General properties of long-run supergames*

Jérôme Renault[†], Tristan Tomala[‡]

June 2011

Abstract

Supergames are repeated games in which a fixed known finite one-shot game is repeated over and over. Information about the actions chosen at each stage is provided by a signalling technology. This papers studies the main properties which are valid over this whole class of games, it both surveys known results and provides new ones.
Key words. Repeated Games, Signals, Folk Theorem.

1 Introduction

The theory of repeated games dates back to the fifties-sixties-seventies with the works of Shapley (1953), Aumann and Maschler (1965-66, re-edited in 1995), Aumann and Shapley (1976, re-edited in 1994) and Rubinstein (1977). These seminal papers have built a theory of dynamic games, where a set of players interact repeatedly and along the play, collect payoffs and information about the data of the game and about the behavior of their opponents. An important sub-class is the set of repeated games with complete information, also named supergames, where the players, the action sets and the payoff functions are known to all players and fixed through time. The first important result for supergames is the well known *Folk Theorem*[1] which characterizes the equilibrium outcomes of an infinitely repeated supergame with *patient players* and *perfect observation.* These assumptions mean that each player maximizes the limit average payoff and

*The authors gratefully thank Pierre Cardaliaguet, Marie Laclau, Ehud Lehrer, Xiaoxi Li and Sylvain Sorin for their comments.

[†]TSE (GREMAQ, Université Toulouse 1 Capitole), 21 allée de Brienne, 31000 Toulouse, France. e-mail: jerome.renault@tse-fr.eu. Jérôme Renault gratefully acknowledges the support of the Agence Nationale de la Recherche, under grant ANR-10-BLAN 0112.

[‡]HEC Paris, Economics and Decision Sciences Department, 1 rue de la Libération, 78351 Jouy-en-Josas Cedex, France. e-mail: tomala@hec.fr. Tristan Tomala gratefully acknowledges the support of the HEC foundation and of the Agence Nationale de la Recherche, under grant ANR-10-BLAN 0112.

[1]This result is generally attributed to Aumann and Shapley, 1976, and Rubinstein, 1977. It has been later generalized to discounted repeated games (Sorin, 1986, Fudenberg and Maskin, 1986) and to finitely repeated games (Benoit and Krishna, 1987).

that all players observe all the actions taken at each stage. The Folk Theorem then says that any feasible payoff vector is an equilibrium outcome, provided that it is individually rational. That is, each player gets at least his *minmax* level which is the payoff obtained when all other players are adversarial. The proof of this result is transparent. Players agree on which sequence of actions to take and unilateral deviations from the agreement are punished by minmax strategies. One sees directly how much this reasoning hinges on perfect observation. In the general model of supergames, actions need not be directly observed. Instead, there is a *signalling structure* modelled as a probability transition from actions to sets of signals, that describes the information obtained by the players through the play. These games are often called *repeated games with imperfect monitoring* and have received a lot of attention due to their wide applicability. An important challenge in the theory is to extend the characterization provided by the Folk Theorem to this larger class of games, namely, to characterize the equilibrium outcomes of the game, given the payoffs and the signalling structure.

This problem is difficult and unsolved in its full generality. The literature on the topic is divided mainly in two branches: undiscounted and discounted games. The study on undiscounted repeated games with imperfect monitoring has been initiated by Lehrer (1989, 1990, 1992a, 1992b, 1992c) who studied two-player games with general signals and n-player games with semi-standard information (players' action sets are divided in equivalence classes which are publicly observed) and obtained characterizations there. The work of Lehrer had many followers (among others Hillas and Liu, 1996, Tomala, 1998, 1999, Renault and Tomala, 2004, Renault et al. 2005)[2]. The literature on discounted game generaly studies a more stringent equilibrium concept than the Nash equilibrium, namely subgame perfect or sequential equilibria. The extension of the Folk Theorem to subgame perfect equilibria for games with perfect observation is due to Fudenberg and Maskin (1986). This work was followed by a detailed study of discounted repeated games with *public signals and public strategies*. These are games where all players observe the same signals and condition their play on these signals only, thereby forgetting their own past actions. This restriction of strategies allows, on one hand, to give a concise definition of subgame perfection, and on the other hand to use dynamic programming methods to characterize the equilibrium payoff set (Abreu et al., 1990, Fudenberg and Levine, 1994, Fudenberg et al., 1994, 2007). Sequential equilibria of repeated games without the assumption of public monitoring are much harder to study and no general characterization is known so far, even for a (reasonably large) subclass of games (see e.g. among others, Kandori and Matsushima, 1998, Ely et al. 2005 and the survey book of Mailath and Samuelson, 2006).

The aim of the present paper is to present in a systematic way the results that are valid for any supergame (that is for any payoff and signalling structures). The paper contains old as well as new results, and is structured as follows. Section

[2]Most of these works consider restricted setups, i.e. two-player games (Hillas and Liu, 1996), specific signals or games (Tomala, 1998, 1999, Renault et al. 2005)

2 presents the basic ingredients of supergames. In this section we first describe the model, the notions of feasible and individually rational payoffs and the main notions of repeated games and equilibrium payoffs (finitely repeated, discounted, and uniform). We also review the classical Folk Theorem, and finally we present several properties valid for any signalling structure, including a new result on the convergence of the T-fold repeated equilibrium payoffs set when T goes to infinity. In section 3, we review the result of Renault and Tomala (2004) which is a characterization of *communication* equilibrium[3] payoffs and which holds for any supergame. This is the tightest upper bound on the set of Nash equilibrium payoff known to date and valid in full generality. We also provide some new applications of this result in subsections 3.5, 3.6, 3.7. Specifically, we spot conditions on signals ensuring that, for any payoff function, communication equilibrium payoffs are the feasible and individually rational payoffs. In section 4, we present some extensions, specific results and open problems.

2 Basic ingredients

2.1 Description of the game

In all the paper we consider a repeated game with signals, or supergame, Γ defined by:

• a finite stage game $G = (N, (A^i)_{i \in N}, (g^i)_{i \in N})$ given by a set of players $N = \{1, ..., n\}$, and for each player i a finite non empty set of actions A^i and a payoff function $g^i : A \longrightarrow I\!R$, where $A = \prod_i A^i$ stands for the set of action profiles. The stage game G is also called the one-shot game.

• a signalling structure $((U^i)_{i \in N}, f)$ where for each player i, we have a finite non empty set of signals U^i, and a signalling function $f : A \longrightarrow \Delta(U)$, where $U = \prod_i U^i$ is the set of signal profiles and $\Delta(U)$ is the set of probability distributions over U.

The game is played as follows: at each stage $t = 1, 2, \ldots$, each player chooses an action in his own set of actions, choices are simultaneous. If $a_t = (a_t^i)_{i \in N} \in A$ is the action profile chosen, a profile of signals $u_t = (u_t^i)_i$ is selected according to the distribution $f(a_t)$. Each player i then observes his signal u_t^i and the game proceeds to stage $t + 1$. The payoff for player i at stage t is then $g^i(a_t)$, but is not necessarily observed by player i, all what player i learns before starting stage $t + 1$ is u_t^i.

[3]A communication equilibrium is a Nash equilibrium of an extended repeated game where players communicate costlessly with a trustworthy mediator between game stages. It allows for almost all kinds of costless communication, for example the players may send and receive private messages between the stages, send emails using the cc: and bcc: functions, etc... However, this communication is disconnected from the fundamentals of the repeated game: player i may send a message to player j saying he has just played action a^i, but player i can not prove, or certify, that he actually played a^i.

In this paper we aim at presenting results without any assumption on the stage game G nor on the observation structure, nevertheless we discuss illustrative examples.

Example 2.1. *An example of a stage game is the famous following "prisoner's dilemma": there are 2 players, $A^1 = \{C^1, D^1\}$ and $A^2 = \{C^2, D^2\}$, and the payoffs are given by the following matrix, with the first, resp. second, coordinate being the payoff to player 1, resp. player 2:*

$$\begin{array}{cc} & \begin{array}{cc} C^2 & D^2 \end{array} \\ \begin{array}{c} C^1 \\ D^1 \end{array} & \begin{pmatrix} (3,3) & (0,4) \\ (4,0) & (1,1) \end{pmatrix} \end{array}$$

The following particular cases of signalling structures are often met in the literature. The most standard case is the one of *perfect observation of actions* (also called *perfect monitoring*) when $U^i = A$, and $u_t^i = a_t$ for each player i. Player i is said to have *trivial observation* if U^i is a singleton (such a player gets no information on the other players' actions). *Payoffs are observable* by player i if this player can compute his current payoff based on his signal and his own action (that is $g^i(a) = \tilde{g}^i(a^i, u^i)$). Signals are *public* when $U^i = U^j$ and $u_t^i = u_t^j$ for all players i and j and stage t.

Example 2.2. *Consider the following 3-player minority game: 3 players have to vote for one of two alternatives A and B, the player (if any) who vote for the less chosen alternative receives a reward of one euro and other players receive zero. The current majority alternative is publicly announced after each stage. This defines a repeated game with public signals and observable payoffs.*

We turn now to the definition of strategies in the repeated game.

Definition 2.3. *A strategy for player i is an element $\sigma^i = (\sigma_t^i)_{t\geq 1}$, where $\sigma_t^i : (A^i \times U^i)^{t-1} \longrightarrow \Delta(A^i)$ gives the lottery played by player i at stage t depending on his current information.*

In words, a strategy defines the distribution of the next action of player i, given his current information, i.e. his history of own actions and signals. Such a strategy is usually called a *behavior* strategy: a player performs a local lottery on his actions at every stage depending on his own past history. This shall be the one and only concept of strategy used in this paper. A *pure* strategy is a behavior strategy that chooses actions in a deterministic fashion and is thus a particular behavior strategy[4]. In the whole paper we use the term strategy to refer to behavior strategies.

[4]Perfect recall of the whole past is a feature of the supergame, thus according to Kuhn's theorem (Kuhn, 1953), mixed strategies (lotteries over the set of pure strategies, endowed with the product sigma-algebra) are equivalent to behavior strategies here.

The set of strategies for player i is denoted by Σ^i, and we write $\Sigma = \prod_{i\in N} \Sigma^i$. A *play* is a sequence of profiles of actions and signals, $\omega = (a_1, u_1, a_2, u_2, \ldots)$ in $(A\times U)^\infty$, with a_t being the action profile played at stage t and u_t being the signal profile observed at stage t. A strategy profile in Σ naturally induces a probability distribution over the set of plays $\Omega = (A \times U)^\infty$ endowed with the product σ-algebra. This probability distribution will be used to compute expected payoffs.

Notations 2.4. *Throughout the paper, we use the following notations. The vector payoff function is denoted by g, that is $g(a) = (g^i(a))_{i\in N}$ for each action profile a. M denotes an upper bound for all absolute values of payoffs: $M = \max_{i,a} |g^i(a)|$. In general if $(E^i)_{i\in N}$ is a collection of sets indexed on players, E stands for $\prod_{i\in N} E^i$. We denote by e^{-i} the current element of $E^{-i} = \prod_{j\neq i} E^j$ and write $e = (e^i, e^{-i}) \in E$ when the i-th component is stressed. If E is a finite set, we let $|E|$ be its cardinality and $\Delta(E)$ be the set of probability distributions over E. An element e in E is identified with the Dirac mass on e. For $p = (p(e))_{e\in E}$ in $\Delta(E)$,* $\operatorname{supp} p$ *denotes the support of p. Given $h : E \to I\!R$, we extend h to $\Delta(E)$ in the usual fashion: $h(p) = \sum_{e\in E} p(e)h(e)$ for $p \in \Delta(E)$. Regarding payoff vectors, we use the Euclidean norm on $I\!R^N$ and the Hausdorff distance between compact subsets of $I\!R^N$: $d(A, B) = \max\{\max_{a\in A} \min_{b\in B} \|a - b\|, \max_{b\in B} \min_{a\in A} \|a - b\|\}$.*

2.2 Feasible and individually rational payoffs

We define now feasible and individually rational payoffs, which only depend on the stage game G.

Informally, a payoff vector is feasible if it is induced by some strategy profile in (some version of) the repeated game. This is captured by the set of payoff vectors achievable with correlated strategies in the stage game, that is $g(\Delta(A)) = \{g(P), P \in \Delta(A)\} = \operatorname{conv} g(A)$. This set is the convex hull of the payoff vectors achievable with pure strategies in the one-shot game, hence a polytope. This set includes all equilibrium payoffs considered in this paper.

Definition 2.5. *The set of feasible payoffs is $F = g(\Delta(A))$.*

We introduce now individual rationality levels (or punishment levels, or minmax levels) which measure the payoff that each player can secure to himself, irrespective of the behavior of other players. Two levels are considered.

Definition 2.6. *For each player i in N, the independent minmax level of player i is:*

$$v^i = \min_{x^{-i}\in\prod_{j\neq i}\Delta(A^j)} \max_{x^i\in\Delta(A^i)} g^i(x^i, x^{-i}).$$

The correlated minmax level of player i is:

$$w^i = \min_{x^{-i}\in\Delta(\prod_{j\neq i} A^j)} \max_{x^i\in\Delta(A^i)} g^i(x^i, x^{-i}).$$

The first one is what player i gets when all other players minimize his payoff using mixed strategies. The second one is obtained when the other players can correlate their actions. It always holds that $w^i \leq v^i$, and there is an equality for 2-player games. We may have $w^i < v^i$ if there are at least 3 players as shown by the following example.

Example 2.7. *In the following 3-player game where $A^1 = \{T, B\}$, $A^2 = \{L, R\}$ and $A^3 = \{W, E\}$, we have $v^1 = v^2 = v^3 = w^1 = w^2 = 0$ and $w^3 = -1/2$.*

$$\begin{array}{cc} & \begin{array}{cc} L & R \end{array} \\ \begin{array}{c} T \\ B \end{array} & \begin{pmatrix} (0,0,0) & (1,-1,1) \\ (-1,0,1) & (0,1,-1) \end{pmatrix} \\ & W \end{array} \qquad \begin{array}{c} \begin{array}{cc} L & R \end{array} \\ \begin{pmatrix} (1,0,-1) & (1,-1,1) \\ (1,0,0) & (2,-1,0) \end{pmatrix} \\ E \end{array}$$

Payoff vectors above the minmax levels are called individually rational.

Definition 2.8. *The set of individually rational payoffs with respect to the independent minmax levels is defined as: $IR = \{u = (u^i)_{i \in N}, u^i \geq v^i \ \forall i \in N\}$. The set of individually rational payoffs with respect to the correlated minmax levels is defined as: $IRC = \{u = (u^i)_{i \in N}, u^i \geq w^i \ \forall i \in N\}$.*

2.3 Nash equilibrium payoffs

We define now the most common average payoff notions and the corresponding equilibrium concepts.

2.3.1 Finitely repeated and discounted games: the sets E_T and E_λ

Definition 2.9. *Given a positive integer T, the T-stage average payoff for player i induced by the strategy profile σ is:*

$$\gamma_T^i(\sigma) = I\!E_\sigma \left(\frac{1}{T} \sum_{t=1}^{T} g^i(a_t) \right).$$

We denote by G_T the T-stage repeated game, that is the game with strategy sets Σ^i and payoff functions γ_T^i for each player i in N, and by E_T the set of Nash equilibrium payoffs of G_T.

Since only the first T stages matter in G_T, it can be seen as a finite extensive-form game.

Definition 2.10. *Given a discount factor λ in $(0,1]$, the λ-discounted payoff for player i induced by the strategy profile σ is:*

$$\gamma_\lambda^i(\sigma) = I\!E_\sigma \left(\sum_{t=1}^{\infty} \lambda (1-\lambda)^{t-1} g^i(a_t) \right) = \sum_{t=1}^{\infty} \lambda (1-\lambda)^{t-1} I\!E_\sigma \left(g^i(a_t) \right).$$

We denote by G_λ the λ-discounted repeated game, that is the game with strategy sets Σ^i and payoff functions γ^i_λ for each player i in N, and we denote by E_λ the set of Nash equilibrium payoffs of G_T.

In the λ-discounted game, receiving a payoff of $1-\lambda$ today is equivalent to receiving a payoff of 1 tomorrow (the discounted game is often parametrized by $\delta = 1-\lambda \in [0,1)$). Notice that considering $T=1$ in Definition 2.9 or $\lambda = 1$ in definition 2.10 leads to the stage game $G = G_1$. This game being finite, it has a mixed Nash equilibrium by the Nash theorem. Repeating such an equilibrium at each stage, independently of the past, constitutes a Nash equilibrium of any finitely repeated game G_T and of any discounted game G_λ. It follows that E_T and E_λ contain E_1 and are non empty. By compactness of the strategy sets (endowed with the product topology) and continuity of the payoff functions, E_T and E_λ are compact subsets of $I\!R^N$. We have thus the following:

Lemma 2.11. *For any T and λ, E_T and E_λ are compact subsets of $I\!R^N$, both including E_1.*

The T-stage repeated game has a finite duration. For the λ-discounted game, one can interpret λ as the probability that the game terminates at the current period, so that $1/\lambda$ represents the expected number of repetitions. With this interpretation the expected duration of G_λ is finite. In both cases, only finitely many game stages have a significant impact on the average payoff. One way to study long-run repeated games is to go for an asymptotic approach and to consider $\lim_{T\to\infty} E_T$ or $\lim_{\lambda\to 0} E_\lambda$.

2.3.2 The set E_∞ of uniform equilibrium payoffs, and the set E_*

The most common equilibrium notion for undiscounted infinitely repeated games is the *uniform equilibrium*[5] (see Sorin 1986, Fudenberg and Levine 1991), which on the one hand allows for some arbitrary small error ε and on the other hand requires a strong property of uniformity in time of the equilibrium strategies.

Definition 2.12. *A strategy profile σ is a uniform Nash equilibrium of Γ if: for each $\varepsilon > 0$, σ is an ε-Nash equilibrium of any finitely repeated game with sufficiently many stages, and $(\gamma^i_T(\sigma))_{i\in N}$ converges as T goes to infinity to a limit called a* uniform Nash equilibrium payoff. *We denote by E_∞ the set of uniform Nash equilibrium payoffs.*

The first condition formally writes:

$$\forall \varepsilon > 0, \exists T_0 > 0, \forall T \geq T_0, \forall i \in N, \forall \tau^i \in \Sigma^i, \gamma^i_T(\tau^i, \sigma^{-i}) \leq \gamma^i_T(\sigma) + \varepsilon.$$

[5] Another approach is to introduce the limit average payoff $\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^T g^i(a_t)$. More precisely, since this limit may not exist for all sequences of actions, one may extend the limit in some way (taking the limsup for instance) and study the game associated with the chosen extension. We do not follow this approach here, except for a few comments in section 4.

Notice that T_0 depends only on ε and is uniform with respect to the strategies τ^i, that is the ε-Nash condition should be *uniform* in T. The following reformulation is useful.

Lemma 2.13. *(Mertens Sorin Zamir, 1994) Let $x = (x^i)_{i \in N} \in I\!R^N$. The following conditions are equivalent:*

(i) x is a uniform equilibrium payoff,

(ii) $\forall \varepsilon > 0$, $\exists \sigma = (\sigma^i)_{i \in N} \in \Sigma$, $\exists T_0$, $\forall T \geq T_0$, $\forall i \in N$, $\forall \tau^i \in \Sigma^i$, $\gamma^i_T(\tau^i, \sigma^{-i}) \leq x^i + \varepsilon$ and $\gamma^i_T(\sigma) \geq x^i - \varepsilon$,

(iii) there exists a sequence $(\varepsilon_k, T_k, \sigma_k)_{k \geq 1}$ such that $(\varepsilon_k)_k$ is a real sequence decreasing to 0, for each k, T_k is a positive integer and σ_k is an ε_k-Nash equilibrium of G_{T_k} with payoff ε_k-close to x.

We briefly recall the proof. The main argument is to construct a strategy in the repeated game by concatenating ε-equilibria of finitely repeated games. The implications $(i) \implies (ii) \implies (iii)$ are clear and it is enough to show that $(iii) \implies (i)$. Choose for each k an integer l_k large enough so that $l_k T_k \varepsilon_k \geq T_{k+1}$. A strategy profile σ is defined by playing σ_1 cyclically l_1 times, then σ_2 cyclically l_2 times, ... , σ_k cyclically l_k times and so on (past history is reset to the empty history whenever a new strategy starts). Simple computations show that σ is a uniform equilibrium with payoff x.

Corollary 2.14. *For each T, $E_T \subseteq E_\infty$. The set E_∞ is compact and convex.*

The first properties follow immediately from Lemma 2.13, and convexity is obtained by alternating between two uniform equilibria with equal time proportions.

Regarding discounted payoffs, the inclusion $E_\lambda \subseteq E_\infty$ seems likely to hold, but to the best of our knowledge this is still an open problem. We have the following lemma.

Lemma 2.15. *Let σ be a uniform equilibrium. Then for each $\varepsilon > 0$, σ is an ε-Nash equilibrium of any discounted game with low enough discount factor:*

$$\forall \varepsilon > 0, \exists \lambda_0 \in (0,1], \forall \lambda \in (0, \lambda_0], \forall i \in N, \forall \tau^i \in \Sigma^i, \gamma^i_\lambda(\tau^i, \sigma^{-i}) \leq \gamma^i_\lambda(\sigma) + \varepsilon,$$

and $\gamma_\lambda(\sigma)$ converges as λ goes to 0 to $\lim_{T \to \infty} \gamma_T(\sigma)$.

Proof: First consider an arbitrary bounded sequence of real numbers $(x_t)_{t \geq 1}$, and denote by $\bar{x}_T$ the Cesaro mean $\frac{1}{T}\sum_{t=1}^T x_t$ and by $\bar{x}_\lambda$ the Abel mean $\sum_{t=1}^\infty \lambda (1-\lambda)^{t-1} x_t$.

$$\begin{aligned}
\bar{x}_\lambda &= \sum_{t=1}^\infty \left(\sum_{T=t}^\infty \lambda (1-\lambda)^{T-1} - \lambda (1-\lambda)^T \right) x_t, \\
&= \sum_{T=1}^\infty (\lambda (1-\lambda)^{T-1} - \lambda (1-\lambda)^T) \sum_{t=1}^T x_t, \\
&= \sum_{T=1}^\infty T \, \lambda^2 \, (1-\lambda)^{T-1} \, \bar{x}_T.
\end{aligned}$$

As a consequence,

$$\limsup_{T\to\infty} \bar{x}_T \geq \limsup_{\lambda\to 0} \bar{x}_\lambda \geq \liminf_{\lambda\to 0} \bar{x}_\lambda \geq \liminf_{T\to\infty} \bar{x}_T. \qquad (1)$$

and the convergence of $\bar{x}_T$ easily implies the convergence of $\bar{x}_\lambda$ to the same limit.

Let now σ be a uniform equilibrium, and write $\gamma^i(\sigma) = \lim_{T\to\infty}\gamma^i_T(\sigma) = \lim_{\lambda\to 0}\gamma^i_\lambda(\sigma)$. Consider $\varepsilon > 0$ and T_0 such that: $\forall T \geq T_0, \forall i \in N, \forall \tau^i \in \Sigma^i, \gamma^i_T(\tau^i, \sigma^{-i}) \leq \gamma^i(\sigma)+\varepsilon$. Define λ_0 such that: $\forall \lambda < \lambda_0$, $\sum_{T=1}^{T_0-1} T\,\lambda^2\,(1-\lambda)^{T-1} \leq \varepsilon$. Consider now $\lambda < \lambda_0$, a player i in N and a strategy τ^i in Σ^i. We have:

$$\begin{aligned}
\gamma^i_\lambda(\tau^i, \sigma^{-i}) &= \sum_{t=1}^{\infty} \lambda(1-\lambda)^{t-1} \mathbb{E}_{\tau^i,\sigma^{-i}}\left(g^i(a_t)\right), \\
&= \sum_{T=1}^{\infty} T\,\lambda^2\,(1-\lambda)^{T-1}\,\gamma^i_T(\tau^i, \sigma^{-i}), \\
&\leq \varepsilon\, M + \sum_{T=T_0}^{\infty} T\,\lambda^2\,(1-\lambda)^{T-1}\,\left(\gamma^i(\sigma)+\varepsilon\right), \\
&\leq \gamma^i(\sigma) + \varepsilon + \varepsilon(2M+\varepsilon),
\end{aligned}$$

and the lemma is proved. □

Remark: All inequalities in (1) might be strict, but Hardy and Littlewood have proved that the convergence of $\bar{x}_\lambda$ also implies the convergence of $\bar{x}_T$ to the same limit (see e.g. Lippman 1969). This implies that for any strategy profile σ, $\lim_{T\to\infty}\gamma_T(\sigma)$ exists iff $\lim_{\lambda\to 0}\gamma_\lambda(\sigma)$ exists, and in case of convergence both limits are equal. Nevertheless, one can construct an example of a strategy profile σ satisfying the conclusion of Lemma 2.15 without being a uniform equilibrium.

Uniform equilibria are approximations of Nash equilibria of finite but long games. The concept is adapted to repeated games with long and uncertain duration. The strategies should be approximately optimal independently of the number of stages, provided it is large enough (or on the discount factor, provided it is low enough, recall Lemma 2.15). By contrast, the equilibria in G_T or G_λ need not have such robustness property as they may crucially depend on the exact value of T and λ. Another important difference when studying E_T or E_λ is that players' strategies are exact best replies to the strategies of the other players. With uniform equilibria, the players only play ε-best replies, where ε vanishes as the number of stages goes to infinity. Most of the present paper focuses on uniform equilibria which are easier to deal with and for which the sharpest characterizations are known. Mathematically, the clearest results are obtained with the set E_∞ whereas results for $\lim_{\lambda\to 0}E_\lambda$ and $\lim_{T\to\infty}E_T$ are more technical.

We finally define another set of equilibrium payoffs which contains all the previously defined sets, and which is fully characterized when allowing for a mediator (see Theorem 3.16).

Definition 2.16. *Let E_* be the set of vectors x in $\mathbb{R}^N$ satisfying: $\forall \varepsilon > 0$, there exists σ in Σ and λ in $(0,1]$ such that σ is an ε-Nash equilibrium of the discounted game G_λ with payoff ε-close to x.*

The set E_* clearly contains E_λ for each discount factor λ. By lemma 2.15, it also contains E_∞[6] and consequently E_T for each T. To the best of our knowledge, there is no example of game where the equality $E_\infty = E_*$ does not hold.

2.4 Benchmark: the standard case of perfect observation

In this section, we assume perfect observation of the action profile and present the main *Folk theorems* for Nash equilibrium payoffs. These results essentially state that *in repeated games with perfect observation and very patient players, the set of equilibrium payoffs is the set of feasible and individually rational payoffs.*

A first observation is that E_*, E_∞, E_T and E_λ are subsets of $F \cap IR$. The inclusion in F is clear. When actions are publicly observed at each stage, given a strategy profile σ^{-i} for players other than i, player i may choose a best-reply to the mixed action profile of the others after each history. This secures a payoff no less than v^i to player i at each stage and thus there exists σ^i such that for each T, $\mathbb{E}_{\sigma^i,\sigma^{-i}}(g^i(a_t)) \geq v^i$. It follows that E_*, E_∞, E_T and E_λ are subsets of $F \cap IR$. The first Folk theorem is the following.

Theorem 2.1. "The" Folk theorem: *In case of perfect observation, the uniform equilibrium payoffs are the feasible and individually rational payoffs: $E_\infty = E_* = F \cap IR$.*

It is difficult to establish who proved this result. As Aumann wrote, it *"has been generally known in the profession for at least 15 or 20 years, but has not been published; its authorship is obscure." (R.J. Aumann, 1981).* This is an "everything is possible" result: any *reasonable* payoff can be achieved at equilibrium. We recall the basic and important proof.

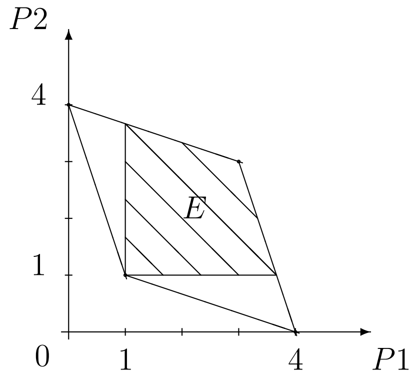
Proof: Fix $u \in F \cap IR$. There exists a play $h = (a_1, \ldots, a_t, \ldots)$ such that for each player i, $\frac{1}{T}\sum_{t=1}^{T} g^i(a_t) \to_{T\to\infty} u^i$. The play h is called the main path of the strategy, and playing according to h for some player i at stage t means playing the i-component of a_t. For each pair of distinct players (i,j), fix $x^{i,j}$ in $\Delta(A^j)$ such that $(x^{i,j})_{j\neq i}$ achieves the minimum in the definition of v^i.

We define now the strategy σ^i of each player $i \in N$. At stage 1, σ^i plays according to the main path, and continues to play according to h as long as the other players do so. If there is a first stage $t \geq 1$ at which some player j does not follow the main path, then σ^i plays the mixed action $x^{j,i}$ at all subsequent

[6] Theorem 3.1 in Fudenberg and Levine 1991, also implies $E_\infty \subseteq E_*$

stages (if several players leave the main path at the same stage, by convention the punished player is the smallest according to a fixed linear order on N). It is easy to see that $\sigma = (\sigma^i)_{i \in N}$ is a uniform equilibrium of the repeated game with payoff u. $\square$

For the Prisoner's Dilemma (Example 2.1) with perfect observation, we have $v^1 = v^2 = 1$, and obtain the following set of equilibrium payoffs:



We now state the discounted and finitely repeated Folk theorems without proof, the reader is referred to Sorin (1992).

Theorem 2.2. Discounted Folk theorem *(Sorin 1986, Fudenberg Maskin 1986) Assume they are 2 players, or that there exists $u = (u^i)_{i \in N}$ in E such that for each player i, $u^i > v^i$. Then, in case of perfect observation, $E_\lambda \xrightarrow[\lambda \to 0]{} F \cap IR$.*

Compared with Theorem 2.1, a condition on the payoff of the stage-game is required. This is mainly due to the fact that a profitable one-shot deviation may be profitable in the discounted game, while in the infinitely repeated undiscounted game, a one-shot gain is offset at the limit. There is thus a need to be able to punish even "small" deviations. The following example (due to Forges, Mertens and Neyman, 1986) is a counter-example to the convergence of E_λ to $F \cap IR$, i.e. to the conclusion of Theorem 2.2, when the assumption that there is a feasible payoff which is strictly individually rational is not met. Consider the payoff matrix,

$$\begin{pmatrix} (1,0,0) & (0,1,0) \\ (0,1,0) & (1,0,1) \end{pmatrix}.$$

This is a 3-player game where player 1 chooses a row, player 2 chooses a column and player 3 has no action. Essentially, this game is a zero-sum game between players 1 and 2, and in each Nash equilibrium of G_λ each of these players independently randomizes at each stage between his 2 actions with equal probabilities. Therefore $E_\lambda = \{(1/2, 1/2, 1/4)\}$ for each λ, whereas $(1/2, 1/2, 1/2) \in F \cap IR$.

Theorem 2.3. Finitely repeated Folk theorem *(Benoît and Krishna, 1987)* *Assume that for each player i there exists x in E_1 such that $x^i > v^i$.*
Then in case of perfect observation, $E_T \xrightarrow[T\to\infty]{} F \cap IR$.

As for Theorem 2.2, there is a condition on the stage-game[7]. It is now required that there is a strictly individually rational Nash payoff. This is mainly a "boundary" effect, a one-shot Nash equilibrium has to be played at the last stage of the game. Regarding the importance of the condition, there is a strong result holding whenever each player receives no more than his independent minmax payoff in any Nash equilibrium of the stage game.

Proposition 2.17. *(Sorin, 1986) In case of perfect observation, if $E_1 = \{v\}$ then $E_T = \{v\}$ for each T .*

The proof is simple and by induction on T. The result applies to the Prisoner's Dilemma of Example 2.1: for each T, $E_T = \{(1,1)\}$, hence the Pareto-optimal equilibrium payoff $(3,3)$ can not be approximated by equilibrium payoffs of finitely repeated games, and there is no convergence of E_T to $F \cap IR$.

2.5 Beyond the perfect observation case, a few general properties

How can we generalize the Folk Theorem to games with imperfect observation? More precisely, a main open problem in repeated games is to compute the set of Nash equilibrium payoffs E_∞ for general stage games and observation structures. It is clear that signals do matter, and that the proof of Theorem 2.1 is not valid for general observation structures. We see five main obstructions for extending the proof, which can be viewed as five main features of repeated games with imperfect observation.

1) A deviation of a player from the main path may not be detected by the other players. Consider for instance the Prisoner's Dilemma with trivial observation for both players. Clearly, no cooperation is possible there, and $E_T = E_\lambda = E_\infty = \{(1,1)\}$ for all T and λ.

2) It follows from the above point that players should not be offered profitable and undetectable deviations. The notion of detectable deviation is not straightforward though. It may be the case that a player is incentivized to follow his equilibrium strategy, neither because his actions maximize his current payoffs, nor because his actions induce correct current signals for the other players, but because the strategy gives player i a superior information on the actions of the players $-i$. This was first noticed by Lehrer (1989,1992), see Example 3.6 in Section 3.

3) A deviation from the main path may be detected by some players, but not by others. This happens for instance when the players are vertices of a graph and

[7] see Benoit Krishna 1985 and Gossner 1995 for a Folk theorem for subgame-perfect equilibria

only observe after each stage the actions played by their neighbours (Ben-Porath and Kahnemann 1996, Renault and Tomala 1998).

4) A deviation may be detected by all players, yet the identity of the deviator remains unknown (Tomala 1999, Renault et al. 2005 and 2008).

5) The independent minmax v^i is the punishment level of player i in case of perfect observation. But the signals may create the possibility to correlate actions and to punish player i below the level v^i. This leads to new punishment levels depending on the observation structure (see Gossner and Tomala, 2007).

Giving a general formula for E_∞ is an open and difficult problem, even for two players. In the next section, we show that allowing players to communicate between game stages allows to give a full characterization of *communication* equilibrium payoffs, which include all uniform equilibrium payoffs. The intuition is two-fold. First, since players have asymmetric information about past histories, they may want to communicate to coordinate their play. Second, their continuation strategies are correlated, thus inducing a correlated equilibrium of the continuation game. Therefore, using communication and correlation devices smoothes the analysis. We recall that a *correlated equilibrium* (Aumann, 1974) of a game **G** is a Nash equilibrium of any extension of **G** where players first receive private signals and then choose actions and receive payoffs as in **G**. A correlated equilibrium of the stage game G induces a *correlated equilibrium distribution*, that is $p \in \Delta(A)$ such that,

$$\forall i \in N, \forall a^i, b^i \in A^i, \sum_{a^{-i} \in A^{-i}} p(a^i, a^{-i}) g^i(a^i, a^{-i}) \geq \sum_{a^{-i} \in A^{-i}} p(a^i, a^{-i}) g^i(b^i, a^{-i}).$$

A correlated equilibrium payoff of G is a vector $g(p)$, where p is a correlated equilibrium distribution. We conclude this section with a few general results valid for any observation structure.

A. Feasibility and individual Rationality. Equilibrium payoffs have to be feasible and individually rational but as mentioned in point 5) above, one has to take care of the punishment level. Precisely, given a strategy profile σ^{-i} of the players different from i, player i can play at each stage a best reply to the expected distribution of actions of the other players. Hence, there exists σ^i such that for each t, $I\!E_{\sigma^i, \sigma^{-i}}(g^i(a_t)) \geq w^i$. Thus,

Lemma 2.18. *E_*, E_∞, E_T and E_λ are subsets of $F \cap IRC$.*

B. Finitely repeated games.

One may wonder if Proposition 2.17 extends directly to general observation structures, the answer is negative.

Lemma 2.19. *With imperfect monitoring, it may happen that* $E_1 = \{v\}$ *and* $E_2 \neq \{v\}$.

Proof: Such an example can be found in Mailath *et al.* (2002), we present here a simple variant with deterministic public signals. The stage game is similar to an example of Moulin and Vial (1978).

$$\begin{array}{cc} & \begin{array}{ccc} l & m & r \end{array} \\ \begin{array}{c} T \\ M \\ B \end{array} & \begin{pmatrix} (0,0) & (2,1) & (1,2) \\ (1,2) & (0,0) & (2,1) \\ (2,1) & (1,2) & (0,0) \end{pmatrix} \end{array}$$

In the stage game there is a unique Nash equilibrium where each player plays each action with probability 1/3, and we have $E_1 = \{(1,1)\} = \{v\}$. However, there is a correlated equilibrium distribution where each entry with payoff $(2,1)$ or $(1,2)$ has probability 1/6, yielding a correlated equilibrium payoff of $(3/2, 3/2)$. Define now the set of signals $U^1 = U^2 = \{Good, Bad\}$, and assume that the public signal is *Bad* if the payoff is $(0,0)$ and *Good* if the payoff is $(2,1)$ or $(1,2)$.

Consider the following strategy profile in the 2-stage game. At the first stage, the players play the mixed Nash equilibrium where each action is played with probability 1/3. If the public signal is *Bad*, they play again this mixed Nash equilibrium at the second stage, but if the signal is *Good*, each player repeats at stage 2 the pure action he played at stage 1. This strategy profile is a Nash equilibrium of the 2-stage game and its payoff is $1/2((1,1)+1/3(1,1)+2/3(3/2,3/2)) = (7/6, 7/6)$. □

Sekiguchi (2001) gave specific conditions on the monitoring structure ensuring that if $E_1 = \{v\}$, then $E_T = \{v\}$ for all T. Without conditions on the monitoring structure, the natural extension of Proposition 2.17 is the following result. Let C_T denote the set of communication equilibrium payoffs of the finitely repeated game G_T (see Section 3). For each T, C_T is convex and compact, and $E_T \subseteq C_T \subseteq F \cap IRC$. When $T = 1$, C_1 simply reduces to the set of correlated equilibrium payoffs of G.

Proposition 2.20. *If* $C_1 = \{w\}$, *then* $w = v$ *and* $C_T = E_T = \{v\}$ *for each* T.

More general conditions on the stage game have been given by Sekiguchi (2005), who studies correlated equilibrium payoffs in the case of full support[8] stochastic signals (all signals have positive probability under all action profiles).

Proof[9] of proposition 2.20: The assumption clearly implies $E_1 = \{w\}$, and thus $w = v$. We proceed by induction, and assume $C_T = \{w\}$ for some T. Consider a communication equilibrium σ of G_{T+1}. The restriction of σ to the stages

[8] With this assumption, Nash equilibrium payoffs and sequential equilibrium payoffs of the T-stage games coincide.

[9] The proof that $E_T = \{v\}$ remains correct by replacing "communication equilibrium" by "correlated equilibrium" throughout.

2,..., $T+1$ defines a communication equilibrium of G_T, hence by assumption we have for each player i:

$$\mathbb{E}_\sigma\left(\frac{1}{T}\sum_{t=2}^{T+1} g^i(a_t)\right) = w^i.$$

Suppose by contradiction that the strategy of stage 1 induced by σ is not a correlated equilibrium of G. There exists a player i who can profitably deviate at the first stage, and obtain an expected payoff no less than w^i at all subsequent stages, by playing a best reply to the expected distribution of actions of the other players. Hence, player i has a profitable deviation in the $T+1$-stage game, a contradiction. It follows that σ plays a correlated equilibrium of G at stage 1, and for each player i, $\mathbb{E}_\sigma\left(g^i(a_1)\right) = w^i$. As a consequence, $\mathbb{E}_\sigma\left(\frac{1}{T+1}\sum_{t=1}^{T+1} g^i(a_t)\right) = w^i$ for each i, and $C_{T+1} = \{w\}$. $\square$

An immediate consequence of Proposition 2.20 is the following corollary: no observation structure can yield cooperation in the finitely repeated prisoner's dilemma.

Corollary 2.21. *In the finitely repeated Prisoner's Dilemma, $C_T = E_T = \{(1,1)\}$ for all T and all signalling structures.*

The following proposition shows that the sequence $(E_T)_T$ behaves somehow like an increasing sequence, in that it always converges (for the Hausdorff distance) to the closure of $\bigcup_{T\geq 1} E_T$. In the statement below, TE_T denotes $\{Tx : x \in E_T\}$.

Proposition 2.22.
1) For all T and T', $TE_T + T'E_{T'} \subseteq (T+T')E_{T+T'}$,
2) $\mathrm{cl}\left(\bigcup_{T\geq 1} E_T\right)$ is convex and compact (cl is the closure operator),
3) $E_T \xrightarrow[T\to\infty]{} \mathrm{cl}\left(\bigcup_{T\geq 1} E_T\right)$.

It follows directly from 1) that $E_T \subseteq E_{kT}$ for all integers k, T and thus $E_T \cup E_{T'} \subseteq E_{TT'}$ for all T, T'.

Proof:
1) This is a concatenation property. Given an equilibrium payoff x in E_T and x' in E'_T, one can construct an equilibrium payoff z of $G_{T+T'}$ by defining strategies which first play an equilibrium of G_T with payoff x and then independently play an equilibrium of $G_{T'}$ with payoff x' . We have $z = \frac{1}{T+T'}\left(Tx + T'x'\right)$, hence the result.

2) Define $E := \mathrm{cl}\left(\bigcup_{T\geq 1} E_T\right)$, which is clearly compact. Given x and x' in $\bigcup_{T\geq 1} E_T$, say $x \in E_T$ and $x' \in E_{T'}$, both x and x' are in $E_{TT'}$, thus $(x+x')/2 \in E_{2TT'} \subseteq E$. Hence E is convex.

3) Consider now the space $\mathcal{E}$ of compact subsets of E endowed with the Hausdorff distance. E being compact metric, $\mathcal{E}$ is also a compact metric space, and to obtain the convergence of E_T to E, it is enough to prove that any cluster point

of $(E_T)_T$ is E (see e.g. Aubin, 1977). We thus consider a subsequence $(E_{\varphi(T)})_T$ converging to some limit E' in $\mathcal{E}$, and prove $E' = E$.

It is clear that $E' \subseteq E$, and we prove now the converse. Let x be in $E_{\hat{T}}$ for some $\hat{T}$, and fix $\varepsilon > 0$. There exists T_0 such that for all $T \geq T_0$, $d(E', E_{\varphi(T)}) \leq \varepsilon$. Consider $T \geq T_0$ large enough so that $\varphi(T) = K\hat{T} + r$, with K and r non negative integers such that $K \geq \frac{2M}{\varepsilon}$ and $r < \hat{T}$. Using property 1) with an arbitrary Nash payoff e in E_1, we have that $u := \frac{K\hat{T}x+re}{\varphi(T)} \in E_{\varphi(T)}$, and $\|x - u\| \leq \frac{r}{K\hat{T}}\|u - e\| \leq \varepsilon$. Hence $d(x, E') \leq 2\varepsilon$ for each ε, and $x \in E'$. It follows that $E \subseteq E'$ and the proof is complete. □

C. Discounted repeated games.

There is also a concatenation property for discounted equilibrium payoffs.

Proposition 2.23. *For all λ in $(0, 1]$ and positive integer K,*

$$E_1 \subseteq E_\lambda \subseteq E_{1-(1-\lambda)^{1/K}}.$$

Proof: Fix λ and K, and define $\mu = 1 - (1 - \lambda)^{1/K} \in (0, \lambda]$. Divide the set of stages into blocks $B_1, \ldots, B_K$ where B_k is the set of positive integers b which are equal to k modulo K (i.e. $b = sK + k$ for some integer s). The discounted payoff of a strategy profile σ in the game with discount factor μ writes:

$$\begin{aligned}
\gamma_\mu(\sigma) &= \sum_{k=1}^{K} \sum_{t \in B_k} \mu(1-\mu)^{t-1} I\!E_\sigma\left(g^i(a_t)\right), \\
\gamma_\mu(\sigma) &= \sum_{k=1}^{K} \sum_{s=0}^{\infty} \mu(1-\mu)^{sK+k-1} I\!E_\sigma(g^i(a_{sK+k})), \\
&= \frac{\mu}{\lambda} \sum_{k=1}^{K} (1-\mu)^{k-1} \sum_{s=0}^{\infty} \lambda(1-\lambda)^{s-1} I\!E_\sigma(g^i(a_{sK+k})).
\end{aligned}$$

Now consider an equilibrium σ of Γ_λ, and define the strategy profile τ which plays independent copies of σ on each block $B_1, \ldots, B_K$: at stage $sK + k$, τ plays what σ plays at stage $s + 1$ after the history of the actions played at stages k, $K + k, \ldots, (s - 1)K + k$. By the previous computation, τ is a Nash equilibrium of Γ_μ. Since $\frac{\mu}{\lambda}\sum_{k=1}^{K}(1-\mu)^{k-1} = 1$, the equilibrium payoff is the one induced by σ in the game Γ_λ. □

A corollary of the above proposition is that if E_λ converges as λ goes to 0, the limit must be cl $\left(\bigcup_{\lambda>0} E_\lambda\right)$. However, to the best of our knowledge, the convergence of E_λ still is an open problem.

3 Communication equilibria

We study here a solution concept which includes Nash equilibria and for which a full characterization obtains. A *communication equilibrium* of the repeated game is a Nash equilibrium of an extended repeated game in which players communicate with an exogenous device between game stages. The introduction of communication devices eases the characterization of equilibria as it copes with many problems inherent to games with imperfect observation. Namely, the device collects information from all players and may spread it, it also correlates their strategies, as players condition their actions on the messages received from it. It is straightforward to see that a Nash equilibrium is also a communication equilibrium since the device may be silent. Characterizing communication equilibrium payoffs thus yield a superset of Nash equilibrium payoffs. In this section, we mainly subsume Renault and Tomala (2004) who characterized the set of communication equilibrium payoffs, thereby providing the tightest upper bound of Nash equilibrium payoffs known to date. We also provide some new examples and applications.

3.1 Communication equilibrium payoffs

Communication equilibria have been introduced by Myerson (1986) and Forges (1986). The idea is to add an exogeneous mediator who communicates with the players between the stages. The mediator has no commitment power, no interest in the game (constant payoffs), and may use any communication method or *communication device*. For instance, the mediator may broadcast messages (public communication) or allow the players to exchange emails, allow for cc or bcc, and so on. A communication device defines a new (or extended) repeated game, and a Nash equilibrium of the extended game is called a communication equilibrium of the original game. A special kind of communication equilibrium deserves our attention, these are called *canonical communication equilibria* which bear the two following features.

1) The extended game is such that at each stage, the mediator first sends privately to each player i, a recommended action in A^i. Then the stage game is played and signals are observed. At the end of each stage, each player i sends back a private message in U^i to the mediator.

2) The equilibrium strategies for the players are *faithful*: each player plays recommended actions and sends back the actually observed signals.

It can be shown that canonical communication equilibrium payoffs exhaust all communication equilibrium payoffs. The idea is to start with an arbitrary communication equilibrium and to construct an equivalent canonical communication equilibrium by letting the mediator operate the device and the strategies of the players. This reasoning is known at the *revelation principle* (see Myerson 1986 or Forges 1986) and is mathematically simple, this is just a reformulation. In what follows, we focus on canonical communication equilibria.

We now give formal definitions. To distinguish between recommendations and actions, and between signals and messages, it is actually convenient to define, for each player i: $R^i = A^i$ (R^i is interpreted as the set of recommendations for player i whereas A^i is the set of actions that player i can take), and similarly $M^i{=}U^i$

(M^i for messages sent back to the mediator, U^i for signals observed by player i).

Definition 3.1. *A canonical communication device is an element $c = (c_t)_{t\geq 1}$, where $c_1 \in \Delta(R)$ and for each $t \geq 2$, c_t is a mapping from $(R \times M)^{t-1}$ to $\Delta(R)$.*

A canonical correlation device corresponds to a strategy of the mediator in the extended game. Given a fixed canonical communication device c, we define an infinitely repeated game Γ_c played as follows:

> - stage 1: the mediator selects a joint recommendation $(r_1^i)_{i\in N}$ in R according to c_1, and sends privately the recommendation r_1^i to each player i. Then the players simultaneously choose actions and receive signals as in the original game. To conclude stage 1, each player i chooses a message m_1^i in M^i that he sends privately to the mediator.
>
> - stage t: the mediator selects a joint recommendation $(r_t^i)_{i\in N}$ in R according to $c_t((r_1^i, m_1^i)_{i\in N}, \ldots, (r_{t-1}^i, m_{t-1}^i)_{i\in N})$ and sends privately r_t^i to each player i. Then the players simultaneously choose actions $a_t = (a_t^i)_{i\in N}$ and observe signals (drawn from $f(a_t)$). To conclude stage t, each player i sends back a private message m_t^i to the mediator.

In Γ_c, each player i in N has a special strategy σ^{i*}: at each stage, σ^{i*} plays the recommendation just received, and sends back to the mediator the signal just observed by player i. We will refer to σ^{i*} as the faithful strategy of player i.

Definition 3.2. *If c is a canonical communication device and if the faithful strategy σ^* is a uniform equilibrium of Γ_c, the limit payoff $(\gamma_c^i(\sigma^*))_{i\in N} \in \mathbb{R}^N$ is called a (canonical) communication equilibrium payoff of the original repeated game Γ.*

We denote by C_∞ the set of communication equilibrium payoffs of the repeated game Γ.

C_∞ clearly contains E_∞: given a uniform equilibrium of the repeated game Γ, one can define a canonical communication device mimicking the strategies of the players. We stick to the uniform equilibrium paradigm for two reasons: 1) this makes the analysis easier and 2) this includes all usual equilibrium payoffs. For instance we show (see Theorem 3.16) that C_∞ contains E_T and E_λ for each T and λ, and probably all equilibrium payoffs of any reasonable version of the repeated game with signals.

This set is convex and compact and as in Lemma 2.18 we easily have:

Lemma 3.3. *$C_\infty \subseteq F \cap IRC$.*

In the case of perfect observation, the proof of the Folk Theorem 2.1 easily adapts to communication equilibria. Note that the appropriate punishment levels for communication equilibria are the correlated minmax levels w^i, for $i \in N$. Indeed, in a punishment phase, the mediator may send correlated recommended actions and thus punish to the correlated minmax level.

Theorem 3.1. *In case of perfect observation:* $C_\infty = F \cap IRC$.

In the sequel, we provide a characterization of C_∞ which is valid in all cases. We need the following definitions that pertain to the strategies of players at a given stage of the extended game.

Definition 3.4. *A* decision *of player i is an element of:*

$$D^i = \{d^i = (\alpha^i, \mu^i), \text{ with } \alpha^i : R^i \longrightarrow A^i \text{ and } \mu^i : R^i \times U^i \longrightarrow M^i\}.$$

The special decision $d^{i*} = (\alpha^i, \mu^i)$ such that $\alpha^i(a^i) = a^i$ and $\mu^i(a^i, u^i) = u^i$ for all a^i and u^i, is called the faithful decision *of player i.*

A decision $d^i = (\alpha^i, \mu^i)$ of player i corresponds to a pure strategy of player i in the one-shot extended game: if player i is recommended the action r^i by the mediator, he plays the action $\alpha^i(r^i)$, then if he observes a signal u^i, he reports the message $\mu^i(r^i, u^i)$. The faithful decision plays what is recommended and reports what is observed. The next notations will be important in the sequel.

Notations 3.5. *Assume that the mediator recommends the action profile a in A, that the players $j \neq i$ play faithfully whereas player i plays according to a mixed decision $\delta^i \in \Delta(D^i)$, i.e. chooses a decision in D^i according to δ^i and plays according to it.*

We denote by $g^i_{\delta^i}(a)$ the expected payoff of player i under this scenario, and by $\psi^i(\delta^i, a) \in \Delta(U)$ the induced distribution of the profile of messages received by the mediator.

Notice that $g^i_{\delta^i}(a)$ does not depend on the message reported by player i under δ^i. To select an element u according to $\psi^i(\delta^i, a)$, one may proceed as follows. First draw $d^i = (\alpha^i, \mu^i) \in D^i$ according to δ^i, then choose an element $\tilde{u} = (\tilde{u}^k)_{k\in N}$ in U according to $f(a^{-i}, \alpha^i(a^i))$. Finally set $u = \big((\tilde{u}^k)_{k\neq i}, \mu^i(a^i, \tilde{u}^i)\big) \in U$.

Our characterization is driven by the following ideas. Given recommended actions and reported signals, can the mediator infer that there was a deviation? When so, how should he adapt the future recommendations so as to punish the deviation? This is the object of the next two subsections.

3.2 Undetectable deviations

If a player can deviate from the recommended action without changing the (distribution of) reported signals, he may have an incentive to do so. Inducing the same reported signals means that other players get the same signals, whereas the deviating player gets at least as much information, i.e. is able to infer the signal he would have observed, had he played faithfully. We start with an example.

Example 3.6. *Consider the Prisoner's Dilemma with the following observation structure. Player 1 has trivial observation whereas Player 2 has signal set $U^2 = \{a, b, c\}$. The payoffs and the signals of Player 2 are given by the following matrix:*

$$
\begin{array}{cc}
 & \begin{array}{cc} C^2 & D^2 \end{array} \\
\begin{array}{c} C^1 \\ D^1 \end{array} & \begin{pmatrix} (3,3)\, a & (0,4\,)c \\ (4,0)\, b & (1,1)\, c \end{pmatrix}
\end{array}
$$

We now prove that $(3,3) \in C_\infty$. The strategy c of the mediator (canonical communication device) has a main path and a punishment phase and is defined as follows:

- On the main path at stage t, the mediator recommends Player 2 to play C^2, and Player 1 to play C^1 with probability $1 - 1/\sqrt{t}$ and D^1 with probability $1/\sqrt{t}$.

- The mediator continues as above as long as Player 2's reported signal matches Player 1's recommended action. Otherwise, he goes the punishment phase.

- In the punishment phase the mediator punishes forever, i.e. recommends (D^1, D^2) at every stage.

The limit payoff under the faithful strategies is $(3,3)$. If Player 1 unilaterally deviates from his faithful strategy, he is immediately detected and the punishment phase starts, so no deviation of Player 1 can be profitable.

Assume now that Player 2 deviates from his faithful strategy. Denote by $I\!\!P$ the probability measure induced on plays by the deviation of Player 2, and by $I\!\!E$ the corresponding expectation operator. Consider a large number of stages T and denote by Z the number of stages t in $\{1, \ldots, T\}$ where:

-Player 2 plays D^2 at stage t and,

-the play is still on the main path at $t+1$ (Player 2 manages to report the correct signal).

The probability of reporting a correct signal while playing D^2 at a given stage $t \geq 2$ is at most $1 - \frac{1}{\sqrt{T}}$. Since the recommendations of the mediator are independent across stages, we have for each integer z:

$$
I\!\!P(Z > z) \leq \left(1 - \frac{1}{\sqrt{T}}\right)^z \leq e^{\frac{-z}{\sqrt{T}}}.
$$

This yields an upper bound for the average payoff of Player 2:

$$
\begin{aligned}
I\!\!E\left(\frac{1}{T}\sum_{t=1}^{T} g^i(a_t)\right) &\leq I\!\!P(Z \leq z) I\!\!E\left(\frac{1}{T}\sum_{t=1}^{T} g^i(a_t) \mid Z \leq z\right) + 4\, e^{\frac{-z}{\sqrt{T}}} \\
&\leq \frac{1}{T}(3(T-z) + 4z) + 4\, e^{\frac{-z}{\sqrt{T}}} = 3 + \frac{z}{T} + 4\, e^{\frac{-z}{\sqrt{T}}}.
\end{aligned}
$$

Choosing $z = T^{3/4}$ yields: $\forall \varepsilon > 0$, $\exists T_0$, $\forall T \geq T_0$, $I\!\!E\left(\frac{1}{T}\sum_{t=1}^{T} g^i(a_t)\right) \leq 3 + \varepsilon$. This proves that σ^* is a uniform equilibrium of Γ_c, and thus $(3,3) \in C_\infty$.

Note that in this example, the deviation of Player 2 does not change the (trivial) signal of Player 1. However, to be able to report the correct signal, Player 2 must play C^2. The incentive for Player 2 to play C^2 is thus to get information about the action of Player 1. Notice that the mediator is needed to get cooperation here: we can show that $E_\infty = \{(1,1)\}$. □

We explain now the role of undetectable deviations. Suppose that at some stage, the mediator recommends the action profile $a = (a^i)_i$. Assume that some player i has an action b^i which does not change the signals of the other players and gives player i as least as much information as a^i. Then at equilibrium, b^i should not give a better payoff to player i than a^i. Otherwise player i would profitably deviate to b^i without changing the continuation play. We now formalize this idea due to Lehrer (1992a).

Definition 3.7. *A correlated distribution of actions $p \in \Delta(A)$ is* robust to undetectable deviations *if $\forall i \in N$, $\forall \delta^i \in \Delta(D^i)$,*

$$(\forall a \in A,\ \psi^i(\delta^i, a) = f(a)) \implies g^i(p) \geq g^i_{\delta^i}(p)$$

We denote $\mathcal{P}$ the set of those distributions. The set of feasible payoffs robust to undetectable deviations is $g(\mathcal{P})$.

If $\psi^i(\delta^i, \cdot) = f(\cdot)$, then player i may deviate (from the faithful decision) and play δ^i while inducing the same distribution of reported signals. Such a deviation is *undetectable by the mediator*. $g(\mathcal{P})$ is thus the set of payoffs feasible by a distribution for which there exists no profitable and undetectable deviaton. Clearly, $\mathcal{P}$ contains the set of correlated equilibrium distributions of the stage game and is thus non-empty. Also, $\mathcal{P}$ is a polytope. The equation $\psi^i(\delta^i, \cdot) = f(\cdot)$ is linear w.r.t. δ^i, $g^i(p)$ is linear w.r.t. p and $g^i_{\delta^i}(p)$ is separately linear w.r.t. δ^i and p. It follows that $\mathcal{P}$ can be represented as the "dual" of a polytope and is thus a polytope.

Remark 3.8. *The expression of $\mathcal{P}$ is simpler for deterministic signals. Assume that each player i has a deterministic signalling function $f^i : A \longrightarrow U^i$, so that player i observes $f^i(a)$ when a is played. The following notions are introduced by Lehrer (1992a). Given two actions a^i and b^i of player i, we say that b^i is "greater" than a^i, and write $b^i \geq a^i$ if:*
(i) $\forall a^{-i} \in A^{-i}$, $\forall j \neq i$, $f^j(b^i, a^{-i}) = f^j(a^i, a^{-i})$ and
(ii) $\forall a^{-i}, b^{-i} \in A^{-i}$, $f^i(a^i, a^{-i}) \neq f^i(a^i, b^{-i}) \Longrightarrow f^i(b^i, a^{-i}) \neq f^i(b^i, b^{-i})$.
In this case,

$$\mathcal{P} = \Big\{ p \in \Delta(A), \forall i \in N, \forall b^i, a^i \in A^i \text{ s.t. } b^i \geq a^i,$$
$$\sum_{a^{-i} \in A^{-i}} p(a^i, a^{-i}) g^i(a^i, a^{-i}) \geq \sum_{a^{-i} \in A^{-i}} p(a^i, a^{-i}) g^i(b^i, a^{-i}) \Big\}.$$

The condition $b^i \geq a^i$ means that both actions induce the same signals for all other players (condition (i), b^i is *equivalent* to a^i) and that player i playing b^i can compute the signal he would have observed had he play a^i (condition (ii), b^i is *more informative* than a^i). It follows directly form this remark that $\mathcal{P}$ is the set of correlated equilibrium distributions when signals are trivial for each player.

This notion together with individual rationality is enough to characterize communication equilibrium payoffs in the 2-player case. Lehrer (1992a) considered 2-player games with deterministic *non-trivial observation* (each player can learn something from his signals[10]) and deterministic signals and showed that the set of *uniform correlated equilibrium* payoffs is $g(\mathcal{P}) \cap IR$ (recall that $IR = IRC$ for 2-player games). Mertens et al. (1994) showed that, under the same assumptions, $C_\infty = g(\mathcal{P}) \cap IRC$. Hillas and Liu (1996) extended Lehrer's result to non-trivial stochastic signalling. The following characterization of C_∞ for any 2-player supergame extends those results and is in Renault and Tomala (2004).

Theorem 3.9. *For 2-player games, the communication equilibrium payoffs are the feasible payoffs that are robust to undetectable deviations and individually rational:*

$$C_\infty = g(\mathcal{P}) \cap IRC.$$

3.3 Jointly rational payoffs

With three or more players, it may be the case that a player deviates from the main path, all players (and the mediator) know that a deviation occurred but players (and the mediator) do not know who has deviated. As a consequence, several players have to be simultaneously punished. Since such collective punishments may not be possible, new constraints on the equilibrium payoffs appear. To get an intuition on these constraints, we start by an example (variant of Example 3.1 in Renault and Tomala 2004).

Example 3.10.

$$\begin{array}{c} \\ T \\ B \end{array} \begin{array}{c} \begin{array}{cc} L & R \end{array} \\ \begin{pmatrix} (0,0,0) & (0,3,0) \\ (3,0,0) & (1,1,0) \end{pmatrix} \\ W \end{array} \begin{array}{c} \begin{array}{cc} L & R \end{array} \\ \begin{pmatrix} (0,2,0) & (0,2,0) \\ (0,2,0) & (0,2,0) \end{pmatrix} \\ M \end{array} \begin{array}{c} \begin{array}{cc} L & R \end{array} \\ \begin{pmatrix} (2,0,0) & (2,0,0) \\ (2,0,0) & (2,0,0) \end{pmatrix} \\ E \end{array}$$

This is a 3-player game, Player 1 *chooses the row, Player* 2 *chooses the column and Player* 3 *chooses the matrix. Player* 3 *has trivial observation whereas Players* 1 *and* 2 *perfectly observe the moves of each other.*

Note that Player 3 has constant payoffs and that the minmax levels are $w^1 = w^2 = w^3 = 0$, thus all feasible payoffs are individually rational. In the case of perfect observation, we have $C_\infty = g(\Delta(A)) = \{(x^1, x^2, 0) \in \mathbb{R}^3_+, x^1 + x^2 \leq 3\}$. In the case of trivial observation for each player, one can check that (T, L, W), (T, R, W) and (B, L, W) are played with probability 0 in any correlated equilibrium of the one-shot game. Thus, $C_\infty = \{(x^1, x^2, 0) \in \mathbb{R}^3_+, x^1 + x^2 = 2\}$.

In the case where Player 3 has trivial observation and Players 1 and 2 observe the moves of each other, we have

[10]For each player i, there exists an action $a^i \in A^i$ and two actions a^{-i}, b^{-i} of the other player such that $f^i(a^i, a^{-i}) \neq f^i(a^i, b^{-i})$.

Claim 3.11. $C_\infty = \{(x^1, x^2, 0) \in g(\Delta(A)), x^1 + x^2 \geq 2\}$.

Proof: Let c be a canonical communication device such that the faithful strategy σ^* is a uniform equilibrium of the extended game Γ_c with payoff $x = (x^1, x^2, 0)$ in $I\!R^3$. Consider the deviation σ^1 of Player 1 which plays B, and reports R at every stage, after any history. Similarly, the deviation σ^2 of Player 2 plays R, and reports B at every stage after any history. The two strategy profiles $(\sigma^1, \sigma^{2*}, \sigma^{3*})$ and $(\sigma^{1*}, \sigma^2, \sigma^{3*})$ induce the same profiles of messages reported by the players to the mediator: Player 1 reports R, Player 2 reports B and Player 3 reports nothing. As a consequence, both profiles induce the same distributions of sequences of actions played by Player 3. For any T, denote $(\lambda^T_W, \lambda^T_M, \lambda^T_E)$ the expected empirical distribution of actions of Player 3, averaged over stages $1, \ldots, T$. At each stage, the payoff of Player 1 playing B is at least 1 if Player 3 plays W, and is 2 if Player 3 plays E, so:

$$\gamma^1_{T,c}(\sigma^1, \sigma^{2*}, \sigma^{3*}) \geq \lambda^T_W + 2\lambda^T_E.$$

Similarly,

$$\gamma^2_{T,c}(\sigma^{1*}, \sigma^2, \sigma^{3*}) \geq \lambda^T_W + 2\lambda^T_M.$$

Hence, $\gamma^1_{T,c}(\sigma^1, \sigma^{2*}, \sigma^{3*}) + \gamma^2_{T,c}(\sigma^{1*}, \sigma^2, \sigma^{3*}) \geq 2$, and the equilibrium condition implies $x^1 + x^2 \geq 2$. Thus, $C_\infty \subseteq \{(x^1, x^2, 0) \in g(\Delta(A)), x^1 + x^2 \geq 2\}$.

Conversely, consider a feasible payoff $x = (x^1, x^2, 0)$ such that $x^1 + x^2 \geq 2$. There exists $\lambda \in [0, 1]$ such that $x^1 \geq 2(1-\lambda)$ and $x^2 \geq 2\lambda$: Players 1 and 2 both prefer the payoff x to the payoff induced by Player 3 playing $\lambda M + (1-\lambda)E$. We construct a communication equilibrium with payoff x as follows. The construction is similar to the strategies used in the proof of the Folk theorem, using a main path and a punishment phase. There is a main path of pure action profiles leading to the payoff x, and the mediator recommends to play the actions of the main path as long as the messages reported by Player 1 and 2 coincide with the recommended actions. Otherwise, the play enters a punishment phase where the mediator selects the recommendation of Player 3 by choosing M with probability λ and E with probability $1 - \lambda$ at each stage, independently across stages. This is clearly a canonical communication equilibrium with payoff x and the proof of the claim is complete. □

We now describe the constraints induced by simultaneous punishments in the general case. In the previous example, we constructed a deviation of Player 1 and a deviation of Player 2 inducing the same reported messages to the mediator. First, we generalize this idea.

Definition 3.12. *Given a subset of players $J \subseteq N$, the set of similar decisions of the players in J is defined as:*

$$SD(J) = \left\{ (\delta^i)_{i \in J} \in \prod_{i \in J} \Delta(D^i), \forall i, j \in J, \forall a \in A, \psi^i(\delta^i, a) = \psi^j(\delta^j, a) \right\}.$$

In Example 3.10, consider the decision d^1 in D^1 which plays constantly B and reports constantly R, and similarly d^2 in D^2 which plays constantly R and reports constantly B. Then (d^1, d^2) is a pair of similar decisions of the first two players. The interpretation is that if $(\delta^i)_{i\in J} \in SD(J)$, then the deviation δ^j of player in $j \in J$ induces the same reported signals as the deviation δ^k of player $k \in K$. It follows that the mediator considers any player in j as a potential deviator and thus must punish simultaneously all the players in J.

$SD(J)$ is defined by linear inequalities and includes the faithful decision profile $(\delta^{i*})_{i\in J}$, thus is a non empty polytope. For each player i, $SD(\{i\}) = \Delta(D^i)$, which may be interpreted as: every deviation of player i makes player i a suspect.

Given a probability distribution q on J, we consider the value of the zero-sum game between the mediator who minimizes the payoff $g \cdot q := \sum_{i\in J} q^i g^i$ and a fictitious adversary who selects an element in $SD(J)$.

Definition 3.13. *For each q in $\Delta(N)$, the* punishment level *associated to q is:*

$$l(q) = \max_{\delta\in SD(\operatorname{supp} q)} \min_{a\in A} \sum_{i\in N} q^i g^i_{\delta^i}(a) = \min_{p\in\Delta(A)} \max_{\delta\in SD(\operatorname{supp} q)} \sum_{i\in N} q^i g^i_{\delta^i}(p).$$

The set of jointly rational payoffs *is* $JR = \left\{x \in I\!R^N, \forall q \in \Delta(N), x \cdot q \geq l(q)\right\}$.

A possible interpretation is that the mediator forms a "belief" on the identity of the deviator and punishes this deviator "on average". A sharper interpretation invokes Blackwell's approachability theory. The condition $\forall q \in \Delta(N), x \cdot q \geq l(q)$ ensures the existence of a strategy of the mediator such that for each player i, the average payoff of player i is asymptotically no more than x^i, when other players play faithfully.

Note that the second equality in the definition of $l(q)$ is due to the minmax theorem. Also, remark that if q is the Dirac measure on i, then $l(q) = w^i$: the associated punishment level is the correlated minmax of player i. As a consequence a payoff $x \in JR$ satisfies $x^i \geq w^i$ for each player i, i.e. $JR \subseteq IRC$. In the perfect observation case, we simply have $JR = IRC$ since the only similar decisions are the faithful ones. Also, in the 2-player case, $JR = IRC$ for any signalling structure (consider the mixed action profile where each player plays a minmax strategy against the other).

3.4 Main result: the general characterization

Theorem 3.14. *(Renault and Tomala, 2004) For any supergame, the communication equilibrium payoffs are the feasible payoffs which are both robust to undetectable deviations and jointly rational,*

$$C_\infty = g(\mathcal{P}) \cap JR.$$

The set $g(\mathcal{P}) \cap JR$ is clearly convex and compact, but since the definition of JR involves infinitely many linear inequalities, $g(\mathcal{P}) \cap JR$ might not be a polytope (see example 3.7 of Renault and Tomala, 2004). In the case of perfect observation, we recover Theorem 3.1: $C_\infty = F \cap IRC$, in the 2-player case we recover Theorem 3.9 and in the case of trivial observation, C_∞ is the set of correlated equilibrium payoffs of the stage game, that is $C_\infty = C_1$.

The proof of this result relies on the construction of an auxiliary 2-player repeated game with incomplete information and signals. In this game, player I represents the deviator in the original game, player II represents the mediator, the state of nature is the identity of a deviating player in N which is known by player I only. A particular subclass of uniform equilibria of the auxiliary game is one-to-one mapped to communication equilibria of the original game. These equilibria are characterized via approachability techniques using a result of Kohlberg (1975) and statistical tests *à la* Lehrer (see Lehrer, 1989), similarly as in Renault, 2000.

We now slightly extend Theorem 3.14. Define a set of communication equilibrium payoffs C_* by analogy with the set E_* of Definition 2.16.

Definition 3.15. *C_* is the set of vectors x in $\mathbb{R}^N$ satisfying: $\forall \varepsilon > 0$, there exists a canonical correlation device c and a discount factor λ in $(0, 1]$ such that the faithful strategy σ^* is an ε-Nash equilibrium of the λ-discounted extended game Γ_c with payoff ε-close to x.*

Theorem 3.16. $C_* = g(\mathcal{P}) \cap JR$.

Clearly, $E_* \subseteq C_*$ and it follows from Lemma 2.15 that $C_\infty \subseteq C_*$. Thus, $g(\mathcal{P}) \cap JR \subseteq C_*$. An immediate corollary is that E_∞, E_*, E_λ for any discount factor λ and E_T for any T are subsets of $g(\mathcal{P}) \cap JR$. Hence all usual[11] equilibrium payoffs of the repeated game are included in $g(\mathcal{P}) \cap JR$.

Proof: We prove now $C_* \subseteq g(\mathcal{P}) \cap JR$. We start by introducing some notations. Let c be a canonical communication device and λ be a discount factor. For each action profile a, denote $\mu(c, \lambda)(a)$ the discounted expected number of times where a is recommended by the mediator when all players play the faithful strategies. That is,

$$\mu(c, \lambda)(a) = \mathbb{E}_{c,\sigma^*}\left(\sum_t \lambda(1-\lambda)^{t-1} 1_{\{a_t = a\}}\right)$$

where $1_{\{a_t=a\}}$ equals 1 if a is recommended at stage t and 0 otherwise. This defines a correlated distribution of actions $\mu(c, \lambda) \in \Delta(A)$ such that the discounted payoff of player i under (c, σ^*) is $g^i(\mu(c, \lambda))$.

For $x \in C_*$, for all $\varepsilon > 0$ there exists a correlation device c_ε and a discount factor λ_ε such that $\|x - g(\mu(c_\varepsilon, \lambda_\varepsilon))\| \leq \varepsilon$. Choose a sequence $\varepsilon_m \to 0$ such that $\mu_m := \mu(c_{\varepsilon_m}, \lambda_{\varepsilon_m})$ converges to some $\mu \in \Delta(A)$. Then, $x = g(\mu)$.

Assume $x \notin g(\mathcal{P})$ thus $\mu \notin \mathcal{P}$. There exists a player i and $\delta^i \in \Delta(D^i)$ such that $\forall a \in A$, $\psi^i(\delta^i, a) = f(a)$ and $g^i_{\delta^i}(\mu) > g^i(\mu)$. By continuity, there exists $\eta > 0$ and an integer m_0 such that for all $m \geq m_0$, $g^i_{\delta^i}(\mu_m) > g^i(\mu_m) + \eta$. Consider the

[11] One may *a fortiori* consider subgame-perfect or sequential equilibria.

strategy τ^i of player i which plays δ^i at every stage. Since $\psi^i(\delta^i, a) = f(a)$ $(\forall a)$, the distribution of signals reported to the mediator is the same under the faithful strategies or under this unilateral deviation of player i, thus the distribution of recommended actions is also the same. The discounted payoff of player i under $(c_m, \tau^i, \sigma^{*-i})$ is thus $g^i_{\delta^i}(\mu_m) \leq g^i(\mu_m) + \varepsilon_m$, as σ^* is an ε_m-Nash equilibrium of the λ_m-discounted extended game. This contradicts $g^i_{\delta^i}(\mu_m) > g^i(\mu_m) + \eta$ for all m large enough.

Assume now $x \notin JR$. Then there exist $q \in \Delta(N)$, $\delta \in SD(q)$ such that $\min_{p \in \Delta(A)} \sum_i q^i g^i_{\delta^i}(p) > \sum_i q^i g^i(\mu)$. By continuity again, there exists $\eta_1 > 0$ and an integer m_1 such that for all $m \geq m_1$ and for all $p \in \Delta(A)$, $\sum_i q^i g^i_{\delta^i}(p) > \sum_i q^i g^i(\mu_m) + \eta_1$. For each i in the support of q, consider the strategy τ^i of player i which plays δ^i at every stage. Since for all i, j in the support of q, $\psi^i(\delta^i, a) = \psi^j(\delta^j, a)$ $(\forall a)$, the distribution of signals reported to the mediator and of recommended actions does not depend on the choice of i in the support of q. For each m, let $\mu_m(\delta)$ be the expected discounted frequency of actions when the communication device is c_m and some player i in the support of q plays τ^i. The discounted payoff of player i under $(c_m, \tau^i, \sigma^{*-i})$ is thus $g^i_{\delta^i}(\mu_m(\delta)) \leq g^i(\mu_m) + \varepsilon_m$, as σ^* is an ε_m-Nash equilibrium of the λ_m-discounted extended game. Thus, $\sum_i q^i g^i_{\delta^i}(\mu_m(\delta)) \leq \sum_i q^i g^i(\mu_m) + \varepsilon_m$. But from the choice of δ, $\sum_i q^i g^i_{\delta^i}(\mu_m(\delta)) > \sum_i q^i g^i(\mu_m) + \eta_1$ for $m \geq m_1$, a contradiction. $\square$

3.5 Applications to neighbouring networks

We provide now examples of computation of the set C_∞. A class of signalling structures of interest is described by *neighbouring networks*, where each player perfectly observes the actions of a subset of players, namely his neighbours. This model is studied by Ben-Porath and Kahnemann (1996) and Renault and Tomala (1998). It is shown in these papers that if the neighbouring graph has good connectivity properties, then any feasible and individually rational payoff is an equilibrium payoff. In particular, strong 2-connectedness[12] of the graph is necessary. We now apply Theorem 3.14 to compute equilibrium payoffs for some graphs which are not strongly 2-connected.

The oriented circle. Assume that there are three players $N = \{1, 2, 3\}$, and for each i in N the observation of player $i + 1$ is the action of player i (by convention if $i = 3$, player $i + 1$ is player 1). Then $g(\mathcal{P}) = g(\Delta(A))$, and $SD(N)$ is a singleton reduced to the faithful decision profile.

An element (δ^1, δ^2) in $SD(\{1, 2\})$ can be parametrized by a mapping $\alpha : A^1 \to \Delta(A^1)$ such that δ^1 plays $\alpha(a^1)$ if recommended a^1 by the mediator, and reports faithfully the action played by player 3, whereas δ^2 plays faithfully the recommendation of the mediator but if a^1 is played by player 1, it reports to

[12] For any two vertices i, j, there exists two directed and disjoint paths from i to j.

the mediator a message drawn from the probability distribution $\alpha(a^1)$. For $q = (q^1, q^2, 0)$ in $\Delta(\{1,2\})$ we have,

$$l(q) = \min_{p\in\Delta(A)} \left(q^2 g^2(p) + q^1 \max_{\alpha : A^1 \to A^1} \sum_{a\in A} p(a) g^1(a^{-1}, \alpha(a^1)) \right).$$

For instance, with the payoffs of example 3.10, we get $l(1/2, 1/2, 0) = 1$. For the 3-player oriented circle with arbitrary payoffs, Theorem 3.14 yields :

$$\begin{aligned} C_\infty = & \Big\{ x = (x^1, x^2, x^3) \in g(\Delta(A)), \forall \lambda \in [0,1], \\ & \lambda x^1 + (1-\lambda) x^2 \geq \min_{p\in\Delta(A)} (\lambda g^2(p) + (1-\lambda) \max_{\alpha : A^1 \to A^1} \sum_{a\in A} p(a) g^1(a^{-1}, \alpha(a^1))), \\ & \lambda x^2 + (1-\lambda) x^3 \geq \min_{p\in\Delta(A)} (\lambda g^3(p) + (1-\lambda) \max_{\alpha : A^2 \to A^2} \sum_{a\in A} p(a) g^2(a^{-2}, \alpha(a^2))), \\ & \lambda x^3 + (1-\lambda) x^1 \geq \min_{p\in\Delta(A)} (\lambda g^1(p) + (1-\lambda) \max_{\alpha : A^3 \to A^3} \sum_{a\in A} p(a) g^3(a^{-3}, \alpha(a^3))) \Big\}. \end{aligned}$$

Independent rooms. Assume that there is a partition of the set of players C_1, $\ldots$, C_K, and that for each i in N, the observation of player i is the actions of the other players in the partition cell $C_{k(i)}$ containing i. One can interpret $C_1, \ldots, C_K$ as independent rooms and each player observes what happens in his own room only. Denote K_1 (resp. K_2, resp. K_{3+}) the set of indexes $k \in \{1, \ldots, K\}$ such that room C_k contains exactly 1 player (resp. exactly 2 players, resp. at least 3 players).

A player alone in his room may play a best reply at each stage without affecting the observation of any other player. Thus,

$$\mathcal{P} = \Big\{ p \in \Delta(A), \forall i \in \cup_{k\in K_1} C_k, \forall a^i \in A^i, b^i \in A^i,$$

$$\sum_{a^{-i}\in A^{-i}} p(a^i, a^{-i}) g^i(a^i, a^{-i}) \geq \sum_{a^{-i}\in A^{-i}} p(a^i, a^{-i}) g^i(b^i, a^{-i}) \Big\}.$$

Let J be a subset of players with at least 2 elements. If J contains players from separate rooms, or if J contains players in a room with at least 3 players, then $SD(J)$ reduces to the faithful decision profile. We are thus left with the case $J = C_k$ for some 2-player room $C_k = \{i, j\}$. In this case, similar decisions $\delta = (\delta^i, \delta^j)$ in $SD(J)$ can be parametrized by two mappings $\alpha^i : A^i \to \Delta(A^i)$ and $\alpha^j : A^j \to \Delta(A^j)$. Under δ^i, player i receiving the recommendation r^i in A^i plays an action a^i according to $\alpha^i(r^i)$, observes the action a^j played by player j and chooses a reported message according to $\alpha^j(r^j)$. Similarly under δ^j, player j receiving the recommendation r^j in A^j plays an action a^j according to $\alpha^j(r^j)$, observes the action a^i played by player i and chooses a reported message according to $\alpha^i(r^i)$. The distribution of the messages (r^i, r^j) reported to the mediator is the independent distribution $\alpha^j(r^j) \otimes \alpha^i(r^i)$, both under the deviation δ^i and under the deviation δ^j. We get for $q = (q^i, q^j) \in \Delta(\{i,j\})$:

$$l(q) = \min_{p \in \Delta(A)} \max_{\alpha^i, \alpha^j} \sum_{a \in A} p(a) \left(q^i g^i(a^{-i}, \alpha^i(a^i)) + q^j g^j(a^{-j}, \alpha^j(a^j)) \right).$$

We obtain $C_\infty = g(\mathcal{P}) \cap JR$, where:

$$\begin{aligned} JR &= \Big\{ x = (x^i)_{i \in N} \in IRC, \forall k \in K_2, \forall i,j \in C_k, \forall \lambda \in (0,1), \\ &\lambda x^i + (1-\lambda) x^j \geq \min_{p \in \Delta(A)} \max_{\alpha^i, \alpha^j} \sum_{a \in A} p(a) \left(\lambda g^i(a^{-i}, \alpha^i(a^i)) + (1-\lambda) g^j(a^{-j}, \alpha^j(a^j)) \right) \Big\}. \end{aligned}$$

In the particular case where there is no single room, $K_1 = \emptyset$, we have:

$$\begin{aligned} C_\infty &= \Big\{ x = (x^i)_{i \in N} \in g(\Delta(A)) \cap IRC, \forall k \in K_2, \forall i,j \in C_k, \forall \lambda \in (0,1), \\ &\lambda x^i + (1-\lambda) x^j \geq \min_{p \in \Delta(A)} \max_{\alpha^i, \alpha^j} \sum_{a \in A} p(a) \left(\lambda g^i(a^{-i}, \alpha^i(a^i)) + (1-\lambda) g^j(a^{-j}, \alpha^j(a^j)) \right) \Big\}. \end{aligned}$$

3.6 The Folk Theorem for communication equilibria

Now, we apply Theorem 3.14 to find conditions ensuring that the Folk Theorem extends to the supergame, namely all feasible and individually rational payoffs are communication equilibrium payoffs: $C_\infty = F \cap IRC$. Intuitively, this is the case if no player has a profitable and undetectable deviation from any distribution of actions, and if the mediator can punish any detected deviation. A simple condition is the following.

Lemma 3.17. *Assume that there are at least 2 players and for all distinct players i, j, for all decisions δ^i in $\Delta(D^i)$, δ^j in $\Delta(D^j)$:*

$$\left(\forall a \in A, \psi^i(\delta^i, a) = \psi^j(\delta^j, a) \right) \implies \left(\forall a \in A,\ g^i_{\delta^i}(a) \leq g^i(a) \text{ and } g^j_{\delta^j}(a) \leq g^j(a) \right),$$

then $C_\infty = F \cap IRC$.

In words, the condition states that if the decision δ^i cannot be distinguished from δ^i by the mediator, then δ^i (resp. δ^j) is not a profitable deviation for player i (resp. player j).

Proof: The condition implies that,

$$(\forall a \in A, \psi^i(\delta^i, a) = f(a)) \implies (\forall a \in A,\ g^i_{\delta^i}(a) \leq g^i(a)).$$

That is, an undetectable deviation of player i does not increase his payoff. It follows that $\mathcal{P} = \Delta(A)$.

Let us compute JR. From the expression of punishments levels in Definition 3.13, $l(q) \leq \min_{p\in\Delta(A)} \sum_i q^i g^i(p)$ for each $q \in \Delta(N)$ with at least two points in its support. It follows that $F \cap IRC \subseteq JR$ and thus $C_\infty = F \cap IRC$. $\square$

We deduce a condition on signals such that $C_\infty = F \cap IRC$ for all payoff functions. Let D^{*i} be the set of mixed decisions of player i which play recommended actions faithfully with probability one. That is, under $\delta^i \in D^{*i}$, $\alpha(a^i) = a^i$ almost surely. Note that this decision may report signals arbitrarily.

Proposition 3.18. *Assume that there are at least 2 players and for all distinct players i, j, for all decisions δ^i in $\Delta(D^i)$, δ^j in $\Delta(D^j)$:*

$$\big(\forall a \in A, \psi^i(\delta^i, a) = \psi^j(\delta^j, a)\big) \implies \big(\delta^i \in D^{*i} \text{ and } \delta^j \in D^{*j}\big).$$

Then, $C_\infty = F \cap IRC$ for any payoff function.

The proof follows directly from the preceding lemma. The condition of this proposition states that if the mediator cannot distinguish δ^i from δ^j, then these decisions always play actions faithfully. To get a better feeling of signalling structures for which this condition is fulfilled, we consider now deterministic signals $(u^i)_i = (f^i(a))_i$ for which we are able to provide simpler conditions.

No Undetectable Deviation. $\forall i \in N$, $\forall a^i, b^i \in A^i$, $(b^i \geq a^i \implies b^i = a^i)$.

This condition says that if $b^i \neq a^i$, either there exist $j \neq i$ and a_{-i} such that $f^j(a^i, a^{-i}) \neq f^j(b^i, a^{-i})$, or there exist a^{-i}, b^{-i} such that $f^i(a^i, a^{-i}) \neq f^i(a^i, b^{-i})$ and $f^i(b^i, a^{-i}) = f^i(b^i, b^{-i})$.

This means that either there is an action profile and a player j who gets different signals under a^i and b^i, or player i may acquire finer information about a^{-i} by playing a^i rather than by playing b^i. In both cases, the deviation from a^i to b^i is detectable by the deviator: player j or player i will report an unexpected signal.

No Similar Deviations. For each pair (i, j) of distinct players, there exists (a^i, a^j) in $A^i \times A^j$ such that:
$\forall (b^i, b^j) \neq (a^i, a^j), \exists k \neq i, j, \exists a^{-ij} \in \prod_{k' \in K} A^{k'}$ s.t. $f^k(b^i, a^j, a^{-ij}) \neq f^k(a^i, b^j, a^{-ij})$.

This condition expresses the ability of the mediator to distinguish the deviations of player i from those of player j. If the mediator recommends (a^i, a^j), then if player i (resp. player j) deviates to some b^i (resp. b^j), there is another player k who gets different signals under the deviation of player i and that of player j. Then, the mediator gets different reported signals under these two deviations. This is satisfied for instance if the actions of each player are directly observed by at least two other players (see Ben-Porath and Kahnemann, 1996).

Proposition 3.19. *(1) Assume that there are two players. The (deterministic) signals satisfy* No Undetectable Deviation *if and only if* $C_\infty = F \cap IRC$ *for all payoff functions.*

(2) Assume that there are at least three players. If the (deterministic) signals satisfy No Undetectable Deviation *and* No Similar Deviations*, then* $C_\infty = F \cap IRC$ *for all payoff functions. If* No Undetectable Deviation *is not satisfied, then there exists a payoff function for which* $C_\infty \neq F \cap IRC$*.*

The conditions here are reminiscent of those found in Fudenberg et al. (1994) and Kandori and Matsushima (1998). In both of these works, the Folk Theorem obtains whenever all deviations can be detected and the deviator is identified. However, the conditions given in these papers are more stringent than ours, because they study a stronger equilibrium concept, namely the sequential equilibrium of discounted games.

To see that *No Undetectable Deviation* is necessary, assume that $b^i \geq a^i$ and $b^i \neq a^i$. One may then construct a payoff function that mimics the Prisoner's Dilemma of Example 2.1 where b^i is the defect action D^i and a^i is the cooperate action C^i. Since b^i strictly dominates a^i and $b^i \geq a^i$, a^i is never played by player i thus excluding some payoff in $F \cap IRC$ from C_∞. To see that it is sufficient for two-player games, remark that there exists a correlated action profile such that each player gets no more than his minmax value.

We leave as an open question the finding of necessary and sufficient conditions on signals to get $C_\infty = F \cap IRC$ for all payoff functions in games with at least three players.

3.7 Applications to public signals

In this section, we study supergames with public signals. That is, there is a set of public signals $\mathbf{U}$ and a transition $f : A \to \Delta(\mathbf{U})$, and when the action profile a is played, a signal u is drawn from $f(a)$ and publicly announced.

Claim 3.20. *In a supergame with public signals extended by communication, one may assume without loss of generality that the mediator observes the actually realized public signals.*

The intuition is that anything that is publicly known can be assumed to be known by the mediator as well. To prove this claim, we construct a (non-canonical) communication device where the mediator issues a vector of recommendations at each stage, one recommendation for each possible public signal, and players' strategies are required to follow the recommendation attached to the actually observed public signals. This shows that allowing the mediator to observe the public signals does not increase the set of equilibrium outcomes. Conversely, in the extended game where the mediator does observe the realized public signals, he may well ignore this additional information, thus the set of equilibrium outcomes does not decrease either.

We assume for the rest of this section that the mediator observes the public signals, so that a decision of each player i is now parametrized by the mapping α^i only. A mixed decision δ^i is then identified with a mapping $\alpha^i : A^i \to \Delta(A^i)$, and for each action profile a, the distribution of reported signals $\psi^i(\delta^i, a)$ is identified with $f(\alpha^i(a^i), a^{-i})$.

Definition 3.21. *Two mixed actions x^i, $y^i \in \Delta(A^i)$ of player i are equivalent if $\forall a^{-i} \in A^{-i}$, $f(x^i, a^{-i}) = f(y^i, a^{-i})$. We write then $x^i \sim^i y^i$.*

That is, the distribution of the public signal is the same under x^i and under y^i, irrespective of a^{-i}. The formula of $g(\mathcal{P}) \cap JR$ is simplified as follows.

Proposition 3.22. *In a game with public signals,*

$$\mathcal{P} = \Big\{p \in \Delta(A), \forall i \in N, \forall \alpha^i : A^i \to \Delta(A^i) \text{ s.t. } \alpha^i(a^i) \sim^i a^i (\forall a^i),$$
$$\sum_{a \in A} p(a) g^i(a^i, a^{-i}) \geq \sum_{a \in A} p(a) g^i(\alpha^i(a^i), a^{-i})\Big\}.$$

For each $J \subseteq N$, a tuple of similar decisions $(\delta^i)_{i \in J} \in SD(J)$ is represented by a tuple of mappings $(\alpha^i)_{i \in J}$ with $\alpha^i : A^i \to \Delta(A^i)$ such that,

$$\forall i, j \in J, \forall a \in A, f(\alpha^i(a^i), a^{-i}) = f(\alpha^j(a^j), a^{-j}).$$

For each $q \in \Delta(N)$,

$$l(q) = \min_{p \in \Delta(A)} \max_{SD(\operatorname{supp} q)} \sum_{i \in N} q^i \sum_a p(a) g^i(\alpha^i(a^i), a^{-i}).$$

The proof is straightforward given that the mediator knows the public signals and that decisions are viewed as mappings $\alpha^i : A^i \to \Delta(A^i)$. The condition $\psi^i(\delta^i, a) = f(a)$ $(\forall a)$ then reduces to $\alpha^i(a^i) \sim^i a^i$ $(\forall a^i)$. Theorem 3.14 then applies so that $C_\infty = g(\mathcal{P}) \cap JR$, but the expression of this set is simpler. As an application, we get simplified conditions for the Folk Theorem for communication equilibria of supergames with public signals.

No Undetectable Deviation. $\forall i \in N$, $x^i, y^i \in \Delta(A^i)$, $(x^i \sim^i y^i \implies x^i = y^i)$.

No Similar Deviations. For all $i \neq j$, for all α^i, α^j,
$$(\forall a \in A, f(\alpha^i(a^i), a^{-i}) = f(\alpha^j(a^j), a^{-j})\ ) \implies (\forall (a^i, a^j), \alpha^i(a^i) = a^i \text{ and } \alpha^j(a^j) = a^j\ ).$$

Proposition 3.23. *(1) Assume that there are two players. The public signals satisfy* No Undetectable Deviation *if and only if $C_\infty = F \cap IRC$ for all payoff functions.*

(2) Assume that there are at least three players. If the public signals satisfy No Undetectable Deviation *and* No Similar Deviations, *then $C_\infty = F \cap IRC$ for all payoff functions. If* No Undetectable Deviation *is not satisfied, then there exists a payoff function for which $C_\infty \neq F \cap IRC$.*

The necessity of No Undetectable Deviation is as in Proposition 3.19. We leave again as an open question the finding of necessary and sufficient conditions on public signals to get $C_\infty = F \cap IRC$ for all payoff functions in games with at least three players.

4 Extensions and conjectures

4.1 Discounted payoffs and open problems

We already mentioned a few open problems related to discounted equilibrium payoffs. We conjecture the following statements, for any repeated game with signals.

C1: E_λ converges, as λ goes to 0, to cl $\left(\bigcup_{\lambda>0} E_\lambda\right)$?

C2: $E_\lambda \subseteq E_\infty$?

C3: $E_\infty = E_*$?

4.2 Getting rid of the mediator

An important open problem is to give a characterization of E_∞ for any supergame without mediator. The first results in this direction are due to Lehrer. In particular, Lehrer (1989) finds a class of signals for which E_∞ is fully characterized. Signals are *semi-standard* if each player's action set is endowed with an equivalence relation $\sim^i$ and when the action profile a is played, the profile of equivalence classes is publicly announced. These signals have the following special features: they are public, the information on player i's action is the same for every player $j \neq i$. It follows that all actions of player j are equally informative on the action of player i. As a consequence $b^i \geq a^i$ if and only if $b^i \sim^i a^i$. Denote $\mathcal{D}$ the set of distributions $p \in \mathcal{P}$ generated by mixed action profiles, i.e. which are products of independent mixed actions. Lehrer (1989) shows that:

Theorem 4.1. *(Lehrer, 1989) In a supergame with semi-standard signals,*

$$E_\infty = \mathrm{conv} g(\mathcal{D}) \cap IR.$$

So far, semi-standard signalling is the only non-trivial specification of signalling structures for which E_∞ is characterized for n-player games and any payoff function. The extension to any supergame is difficult. As an interesting example, even the specific and apparently simple case of two players where one of them has trivial observation is still unsolved. Here are two open questions of interest.

1. Give conditions on signals that ensure $E_\infty = C_\infty$. Similarly, find conditions for which C_∞ coincides with uniform *correlated* equilibrium payoffs, where the mediator sends message only once, at a pre-play stage.

2. Find non-trivial conditions on signals, other than semi-standard, for which E_∞ can be characterized for any payoff function.

Another interesting sub-problem pertains to the punishment level in the repeated game. Say that player i can be forced to the payoff z^i in $I\!R$ if: $\forall \varepsilon > 0$, $\exists \bar{\sigma}^{-i} \in \Sigma^{-i}$, $\exists \bar{T}$ s.t.: $\forall \sigma^i \in \Sigma^i$, $\forall T \geq \bar{T}$, $\gamma^i_T(\sigma^i, \bar{\sigma}^{-i}) \leq z^i + \varepsilon$. The punishment level v^i_∞ of player i in the supergame is defined as follows (see Renault Tomala 1998):

$$v^i_\infty = \inf\{z^i \in I\!R, \text{ player i can be forced to } z^i\}.$$

In general, punishment levels depend on the stage game but also on the signalling structure. The following example from Renault Tomala (1998) is solved in Gossner and Tomala (2007).

Example 4.2. *This is a 3-player game where Player* 1 *chooses the row, Player* 2 *chooses the column and Player* 3 *chooses the matrix. The payoff of Player* 3 *is the following.*

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \\ W \end{array} \qquad \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \\ E \end{array}$$

The signalling structure is such that Players 1 *and* 2 *have perfect observation whereas Player* 3 *observes the action of Player* 2 *only.*

On this example, one has $v^3 = -\frac{1}{4}$ and $w^3 = -\frac{1}{2}$. The independent minmax is obtained when Players 1 and 2 randomize independently and evenly. The correlated minmax is obtained when Players 1 and 2 play (A, A) and (B, B) with respective probabilities $(\frac{1}{2}, \frac{1}{2})$. Let us denote v^3_∞ the minmax in the repeated game, that is the smallest number v such that for all $\varepsilon > 0$, Players 1 and 2 have a strategy forcing the average payoff of Player 3 below $v + \varepsilon$, in all long enough repeated games. It is easy to see that $v^3_\infty \leq -\frac{3}{8}$ by considering the following strategy. Players 1 and 2 randomize independently and evenly at odd stages, and at even stages they play the action selected by Player 1 at the previous stage. Under this strategy, the payoff of Player 3 is $-\frac{1}{4}$ at odd stages. At every even stage, the distribution of actions of Players 1 and 2, conditional on the signals of Player 3, is (A, A) with probability $\frac{1}{2}$ and (B, B) with probability $\frac{1}{2}$.

Using tools borrowed from Information Theory, Gossner and Tomala (2007) show that $v^3_\infty = (x^2 + (1-x)^2)/2$, where x is the unique solution in $[0, \frac{1}{2}]$ of the equation $-x \log_2 x - (1-x)\log_2(1-x) = \frac{1}{2}$. Gossner and Tomala characterize the punishment level of a single player in repeated games as the value of a static optimization problem on a set of probability distributions under entropy constraints. Therein, the following assumptions are used: all players $j \neq i$ have the same information which includes the information of player i, the signal of player i does not depend on his own action. Computing the punishment level without these assumptions is an open problem.

4.3 Banach limits

An approach to undiscounted infinitely repeated games, alternative to uniform equilibria, is to define a limit payoff for any stream of stage payoffs, usually by extending the limit operator to all bounded sequences.

Definition 4.3. *Let L be a bounded measurable real-valued mapping defined on bounded sequences of real numbers and which coincides with the limit for converging sequences. The L-payoff for player i induced by the strategy profile σ is $\gamma^i_L(\sigma) = \mathbb{E}_\sigma(L(\bar{g}^i))$, where $\bar{g}^i$ is the sequence $\left(\frac{1}{T}\sum_{t=1}^T g^i(a_t)\right)_T$. An L-equilibrium of Γ is a Nash equilibrium of the game with strategy sets Σ^i and payoff functions γ_L.*

An example is when L is a Banach limit, i.e. a linear mapping defined on real bounded sequences such that $\liminf \leq L \leq \limsup$. The existence of Banach limits can be proved using the Hahn-Banach Theorem. One can also define equilibria using the $\liminf$ or the $\limsup$, see for instance Lehrer (1989). A lower equilibrium is a strategy profile σ such that for each player i, $\gamma^i_T(\sigma)$ converges as T goes to infinity, and for each player i and strategy τ^i, $\liminf_T \gamma^i_T(\tau^i, \sigma^{-i}) \leq \lim_T \gamma^i_T(\tau^i, \sigma^{-i})$. The upper equilibrium is defined similarly using the limsup. Denote $\underline{E}$ the set of lower equilibrium payoffs, $\overline{E}$ the set of upper equilibrium payoffs and E_L the set of equilibrium payoffs for the Banach limit L. It is straightforward that for any Banach limit L:

$$E_\infty \subseteq \overline{E} \subseteq E_L \subseteq \underline{E}$$

and that all these sets equal $F \cap IR$ for supergames with perfect observation.

Lehrer (1989) showed that for supergames with signals, the set of lower equilibrium payoffs can be a strict superset of the other equilibrium payoff sets. A simple example, somewhat different from the one given by Lehrer, is provided by Example 3.10. In this example, we claim that $\underline{C} = F \cap IRC$ while C_∞ is a strict subset. The equilibrium construction is as follows. The mediator recommends actions following a pure sequence of actions achieving the target limit payoff. Upon observing a deviation, the mediator recommends player 3 to punish either player 1 (play M) or player 2 (play E) in the following deterministic way. First, play M for a large number N_1 of stages. Then, play E for a much larger number N_2 of stages in such a way that N_1 is negligible with respect to N_2. Then, revert to playing E for N_3 stages with $N_1 + N_2$ negligible with respect to N_3. Continue to alternate like this, each time making the sum of sizes of past blocks negligible with respect to the size of the new block. Then, the liminf of player 1's payoff is found by computing the average at the end of an odd numbered block, and is close to 0.

It is an open problem to decide whether $E_\infty = E_L$ for any supergame. Lehrer (1989) proved that it is the case for semi-standard signalling. The proof of Theorem 3.14 (Renault and Tomala, 2004) shows that $C_\infty = C_L$. Both proofs rely on the characterization. In lack of a characterization of E_∞, another kind of proof should be found, using the specific structure of supergames, where the same known stage game is repeated over and over.

References

[1] Abreu D., D. Pearce et E. Stacchetti. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 58, 1041–1063, 1990.

[2] Aubin J.P. Applied Abstract Analysis. Wiley, 1977.

[3] Aumann R.J. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, **1**, 67-96, 1974.

[4] Aumann R.J. and L. S. Shapley. Long-term competition—A game theoretic analysis. In N. Megiddo, editor, *Essays on game theory*, pages 1–15. Springer-Verlag, New-York, 1994.

[5] Aumann R.J. and M. Maschler. Repeated games with incomplete information, with the collaboration of R. Stearns. M.I.T. Press, 1995.

[6] Benoit J-P. and V. Krishna Finitely repeated games, *Econometrica*, **53**, 905–922, 1985.

[7] Benoit J.-P. and V. Krishna. Nash equilibria of finitely repeated games, *International Journal of Game Theory*, **16**, 197–204, 1987.

[8] Ben-Porath E. and M. Kahneman. Communication in repeated games with private monitoring. *Journal of Economic Theory*, **70**, 281–297, 1996.

[9] Forges F. An Approach to Communication Equilibria. *Econometrica*, 54, 1375–1385, 1986.

[10] F. Forges, J.-F. Mertens, and A. Neyman. A counterexample to the folk theorem with discounting. *Economics Letters*, 20:7–7, 1986.

[11] Fudenberg D. and E. Maskin. The Folk Theorem in repeated games with discounting or with incomplete information. *Econometrica*, **54**, 533–554, 1986.

[12] Fudenberg D. and D. Levine. An approximate Folk Theorem with imperfect private information. *Journal of Economic Theory*, **54**, 26–47, 1991.

[13] Fudenberg D. and D. Levine. Efficiency and observability with long-run and short-run players. *Journal of Economic Theory*, 62:103–135, 1994.

[14] Fudenberg D., D. Levine, and E. Maskin. The folk theorem with imperfect public information. *Econometrica*, 62:997–1039, 1994.

[15] Fudenberg D., D. Levine, and S. Takahashi. Perfect public equilibrium when players are patient. *Games and Economic Behavior*, 61:27–49, 2007.

[16] Gossner O. The Folk Theorem for finitely repeated games with mixed strategies. *International Journal of Game Theory*, **24**, 95–107, 1995.

[17] Gossner O. and T. Tomala. Secret correlation in repeated games with imperfect monitoring. *Mathematics of Operations Research*, 32, 413–424, 2007.

[18] Hillas J. and M. Liu. Repeated Games with Partial Monitoring: the Stochastic Signaling Case. Game Theory and Information 9605001, EconWPA, 1996.

[19] Ely J., J. Hörner, and W. Olszewski. Belief-free equilibria in repeated games. *Econometrica*, 73:377–415, 2005.

[20] Kandori M. and H. Matsushima. Private observation, communication and collusion. *Econometrica*, 66:627–652, 1998.

[21] Kohlberg E. Optimal strategies in repeated games with incomplete information. *International Journal of Game Theory*, 4, 7–24, 1975.

[22] Kuhn H.W. Extensive games and the problem of information. in *Contributions to the Theory of Games* (Kuhn and Tucker, eds.), vol. II, pp. 193-216. Annals of Mathematical Studies 28, Princeton University Press, 1953.

[23] Lehrer E. Lower Equilibrium Payoffs in Two-Player Repeated Games with Non-observable Actions. *International Journal of Game Theory*, 18, 57–89, 1989.

[24] Lehrer E. Nash equilibria of n-player repeated games with semi-standard information. *International Journal of Game Theory*, 19, 191–217, 1990.

[25] Lehrer E. Correlated Equilibria in two-Player Repeated Games with non-Observable Actions. *Mathematics of Operations Research*, 17, 175–199, 1992a.

[26] Lehrer E. On the Equilibrium Payoffs Set of two-Player Repeated Games with Imperfect Monitoring. *International Journal of Game Theory*, 20, 211–226, 1992b.

[27] Lehrer E. Two-player repeated games with nonobservable actions and observable payoffs. *Mathematics of Operations Research*, 17, 200–224, 1992c.

[28] Lippman S. Criterion Equivalence in Discrete Dynamic Programming. Operations Research 17, 920-923, 1969.

[29] Mailath G., S. Matthews and T. Sekiguchi. Private strategies in Finitely Repeated Games with Imperfect Public Monitoring. Contributions to Theoretical Economics, Vol. 2, article 2, 2002.

[30] Mailath G. and L. Samuelson. *Repeated Games and Reputations: Long-Run Relationships.* Oxford University Press, 2006.

[31] Mertens J-F, S. Sorin et S. Zamir. Repeated games. CORE discussion paper 9420, Louvain-la-Neuve, 1994.

[32] Moulin H. and J.-P. Vial. Strategically zero-sum games: The class of games whose completely mixed equilibria cannot be improved upon. International Journal of Game Theory, vol. 7, issue 3/4, p.201–221, 1978.

[33] Myerson, R. Multistage games with communication. *Econometrica*, 54, 323–358, 1986.

[34] Renault J. and T. Tomala. Repeated Proximity Games. *International Journal of Game Theory*, 2, 539–559, 1998.

[35] Renault J. 2-player repeated games with lack of information on one side and state independent signalling. *Mathematics of Operations Research*, 25, 552–572, 2000.

[36] Renault J., S. Scarlatti and M. Scarsini. A Folk Theorem for minority games. *Games and Economic Behavior*, 53, 208–230, 2005.

[37] Renault J., S. Scarlatti and M. Scarsini. Discounted and Finitely Repeated Minority Games with Public Signals. Mathematical Social Sciences, vol.56, pp.44–74, 2008.

[38] Renault J. and T. Tomala. Communication equilibria in supergames. *Games and Economic Behavior*, 49, 313–344, 2004.

[39] Rubinstein, A. Equilibrium in supergames. Center for Research in Mathematical Economics and Game Theory, Research Memorandum 25, 1977.

[40] Sekiguchi T. A negative result in finitely repeated games with product structure. *Economic Letters*, 74, 67-70, 2001.

[41] Sekiguchi T. Uniqueness of equilibrium payoffs in finitely repeated games with imperfect monitoring. *Japanese Economic Review*, 56, 317-331, 2005.

[42] Shapley, L.S. Stochastic games. *Proceedings of the National Academy of Sciences of the U.S.A.*, 39, 1095-1100, 1953.

[43] Sorin, S. Repeated games with complete information. In Handbook of Game Theory with Economic Applications, R.J. Aumann and S. Hart (eds.), Elsevier Science Publishers, 71-107, 1992.

[44] Sorin, S. On repeated games with complete information. *Mathematics of Operations Research*, **11**, 147–160, 1986.

[45] Tomala, T. Pure equilibria of repeated games with public observation. *International Journal of Game Theory*, 27(1), 93-109, 1998.

[46] Tomala, T. Nash equilibria of repeated games with observable payoff vector. *Games and Economic Behavior*, 28, 310–324, 1999.