2. We let τ_σ be a nonrevealing strategy that ε-exhausts the information contained in σ, and choose N as in Step 2 of Lemma 26. Denote by $\nu = 1 + \min\{n \geq 1, i_n \notin I_{\omega_n}\}$. Player 2 plays according to τ_σ up to stage $\min\{\nu, N\}$.

- If $\nu \leq N$, from stage ν on he defends $\phi_n^1(p_\nu, s, \omega_\nu)$.
- If $\nu > N$, we first use the idea of Lemma 27, with the roles of the two players exchanged. Specifically, we define a nonrevealing strategy τ_N^σ that exhausts the probability of leaving the initial communicating set, given the strategy σ_N induced by σ after stage N. Choose N' such that $\mathbf{P}_{p_N,s,\omega_N,\sigma_N,\tau_N^\sigma}(\nu > N') \leq \varepsilon$. Player 2 plays τ_N^σ up to stage $\min\{\nu, N+N'\}$.
  - If $\nu \leq N+N'$, player 2 switches to a strategy that defends $\phi_n(p_\nu, s, \omega_\nu)+\varepsilon$.
  - If $\nu > N+N'$, following Steps 2 and 3 of Lemma 26, player 2 starts to play in blocks of length N_1. In the bth block he forgets past play and follows a strategy that defends $\mathrm{vex}_s\widetilde{u}(p_{N+N'+bN_1}, s, \omega_{N+N'+bN_1})$ in the restricted game $\widetilde{\Gamma}_R^2(p_{N+N'+bN_1}, s, \omega_{N+N'+bN_1})$ against the average continuation strategy $\sigma_{N+N'+bN_1}^{p_{N+N'+bN_1}}$ of player 1.

We now turn to the third assertion. We first prove that player 2 can guarantee $\overline{v}$. By following Steps 2, 3, and 4 of Lemma 25, one proves that player 2 guarantees $\mathrm{cav}_p\max\{\mathrm{cav}_p\widetilde{u},\ \max_{\omega'\in C_\omega, i\notin I_{\omega'}}\mathbf{E}[\overline{v} \mid \omega', i]\}(p,s,\omega)$. Hence, by Lemma 10 (with the roles of the two players exchanged), he can guarantee $\mathrm{vex}_s\mathrm{cav}_p\max\{\mathrm{cav}_p\widetilde{u},$ $\max_{\omega'\in C_\omega, i\notin I_{\omega'}}\mathbf{E}[\overline{v} \mid \omega', i]\} = \overline{v}$.

We now prove that player 1 can defend $\overline{v}$. We first follow Step 2 of Lemma 26. Given τ, we let σ^τ be a strategy in $\widetilde{\Gamma}_R^2(p,s,\omega)$ that exhausts the information contained in τ, and we choose N such that $\mathbf{E}_{p,s,\omega,\sigma^\tau,\tau}[\sum_{n=N}^\infty \|p_n - p_{n+1}\|_1^2] \leq \varepsilon$. Player 1 plays σ^τ up to stage N. He then switches to a strategy that guarantees $\mathrm{cav}_p\max\left\{\mathrm{cav}_p\widetilde{u}(\cdot, s_N, \omega), \max_{\omega'\in C_\omega, i\notin I_{\omega'}}\mathbf{E}[\overline{v} \mid \omega', i]\right\}(p, s_N)$ in $\widetilde{\Gamma}_R^1(p, s_N, \omega_N)$, as in the proof of Proposition 25. The result follows. $\square$

REFERENCES

[1] R. J. Aumann and M. B. Maschler, *Repeated games of incomplete information: The zero-sum extensive case*, in Report of the U.S. Arms Control and Disarmament Agency ST-143, Washington, D.C., 1968, Chapter III, pp. 37–116.

[2] R. J. Aumann and M. B. Maschler, *Repeated Games with Incomplete Information*, MIT Press, Cambridge, MA, 1995.

[3] R. J. Aumann, M. B. Maschler, and R. E. Stearns, *Repeated games of incomplete information: An approach to the non-zero sum case*, in Report of the U.S. Arms Control and Disarmament Agency ST-143, Washington, D.C., 1968, Chapter IV, pp. 117–216.

[4] D. Blackwell, *An analog of the minimax theorem for vector payoffs*, Pacific J. Math., 6 (1956), pp. 1–8.

[5] J. A. Filar, *Ordered field property for stochastic games when the player who controls transitions changes from state to state*, J. Optim. Theory Appl., 34 (1981), pp. 503–515.

[6] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 1988.

[7] R. Laraki, *On the regularity of the convexification operator on a compact set*, J. Convex Anal, to appear.

[8] J.-F. Mertens, *Repeated games*, Proceedings of the International Congress of Mathematicians, Berkeley, CA, 1986, A. M. Gleason, ed., American Mathematical Society, Providence, RI, 1987, pp. 1528–1577.

[9] J.-F. Mertens and A. Neyman, *Stochastic games*, Internat. J. Game Theory, 10 (1981), pp. 53–66.

[10] J.-F. Mertens, S. Sorin, and S. Zamir, *Repeated Games*, CORE Discussion Paper 9420-2, 1994.

[11] J.-F. Mertens and S. Zamir, *The value of two-person zero-sum repeated games with lack of information on both sides*, Internat. J. Game Theory, 1 (1971), pp. 39–64.

[12] J.-F. Mertens and S. Zamir, *Minmax and maxmin of repeated games with incomplete information*, Internat. J. Game Theory, 9 (1980), pp. 201–215.

[13] E. Milman, *Uniform Properties of Stochastic Games and Approachability*, Master Thesis, Tel Aviv University, 2000.

[14] T. Parthasarathy and T. E. S. Raghavan, *An orderfield property for stochastic games when one player controls transition probabilities*, J. Optim. Theory Appl., 33 (1981), pp. 375–392.

[15] T. E. S. Raghavan and J. A. Filar, *Algorithms for stochastic games—a survey*, Z. Oper. Res., 35 (1991), pp. 437–472.

[16] J. Renault, *The Value of Markov Chain Games with Lack of Information on One Side*, preprint.

[17] D. Rosenberg and N. Vieille, *The maxmin of recursive games with incomplete information on one side*, Math. Oper. Res., 25 (2000), pp. 23–35.

[18] K. W. Ross and R. Varadarajan, *Multichain Markov decision processes with a sample path constraint: A decomposition approach*, Math. Oper. Res., 16 (1991), pp. 195–207.

[19] N. Shimkin and A. Shwartz, *Guaranteed performance regions in Markovian systems with competing decision makers*, IEEE Trans. Automat. Control, 38 (1993), pp. 84–95.

[20] S. Sorin, *Big match with lack of information on one side* I, Internat. J. Game Theory, 13 (1984), pp. 201–255.

[21] S. Sorin, *Big match with lack of information on one side* II, Internat. J. Game Theory, 14 (1985), pp. 173–204.

[22] S. Sorin, *A First Course on Zero-Sum Repeated Games*, Mathematiques et Applications 37, Springer-Verlag, Berlin, 2002.

[23] S. Sorin and S. Zamir, *Big match with lack of information on one side* II, in Stochastic Games and Related Topics, T. E. S. Raghavan et al., eds., Kluwer, Dordrecht, The Netherlands, 1991, pp. 101–112.

[24] S. Zamir, *Repeated games of incomplete information: zero-sum*, in Handbook of Game Theory with Economic Applications, Vol. 1, R. J. Aumann and S. Hart, eds., Elsevier Science Publishers B.V., North-Holland, Amsterdam, 1992, pp. 109–154.