Learning with Minimal Information in Continuous Games[*]

Sebastian Bervoets[†] Mario Bravo[‡] and Mathieu Faure[§]

September 5, 2018

Abstract

While payoff-based learning models have almost exclusively focused on finite actions games, where players can test every action, it is harder to design such learning processes for continuous games. This is what we do in this paper. In our process, players update their actions according to variations in payoff between current and past actions. It is simple to implement by players as it requires no sophistication and no knowledge of the game. We show that this process is well-behaved, i.e. actions remain in the state space, and converges to Nash in large classes of games.

Keywords: Payoff-based Learning, Continuous games, Nash equilibrium, Stochastic approximation.

[†]Aix-Marseille University, CNRS, EHESS, Centrale Marseille, AMSE. Email: sebastian.bervoets@univ-amu.fr

[‡]Universidad de Santiago de Chile, Departamento de Matemática y Ciencia de la Computación. Email: mario.bravo.g@usach.cl

[§]Aix-Marseille University, CNRS, EHESS, Centrale Marseille, AMSE. Email: mathieu.faure@univ-amu.fr

1 Introduction

In this paper we construct a simple stochastic learning rule with the three following properties: (i) it is designed for games with continuous action sets; (ii) it requires no sophistication from the players and (iii) it converges to Nash equilibria in large classes of games. The question of convergence to Nash equilibria by agents playing a game repeatedly has given rise to a large body of literature on learning. One branch of this literature explores whether there are learning rules - deterministic or stochastic - which would converge to Nash equilibria in any game (see i.a. Hart and Mas-Colell (2003), Hart and Mas-Colell (2006), Babichenko (2012), Foster and Young (2006), Germano and Lugosi (2007)). Another branch, to which this paper contributes, focuses on specific learning rules and on the understanding of their asymptotic behavior.

Both branches have almost exclusively addressed the issue of learning in discrete games (i.e. games where the set of strategies is finite). However, many economic variables such as price, effort, time allocation, are non-negative real numbers, and thus are continuous. Typical learning models that have been designed for discrete games cannot be adapted to continuous settings without major complications, since they usually rely on assigning a positive probability to choosing each action, which cannot be done in continuous games. In this paper we introduce a learning rule designed for continuous games, which we call the *dampened gradient approximation process* (DGAP), and we analyze its behavior in several well-known classes of games.

Learning rules can be more or less demanding in terms of players' sophistication and of the amount of information required to implement them. The DGAP belongs to the category of so-called *payoff-based* or *completely uncoupled* learning rules, meaning that players know nothing about the payoff functions (neither theirs nor those of their opponents), and they know nothing about the other players' actions, nor about their payoffs. They may not even be aware that they are playing a game. They only observe their own realized payoffs after each iteration of the game and make decisions based on these observations.

Agents aim at maximizing their payoffs by choosing an action. If players knew the gradient of their utility function at every point, a natural learning process in continuous games would be for agents to follow a gradient method (see for instance Arrow and Hurwicz (1960)). However, because players neither know the payoff functions nor observe the others' actions, they would be unable to compute these gradients.

In DGAP, agents construct an approximation of the gradient at the current action profile, by randomly exploring the effects of increasing or decreasing their actions by small increments. The agents use the information collected from this exploration to choose a new action: if the effect revealed is an increase (resp. decrease) in payoff, then players move in the same (resp. opposite) direction, with an amplitude proportional to the approximated gradient.

Although this procedure resembles a gradient learning process, there are two major differences from a standard gradient method. First, the DGAP is a random process instead of a deterministic dynamical system. Second, in the standard gradient method with non-negative actions, players' behaviors are discontinuous at the boundary of the strategy space (see Arrow and Hurwicz (1960)). In order to avoid such discontinuity in players' behavior, we assume that changes in actions are dampened as they approach the boundary. Hence the name of our learning process.

We first prove that this process is well-defined - i.e. players' actions always remain non-negative (Proposition 1). Then we analyze its convergence properties and find that contrary to discrete games, where convergence to Nash of specific learning processes is generally difficult to obtain even for two- or three-player games, convergence is obtained in large classes of games with arbitrary numbers of players. We restrict to strongly single-peaked payoff functions[1], focusing our attention on three classes of games that are of particular interest for economics and have been extensively analyzed in the learning literature: games with strategic complements, a class of games containing all potential games, and all games where the set of Nash equilibria is finite. This last class includes all games with a unique Nash equilibrium, such as strictly concave games and many of the generalized continuous zero-sum games.

The DGAP is a stochastic process, the random part being the direction chosen for the exploration. We analyze its (random) set of accumulation points, called the *limit set*, by resorting to *stochastic approximation theory*. This theory tells us that the long-run behavior of the stochastic process is driven by some underlying *deterministic dynamical system*. We thus start by showing that the deterministic system that underlies our specific stochastic learning process is a dampened gradient system (Proposition 1). We also show that all the Nash equilibria of a game are stationary points - otherwise called zeros - of this dynamical system, although other

---

[1]See Hypothesis 1 for a proper definition.

points may also be stationary. However, we prove (Proposition 2) that non-Nash stationary points are necessarily unstable[2]. This is done in Section 2, where we also detail the DGAP and provide the necessary definitions.

Stochastic approximation theory tells us that the stationary points of the underlying dynamical system are plausible candidates for the limit set of the random process. Yet it does not provide general criteria for excluding some of these candidates so as to obtain more precise predictions. This is actually one of the major difficulties in the field (see for instance Benaïm and Faure (2012)). While the conceptual contribution of this paper lies in providing a natural learning process for games with continuous action sets, our technical contribution lies in providing precise statements on the structure of the limit set of the DGAP. Each result that we get is different, in the sense that it uses a different mathematical tool. It is remarkable that almost all our results hold with probability one, which is in general very difficult to obtain.

In Section 3, we analyze *games with strategic complements*. We show (Theorem 1) that the DGAP cannot converge[3] to an unstable Nash equilibrium. Furthermore, we prove that the process will almost surely converge to a Nash equilibrium which is stable, except in very specific cases involving the structure of interactions between players: non-convergence might occur under a condition called *bipartiteness*.

In Section 4, we analyze a class of games that we call *locally ordinal potential games*. This class contains all the ordinal potential games, which in turn contain all the potential games. We have three results (in Theorems 2 and 3). First, the limit set of the DGAP is always contained in the set of stationary points of the dynamics. When equilibria are isolated, this implies that the process converges to a Nash equilibrium with probability one, since we prove that the process cannot converge to a non-Nash stationary point in these games. Second, we show that under the condition of non-bipartiteness, the DGAP converges to a Nash equilibrium which is stable when equilibria are isolated. Third, although convergence to unstable stationary points (possibly non-Nash) cannot be ruled out in general (i.e. when equilibria are not isolated), we characterize the set of stable stationary points. We prove that they are local maxima of the locally ordinal potential function, that they are neces-

---

[2]Throughout the paper, several notions of stability will be used. They are all recalled in Section 2.

[3]Because the process is stochastic, the notion of convergence we use here is that of almost sure convergence.

sarily connected components of Nash equilibria, and that they are necessarily stable equilibria of another, unrelated dynamical system: the Best-Response dynamics.

Finally, in Section 5, we consider all games for which stationary points are isolated. This class includes the vast majority of games studied in economics. We cannot prove precise and general convergence results, since there is no guarantee that the limit set of the process will be included in the set of stationary points. Still, we state two results. First, DGAP will converge to a stable Nash equilibrium with positive probability in all these games. Second, we exclude convergence to what we call undesirable stationary points, i.e. those that are non-Nash, and unstable Nash equilibria.

Also in Section 5, we focus on games with a unique Nash equilibrium that have previously been analyzed, either by Arrow and Hurwicz (1960) or by Rosen (1965). They examined dynamical systems which are either discontinuous or complex gradient systems, obtaining convergence to the unique Nash equilibrium. We obtain the same results - for our process - with convergence to the unique equilibrium with probability one. Thus, another contribution of this paper is to show that we can use a gradient system which is both simple and continuous, and preserve the convergence properties.

Related Literature

As mentioned earlier, the learning literature has essentially focused on finite action games. Many rules have been proposed and studied, both in the non payoff-based and in the payoff-based contexts. In the former, the most widely-explored adaptive process is fictitious play (introduced in Brown (1951)), where players' average actions are shown to converge[4] for 2-player zero-sum games (Robinson (1951)), for $2 \times N$ games (Miyazawa (1961) and Berger (2005)), for potential games (Monderer and Shapley (1996)). Convergence is also obtained for *stochastic fictitious play*, introduced by Fudenberg and Kreps (1993), in 2×2 games (Benaïm and Hirsch (1999)), zero-sum and potential games (Hofbauer and Sandholm (2002)) and supermodular games (Benaïm and Faure (2012)). However, it has been shown that fictitious play does not always converge to Nash once there are at least 3 actions per player (Shapley (1964)). Other non payoff-based learning rules include hypothesis testing (Foster and Young (2003)) or calibrated forecasting (Kakade and Foster (2008)). Our contribution differs in both dimensions: we focus on continuous games and on payoff-based procedures.

---

[4]In all of the following papers, the convergence notions differ. This has implications in terms of the scope of each result. For details the reader should refer directly to the papers.

Many payoff-based learning rules have been explored in the context of discrete games. Such rules include the popular class of reinforcement learning procedures (see Börgers and Sarin (1997) or Erev and Roth (1998) for pioneer work). These procedures have been studied in very specific finite games: 2×2 games in Posch (1997), 2-player games with positive payoffs in Börgers and Sarin (1997), Beggs (2005), Hopkins and Posch (2005) or Laslier et al. (2001). On the same topic, see also Leslie and Collins (2005), Cominetti et al. (2010), Bravo and Faure (2015) and Bravo (2016). However, it is not known if these procedures converge to Nash in more general games.

Other payoff-based procedures for discrete games have been proposed, including: Regret-testing (Foster and Young (2006)) which converges in any 2-person game; Generalized regret-testing (Germano and Lugosi (2007)) which converges in any generic n-person game; Experimentation dynamics (Marden et al. (2009)) which converge to Nash in the class of n-person weakly acyclic games; Trial and error (Young (2009)) which comes close to Nash equilibrium a large fraction of the time; Aspirational learning (Karandikar et al. (1998)) which may fail to converge even in 2×2 games.

The literature on continuous game is sparser, and a distinction can also be made between procedures which are demanding in terms of sophistication and knowledge of the players, and procedures which are of the payoff-based type. In the first category, Arrow and Hurwicz (1960) prove that when all players' payoff functions are strictly concave, the gradient method converges to the unique Nash equilibrium in generalized zero-sum games. Rosen (1965) studies a gradient method in concave n-person games with a unique equilibrium, and shows that this unique equilibrium is globally asymptotically stable for some weighted gradient system, with suitably chosen weights. In a recent paper, Mertikopoulos (2018) studies a gradient-like stochastic learning algorithm where agents receive erroneous information about their gradients in the context of concave games, and shows that whenever this process converges, it does so to a Nash equilibrium. Using a different approach, Perkins and Leslie (2014) adapt stochastic fictitious play to games with continuous action sets, and show that it converges in 2-player zero-sum games. Our contribution differs from these in analyzing a payoff-based learning process.

The two papers most closely related to ours are those analyzing payoff-based procedures designed for continuous games. Dindos and Mezzetti (2006) consider a stochastic adjustment process called the better-reply. At each step, agents are sequentially picked to play a strategy chosen at random, while the other players do not

6

move. The agent then observes the hypothetical payoff that this action would yield, and decides whether to stick to this new strategy or to go back to the previous one. This process converges to Nash when actions are either substitutes or complements around the equilibrium in games called aggregative games, with quasi-concave utility functions. However, their contribution differs from ours in several respects, the most important being that agents revise their strategies sequentially. The driving force for convergence is that with positive probability, every player will be randomly drawn as many times as necessary to approximate a best response. In our paper, we assume that all players move simultaneously. In that case, it is easy to construct a simple game where simultaneity drives the better-reply adjustment process to cycle.

The second related paper, by Huck et al. (2004), considers another type of payoff-based learning process, called Trial and error - but which has no link with the Young (2009) procedure - in the context of the Cournot oligopoly, where players move simultaneously, as in our paper. Players choose a direction of change and stick to this direction as long as their payoff increases, changing as soon as it decreases. The authors show that the process converges, but it does so to the joint-profit maximizing profile and not to the (unique) Nash equilibrium of the game. While our paper is similar to theirs in spirit, unlike them, we do not focus on a specific game[5], we allow for multiple equilibria, including continua of equilibria, and we get convergence to Nash. Our process differs in two respects. First, we consider an exploration stage where the players decide in which direction they will be moving. This exploration stage seems to make our agents less naive. Second, the amplitude of the moves is constant in their paper while in ours, amplitude depends on the variation in payoffs after an exploration stage. This might explain why we obtain convergence to Nash while they do not.

A group of papers coming from a different literature deserves mention here: in the literature on evolutionary dynamics in population games, many dynamical systems have been suggested in the context of infinite populations choosing strategies among a finite set. Recently, several papers have extended these dynamical systems to continuous strategy spaces. The continuous strategy version of the replicator dynamic has been studied by Bomze (1990), Oechssler and Riedel (2001), Oechssler and Riedel (2002), and Cressman (2005), while the Brown-von Neumann-Nash (BNN) dynamic

---

[5]Although, as the authors suggest, the intuition for their result might carry through to other games.

has been extended by Hofbauer et al. (2009). Lahkar and Riedel (2015) extend the logit dynamics, Cheung (2014) adapts the pairwise comparison dynamics and Cheung (2016) works with the class of imitative dynamics (which includes the replicator dynamics). Although this group of papers deals with dynamical systems for continuous games, their contexts are totally different (continuum of players). Moreover, their main goal is to define the extensions of existing dynamics, and to see whether they are well-defined and share the same properties as their discrete-strategy counterparts.

2 The model

2.1 Definitions and hypothesis

Let $\mathcal{N} = \{1, \ldots, N\}$ be a set of players, each of whom repeatedly chooses an action from $X_i = [0, +\infty[$. An action $x_i \in X_i$ can be thought of as an effort level chosen by individuals, a price set by a firm, a monetary contribution to a public good, etc. Let $X = \times_{i=1,\ldots,N} X_i$. We denote by ∂X the boundary of X, i.e. $\partial X := \{x \in X; x_i = 0 \text{ for some } i \in \mathcal{N}\}$. And we let $\mathrm{Int}(X) := X \setminus \partial X$ denote the interior of X.

At each period of time, players observe a payoff that is generated by an underlying repeated game $\mathcal{G} = (\mathcal{N}, X, u)$, where $u = (u_i)_{i=1,\ldots,N}$ is the vector of payoff functions. Players know nothing about the payoff functions, nor about the set of opponents. In this paper we will examine several classes of underlying games, each class being defined by different properties on the functions u_i. However, we will always make the two following standing assumptions:

Hypothesis 1 *For any i, the payoff map u_i is assumed to be $\mathcal{C}^1$ on $\mathbb{R}_+^N$ and with the property that, for any $x_{-i} \in X_{-i}$, there exists $M(x_{-i}) \in X_i$ such that the map $x_i \mapsto \frac{\partial u_i}{\partial x_i}(x_i, x_{-i})$ is strictly positive for $x_i < M(x_{-i})$ and strictly negative for $x_i > M(x_{-i})$.*

Hypothesis 1 implies that best responses are unique and $\mathrm{BR}_i(x_{-i}) = M(x_{-i})$. This assumption is verified for instance if $x_i \mapsto u_i(x_i, x_{-i})$ is strictly concave, $\frac{\partial u_i}{\partial x_i}(0, x_{-i}) > 0$ and $\lim_{x_i \to +\infty} \frac{\partial u_i}{\partial x_i}(x_i, x_{-i}) < 0$.

In the games we consider, interactions between players can be very general. They can be heterogeneous across players and they can be of any sign. However we assume that externalities are symmetric in sign:

Hypothesis 2 *Games are assumed to have symmetric externalities, i.e. $\forall i \neq j$ and $\forall x$,*

$$\mathrm{sgn}\left(\frac{\partial u_i}{\partial x_j}(x)\right) = \mathrm{sgn}\left(\frac{\partial u_j}{\partial x_i}(x)\right)$$

where $\mathrm{sgn}(a) = 0$ if $a = 0$.

Most of the continuous games in the economics literature fall into this class. Note that a game with symmetric externalities does not require them to be of equal intensity. Also, symmetric externalities allow for patterns where i exerts a positive externality on individual j and a negative externality on individual k. Note finally that symmetric externalities do not imply that $\mathrm{sgn}\left(\frac{\partial u_i}{\partial x_j}(x)\right) = \mathrm{sgn}\left(\frac{\partial u_i}{\partial x_j}(x')\right)$ for $x \neq x'$.

Some of our results will depend on the pattern of interactions in the game $\mathcal{G}$. We capture this pattern by an *interaction graph*, defined as follows. Let $x = (x_1, \ldots, x_N)$ be an action profile. The interaction graph at profile x is given by the matrix $\mathbf{G}(x)$ where $g_{ii}(x) = 0$ and, for $i \neq j$, $g_{ij}(x) = 1$ if $\frac{\partial u_i}{\partial x_j}(x) \neq 0$ and $g_{ij}(x) = 0$ otherwise. Note that the interaction graph is local, in the sense that it depends on the vector of actions. Thus $\mathbf{G}(x)$ can either be constant on X or change as x changes. Note also that the interaction graph of a game satisfying Hypothesis 2 is symmetric.

We now provide two examples of games satisfying Hypothesis 1 and 2 and describe the interaction graphs.

Example 1 (Public good game) *Players contribute an effort $x \in [0; +\infty[$ to a public good. The payoff of player i is $u_i(x) = b_i(x_i + \sum_{j \neq i} \delta_{ij} x_j) - c_i.x_i$, where $c_i > 0$ is the marginal cost of effort for i, $1 \geq \delta_{ij} \geq 0$ is a measure of subtitutability between i and j's efforts, and $b_i(.)$ is a differentiable, strictly increasing concave function. In some contexts, not all players will benefit from one player's contribution (see Bramoullé and Kranton (2007)). This is why we leave the possibility that $\delta_{ij} = 0$ for some pairs (i, j).*

This game satisfies Hypothesis 1 and 2. Further, the interaction graph is constant since it does not depend on the action profile x: $g_{ij}(x) = 0$ if $\delta_{ij} = 0$, while $g_{ij}(x) = 1$ if $\delta_{ij} > 0$.

Example 2 (Aggregate demand externalities) *In macro-economics, aggregate-demand-externality models (see for instance Fudenberg and Tirole (1991)) are games*

satisfying Hypothesis 1 and 2. For instance, search models à la Diamond enter this class, where players exert a search effort x_i, with payoff functions $u_i(x) = \alpha x_i \sum_{j \neq i} x_j - c(x_i)$, and where $c(x_i)$ is the cost of searching, $x_i x_j$ is the probability that i and j end up partners and α is the gain when a partner is found. The interaction graph of such a game depends on the profile x, since $\frac{\partial u_i}{\partial x_j}(x) = \alpha x_i$, thus $g_{ij}(x) = 0$ if $x_i = 0$, while $g_{ij}(x) = 1$ if $x_i \neq 0$.

This game can be generalized to local search models with payoffs $u_i(x) = \alpha x_i \sum_{j \neq i} \delta_{ij} x_j - c(x_i)$, where, as above, $\delta_{ij} \geq 0$ and where $\delta_{ij} = 0$ means that individual j exerts no externality on individual i. A popular game analyzed in the network literature in a different context is the game introduced in Ballester et al. (2006), which is actually a local aggregate-demand-externality model, where $c(x_i) = \frac{1}{2}x_i^2 - x_i$, and $\delta_{ij} \in \{0, 1\}$. In that case, the interaction graph only partly depends on x, since $\delta_{ij} = 0$ implies $g_{ij}(x) = 0$ for all x.

We denote by NE the set of Nash equilibria. In many economics applications, Nash equilibria would consist of isolated points. Examples 1 and 2 provide examples of games where the set of Nash equilibria is generically finite. However, in what follows we will sometimes deal with a continuum of equilibria. This is the case for instance in Example 1, when $\delta_{ij} \in \{0, 1\}$ (see Bervoets and Faure (2018)). Because we wish to be as general as possible, we will consider *connected components* of NE:

Definition 1 *Let Λ be a compact connected subset of NE and let $N^\delta(\Lambda) := \{y \in X : d(y, \Lambda) < \delta\}$. We say that Λ is a* connected component *of NE if there exists $\delta > 0$ such that $N^\delta(\Lambda) \cap NE = \Lambda$.*

2.2 The Learning Process

We consider a payoff-based learning process in which agents construct a partial approximation of the gradient by exploring the effects of deviating in one direction that they chose at random at every period. This information allows agents to choose a new action depending on what they just learned from the exploration stage. Here we detail what agent i does, bearing in mind that every agent simultaneously uses the same rule.

• At the beginning of round n, agent i is playing action $x_n^i := e_{2n}^i$ and is enjoying the associated payoff $u_i(e_{2n}^i, e_{2n}^{-i})$. Player i then selects his actions e_{2n+1}^i and $e_{2n+2}^i(= x_{n+1}^i)$ as follows.

• *Exploration stage* - Player i plays a new action e_{2n+1}^i, chosen at random around his current action e_{2n}^i. Formally, let $(\epsilon_n^i)_n$ be a sequence of i.i.d random variables such that $\mathbb{P}(\epsilon_n^i = 1) = \mathbb{P}(\epsilon_n^i = -1) = 1/2$. At period n, ϵ_n^i is drawn and player i plays $e_{2n+1}^i := e_{2n}^i + \frac{1}{n+1}\epsilon_n^i$.

• *Updating stage* - Player i observes his new payoff, and computes $\Delta u_{n+1}^i := u_i(e_{2n+1}^i, e_{2n+1}^{-i}) - u_i(e_{2n}^i, e_{2n}^{-i})$. This quantity provides i with an approximation of his payoff function's gradient. Using this information, player i updates his action by playing $e_{2n+2}^i := e_{2n}^i + \epsilon_n^i \Delta u_{n+1}^i e_{2n}^i$. Thus, when Δu_{n+1}^i is positive, player i follows the direction that he just explored, while he goes in the opposite direction when Δu_{n+1}^i is negative.

• Period n ends. We set $x_{n+1}^i := e_{2n+2}^i$ and agent i gets the payoff $u_i(e_{2n+2}^i, e_{2n+2}^{-i})$. Round $n + 1$ starts.

Let $x_n = e_{2n}$ and $\mathcal{F}_n$ be the history generated by $\{e_1, ..., e_{2n+1}\}$. Studying the asymptotic behavior of the random sequence $(e_n)_n$ amounts to studying the sequence $(x_n)_n$. Hence the focus of this paper is on the convergence of the random process $(x_n)_n$.

The next proposition shows that the process is well-defined, in the sense that it always remains within the admissible region (i.e. the actions stay positive). It also proves that the DGAP is a discrete time stochastic approximation process.

Proposition 1 *The iterative process is such that $x_n^i > 0$ for all i.*
It can be written as

$$x_{n+1} = x_n + \frac{1}{n+1}\left(F(x_n) + U_{n+1} + \xi_{n+1}\right),\tag{1}$$

where

(i) $F(x) = (F_i(x))_i$ *with* $F_i(x) = x_i \frac{\partial u_i}{\partial x_i}(x_i, x_{-i})$,

(ii) U_{n+1} *is a bounded martingale difference (i.e* $\mathbb{E}(U_{n+1} \mid \mathcal{F}_n) = 0$),

(iii) $\xi_n = \mathcal{O}(1/n)$.

All our proofs are in the appendix.

The iterative process (1) is a discrete time stochastic process with step $\frac{1}{n+1}$.[6] If there were no stochastic term, the process (1) would write

$$x_{n+1} = x_n + \frac{1}{n+1} F(x_n),$$

which corresponds to the well-known Euler method, a numerical procedure for approximating the solutions of the deterministic ordinary differential equation (ODE)

$$\dot{x} = F(x). \tag{2}$$

Although the (stochastic) process (1) differs from the (deterministic) process (2) because of the random noise, the asymptotic behavior of (2) will inform us on the asymptotic behavior of (1).[7]

Remark 1 *In the standard gradient method (see for instance Arrow and Hurwicz (1960)), the dynamical system is defined as $\dot{x} = H(x)$ where*

$$H_i(x) = \begin{cases} \frac{\partial u_i}{\partial x_i}(x) & \text{unless } x_i = 0 \text{ and } \frac{\partial u_i}{\partial x_i}(x) < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

The function H is thus discontinuous, since the process could otherwise leave the admissible space. Conversely, the dynamical system underlying the DGAP is continuous: $F_i(x) = x_i \frac{\partial u_i}{\partial x_i}(x)$. The role played by the multiplicative factor x_i is to dampen the variations of the state variable and ensure that it will never reach the boundary - although it can converge to it. This is not unreasonable behavior: the gradient system assumes that players crash onto the boundary, whereas we assume that the closer they get to the boundary, the smaller their movements become.

---

[6]Note that $\sum_n \frac{1}{n+1} = \infty$ and $\lim_{n \to \infty} \frac{1}{n+1} = 0$. It is important that the sum diverges, to guarantee that the process does not get "stuck" anywhere, unless agents want to stay where they are. Further, it is important that the terms go to zero, so that the process can "settle" when agents want to. In fact, the term $\frac{1}{n+1}$ can be replaced by any step of the form α_{n+1}, where $\sum_n \alpha_{n+1} = \infty$ and $\lim_{n \to \infty} \alpha_{n+1} = 0$, without affecting the results.

[7]Stochastic approximation theory (see Benaïm (1996) or Benaïm (1999) for instance) tells us that, as periods unfold, the random process gets arbitrarily close to the solution curve of its underlying dynamical system. In other words, given a time horizon $T > 0$ - however large it might be - the process shadows the trajectory of some solution curve between times t and $t + T$ with arbitrary accuracy, provided t is large enough.

In Rosen (1965), the author studies another gradient method also ensuring that the system never leaves the state space. The system is given by

$$H_i(x) = r_i \frac{\partial u_i}{\partial x_i}(x) + \sum_{j=1}^{k} \lambda_j \frac{\partial h_j}{\partial x_i}(x), \tag{4}$$

where the h_j functions are the constraints defining the convex and compact set R where x lives (i.e. $R = \{x; h(x) \geq 0\}$), r_i and λ_j are appropriately chosen weights guaranteeing that the system will always remain within the set R.

2.3 Limit sets

The focus of this paper is on the asymptotic behavior of the random process $(x_n)_n$. Hence we are interested in its *limit set*[8].

Definition 2 (Limit set of $(x_n)_n$) *Given a realization of the random process, we denote the limit set of $(x_n)_n$ by*

$$\mathcal{L}\left((x_n)_n\right) := \{x \in X; \ \exists \ a \ subsequence \ x_{n_k} \ such \ that; \ x_{n_k} \to x\}. \tag{5}$$

Note that the limit set of the learning process is a random object, because the asymptotic behavior of the sequence $(x_n)_n$ depends on the realization of the random sequence $(\epsilon_n)_n$, drawn at every exploration stage.

Proposition 1 allows us to make use of stochastic approximation theory, which provides a characterization of the candidates for $\mathcal{L}\left((x_n)_n\right)$[9]. In particular the ω-limit sets of $\dot{x} = F(x)$[10] lie among these candidates. However, several difficulties remain:

---

[8]In the remainder of the paper, we will always place ourselves on the event $\{\limsup_n \|x_n\| < +\infty\}$, i.e. we will abstract from the possible realizations which take the process to infinity.

[9]In Benaïm (1999), it is established that on $\{\limsup_n \|x_n\| < +\infty\}$, the limit set of (x_n) is always compact, invariant and attractor-free. This class of sets is called *internally chain transitive* (ICT). These sets can take very complicated forms, but they conveniently include the zeroes of F and the ω-limit set of any point x (if non-empty).

[10]Let $\varphi(x, t)$ denote the flow of $F(.)$, i.e. the position of the solution of (2) with initial condition x, at time t. Then, the ω-limit set of x is given by $\omega(x) := \{z \in X; \ lim_{k \to \infty} \varphi(x, t_k) = z \text{ for some } t_k \to \infty\}$. Notice that by the regularity assumption on $u(.)$, F satisfies the Cauchy-Lipschitz condition that guarantees that, for all $x \in X$, φ is well-defined and unique. We consider the restriction of φ on $X(= \mathbb{R}_+^N)$, since $X = \mathbb{R}_+^N$ is invariant for its flow, and our random process (1) always remains in the positive orthant.

first, there might be other candidates that are not ω-limit sets of the underlying ODE. Moreover, this theory does not provide general criteria to systematically exclude any of these candidates, nor to confirm that they are indeed equal to $\mathcal{L}\left((x_n)_n\right)$.

The stationary points of the dynamical system (2) are particular ω-limit sets that will be of interest to us, as they contain all the Nash equilibria of the underlying game. The set of stationary points, denoted $Z(F)$, will be called the *zeros of F*: $Z(F) := \{x \in X; F(x) = 0\}$. For convenience, we drop the reference to F and simply write Z.

Observe that $F_i(x) = x_i \frac{\partial u_i}{\partial x_i}(x)$. Thus, $x \in Z \iff \forall i \in \mathcal{N}, \left(x_i = 0 \text{ or } \frac{\partial u_i}{\partial x_i}(x_i, x_{-i}) = 0\right)$, while $x \in NE \iff \forall i \in \mathcal{N}, \left(\frac{\partial u_i}{\partial x_i}(x_i, x_{-i}) = 0, \text{ or } x_i = 0 \text{ and } \frac{\partial u_i}{\partial x_i}(x_i, x_{-i}) \leq 0\right)$. This implies that all the Nash equilibria of the game are *included* in the set of zeros of F. Unfortunately, Z contains more than the set of Nash equilibria. We call $x \in Z \setminus NE$ an *other zero (OZ)* of the dynamical system: $OZ = \left\{x : F(x) = 0 \text{ and } \exists i \text{ s.t. } x_i = 0, \frac{\partial u_i}{\partial x_i}(x) > 0\right\}$. We have the following partition of F:

$$Z = NE \cup OZ. \tag{6}$$

Note that ∂X might contain some points in NE, however $OZ \subset \partial X$.

Convergence or non-convergence of our random process to a given point or set will sometimes depend on the *stability* of the latter with respect to the deterministic dynamical system $\dot{x} = F(x)$. In different sections we use various notions of stability, which we recall here.

Let $\hat{x} \in Z$. The point $\hat{x}$ is *asymptotically stable* (denoted by $\hat{x} \in Z^{AS}$) if it uniformly attracts an open neighborhood W of itself: $\lim_{t \to +\infty} \sup_{x \in W} \|\varphi(x, t) - \hat{x}\| = 0$, where $\varphi(x, t)$ denotes the flow of $F(\cdot)$. The point $\hat{x}$ is *linearly stable* (denoted by $\hat{x} \in Z^{LS}$) if for any $\lambda \in \mathrm{Sp}(DF(\hat{x}))$ - where $DF(\hat{x})$ is the Jacobian matrix of F evaluated at $\hat{x}$ and $\mathrm{Sp}(M)$ is the spectrum of matrix M - we have $\mathrm{Re}(\lambda) < 0$ and $\hat{x}$ is *linearly unstable* (denoted by $\hat{x} \in Z^{LU}$) if there exists $\lambda \in \mathrm{Sp}(DF(\hat{x}))$ such that $\mathrm{Re}(\lambda) > 0$. Note that if $\hat{x}$ is hyperbolic (that is $\mathrm{Re}(\lambda) \neq 0$ for any $\lambda \in \mathrm{Sp}(DF(\hat{x}))$) then it is either linearly stable or linearly unstable. We denote the set $Z \setminus Z^{LU}$ by Z^S and by a slight abuse of language, we will call all points in Z^S *stable*.

We have the following inclusions:

$$Z^{LS} \subset Z^{AS} \subset Z^S.$$

Proposition 2 *We have $OZ \subset Z^{LU}$. As a consequence, $Z^S \subset NE$.*

To prove this, we take x in OZ and pick an individual such that $x_i = 0$ and $\frac{\partial u_i}{\partial x_i} > 0$. We then show that $\frac{\partial u_i}{\partial x_i}$ is an eigenvalue of $DF(x)$. The direct consequence of Proposition 2 is that if the limit set $\mathcal{L}((x_n)_n)$ contains stable stationary points, they must be stable Nash equilibria.

In view of Proposition 2, we will use the following notations in the remainder: $NE^{LS} := Z^{LS}$, $NE^{AS} := Z^{AS}$ and $NE^S := Z^S$.

As mentioned earlier, we will sometimes be dealing with connected components of NE instead of isolated points. We will thus use the concept of *attractor* (see Ruelle (1981)):

Definition 3 *Let $S \subset X$ be invariant for the flow φ. Then a set $A \subset S$ is an attractor for $\dot{x} = F(x)$ if*
(i) A is compact and invariant;
(ii) there exists an open neighborhood U of A with the following property:

$$\forall \epsilon > 0, \ \exists T > 0 \ such \ that \ \ \forall x \in U, \ \forall t \geq T, \ d(\varphi(x,t), A) < \epsilon.$$

An attractor for a dynamical system is a set with strong properties: it uniformly attracts a neighborhood of itself.

Remark 2 *Let $\hat{x} \in Z$ be an isolated stationary point of $\dot{x} = F(x)$. Then $\hat{x}$ is asymptotically stable if and only if $\{\hat{x}\}$ is an attractor for $\dot{x} = F(x)$.*

We turn to the analysis of several classes of games.

3   Strategic complements

Definition 4 *A game $\mathcal{G}$ is a game with strategic complements if payoff functions are such that $\frac{\partial^2 u_i}{\partial x_i \partial x_j}(x) \geq 0$ for all $i \neq j$.*

Games with strategic complements have nice structured sets of Nash equilibria (Vives (1990)), and offer nice convergence properties for specific dynamical systems. However, it can be difficult to obtain convergence to Nash for general learning procedures. There are several reasons for this that we illustrate here through two examples.

First, consider the Best-Response dynamics. Under Hypothesis 1, best-response functions are differentiable and strictly increasing. In that case, Vives (1990) proves in Theorem 5.1 and Remark 5.2 that, except for a specific set of initial conditions, the Best-Response dynamics, whether in discrete or in continuous time, monotonically converges to an equilibrium point. Unfortunately, in our case this set of problematic initial conditions cannot be excluded, in particular because the process is stochastic. It could be that the stochastic process often passes through these points, in which case it is known to possibly converge to very complicated sets[11]. In order to study convergence of the DGAP, we thus need to consider all possible trajectories and cannot rely on existing results.

Second, consider the standard reinforcement learning stochastic process, whose mean dynamics are the replicator dynamics. As shown in Posch (1997), the process can converge with positive probability to stationary points that are not only unstable, but also non-Nash. Examples can be constructed with 2 players, each having 2 strategies, supermodular payoff matrices with a unique strict Nash equilibrium, which is, moreover, found by elimination of dominated strategies. Yet even then, the learning process converges with positive probability to any other combination of strategies. This happens because there are some stationary points of the dynamics where the noise generated by the random process is null.

These two examples illustrate how, despite the games' appealing properties, convergence to Nash is neither guaranteed nor easy to show when it occurs. We show that the DGAP will converge. In order to get our result, we first need to prove that no point in ∂X, the boundary of the state space, will be included in the limit set of the process. We start by imposing a simple and natural hypothesis.

Hypothesis 3 *For any agent i,*

$$\frac{\partial u_i}{\partial x_i}(0,0) > 0.$$

Hypothesis 3 guarantees that players want to move away from the origin. Because of strategic complementarities, this also implies that players want to move away from any point of ∂X (since $\frac{\partial u_i}{\partial x_i}(0, x_{-i}) > \frac{\partial u_i}{\partial x_i}(0,0)$). However, despite the fact that all players prefer to move away from the boundary, it is not clear why the stochastic process should remain at a distance from this boundary. The difficulty

---

[11]See Hirsch (1999).

comes from the following fact: assume players start close to the boundary. Then, at the exploration stage, some decrease their efforts while others increase theirs. Although complementarities imply that the players who decreased their efforts would have been better-off if they had instead increased them, they could still end up with a better payoff than before the exploration, and thus continue decreasing at the updating stage, getting closer to ∂X.

The following proposition proves that this will not happen in the long run.

Proposition 3 *Under Hypothesis 3, there exists $a > 0$ such that $\mathcal{L}((x_n)_n) \subset [a, +\infty[^N$ almost surely.*

From the mathematical point of view, the major problem to obtain Proposition 3 is to show that a stochastic approximation algorithm like the one given by (1) is pushed away from an invariant set for F where the noise term vanishes. In fact, there is almost no general result along these lines in the literature.

The proof of Proposition 3 is long and technical, but the idea goes as follows: among the players close to the boundary, the player exerting the least effort will increase his effort on average. Unfortunately, this does not imply that the smallest effort also increases, since another player may have decreased his. We thus construct a stochastic process which is a suitable approximation of the smallest effort over time. We then show that this new process cannot get close to the boundary, and because it is close asymptotically to our process, we are able to conclude.

Definition 5 *The interaction graph $\mathbf{G}(x)$ is said to be bipartite at $x \in X$ if the set N of players can be partitioned into N_1 and N_2 such that for any pair of players i and j we have*

$$g_{ij}(x) = 1 \implies (i \in N_1 \text{ and } j \in N_2) \text{ or } (i \in N_2 \text{ and } j \in N_1).$$

An interaction graph is non-bipartite on a set A if for all $x \in A$, $\mathbf{G}(x)$ is not bipartite.

We are now ready to state the main result of this section.

Theorem 1 *Consider a game of strategic complements and smooth payoff functions, and assume that Hypothesis 3 holds. Then*

(i) The learning process (x_n) cannot converge to an unstable Nash equilibrium:

$$\forall \, \tilde{x} \in NE^{LU}, \; \mathbb{P}\left(\lim_n x_n = \tilde{x}\right) = 0.$$

(ii) If, in addition, the interaction graph is non-bipartite on $\text{Int}(X)$, the learning process (x_n) almost surely converges to a stable Nash equilibrium:

$$\mathbb{P}\left(\exists \, x^* \in NE^S : \; \lim_n x_n = x^*\right) = 1.$$

This result is very tight. Because the hypotheses of the theorem are verified for most common economic models we can think of, this theorem guarantees that the learning process will not only converge to Nash in most cases, it will additionally converge to a stable equilibrium. In cases where the interaction graph is bipartite, we cannot guarantee that the process will not converge to general unstable sets[12]. However, we can still exclude convergence to linearly unstable equilibria by point (i).

Let us provide some insights on the bipartiteness condition. As in Posch (1997), one potential issue is that the random process could get stuck around stationary points of the underlying dynamics if the random noise is zero at these stationary points. More precisely, a stationary point is unstable if there is some direction along which the system "escapes" the stationary point. But the system has to be able to follow that direction, otherwise it will get stuck. The random process plays precisely this role here: it allows the system to escape, as long as the unstable direction component of the random noise $(U_n)_n$ is not zero at that point. At an unstable equilibrium, we can show that the noise is not zero in the unstable direction and this guarantees the non-convergence result of part (i). The non-bipartiteness of the network guarantees that the noise $(U_n)_n$ has the property of being *uniformly exciting* everywhere in $\text{Int}(X)$, which guarantees that the process can escape in any direction. This yields part (ii). When the network is bipartite, this property does not hold and we cannot guarantee that the process will not get stuck in an unstable set.

Note that the bipartiteness condition does not imply that the process will not converge to an element of Z^S. However, we provide two examples in the appendix (Examples E.1 and E.2) in which we show that the noise can vanish on bipartite networks in games that have either no strategic complements or no symmetric externalities. In our examples the noise vanishes at unstable equilibria.

---

[12]Linearly unstable equilibria are unstable sets, but unstable sets also include much more complex structures.

4 Locally ordinal potential games

We introduce a class of games that we call the *locally ordinal potential games*. Recall that a game $\mathcal{G}$ is a *potential game* (*PG*) if there is a function $P : X \to \mathbb{R}$ such that for all $x_{-i} \in X_{-i}$, for all $x_i, x_i' \in X_i$, we have $u_i(x_i, x_{-i}) - u_i(x_i', x_{-i}) = P(x_i, x_{-i}) - P(x_i', x_{-i})$, and an *ordinal potential game* (*OPG*) if $u_i(x_i, x_{-i}) - u_i(x_i', x_{-i}) > 0 \iff P(x_i, x_{-i}) - P(x_i', x_{-i}) > 0$.

Definition 6 *A game $\mathcal{G}$ is a locally ordinal potential game (LOPG) if there is a differentiable function $P : X \to \mathbb{R}$ such that*

$$\mathrm{sgn}\left(\frac{\partial u_i}{\partial x_i}(x) \right) = \mathrm{sgn}\left(\frac{\partial P}{\partial x_i}(x) \right)$$

The class of *LOPG* is large, in the sense that $PG \subset OPG \subset LOPG$ when P is differentiable. It also contains many games of economic interest. For instance, both examples 1 and 2 are locally ordinal potential games.

The generality of our results depends on the structure of the set of stationary points of the game under consideration, and in particular on whether it consists of isolated points or not. For instance, the public good game of example 1 generically has a finite number of isolated zeros, but can have continua of equilibria for certain values of the substitutability parameter[13].

Theorem 2 *Let $\mathcal{G}$ be an LOPG and P be sufficiently regular. Then*

(i) $\mathbb{P}\left(\mathcal{L}(x_n)_n \subset Z \right) = 1.$

(ii) If $\mathcal{G}$ has isolated zeros, then

$$\mathbb{P}\left(\exists\, x^* \in NE : \ \lim_n x_n = x^* \right) = 1.$$

If, in addition, the interaction graph is non-bipartite on NE, then

$$\mathbb{P}\left(\exists\, x^* \in NE^S : \ \lim_n x_n = x^* \right) = 1.$$

For any *LOPG*, the only set to which the stochastic learning process can converge is the set of zeros of F. Complex ω-limit sets of the dynamical system, which are

---

[13]See Bervoets and Faure (2018) for more details.

non-zeros, can be discarded. We cannot, however, be sure that the process will not reach a set containing other zeros, thus we cannot guarantee convergence to the set of Nash equilibria. When zeros are isolated, however, convergence to Nash is proved by the conjunction of the first point and the fact that the process cannot converge to an isolated other zero. Furthermore, we prove that the DGAP cannot converge to a linearly unstable Nash $\hat{x}$ if $\mathbf{G}(\hat{x})$ is non-bipartite (on this, we provide more details in Section 5).

When zeros are non-isolated, we cannot guarantee that the DGAP will converge to a stable set. We can use Benaïm (1999) to show that $\mathbb{P}\left(\mathcal{L}(x_n)_n \subset A\right) > 0$ on the event $\{x_0 \in \mathcal{B}(A)\}$, for any attractor A of the ODE (2), where $\mathcal{B}(A)$ is the basin of attraction of A. Combining this observation with point (i) of Theorem 2, we get the following important implication: if a connected set Λ is an attractor for $\dot{x} = F(x)$, then Λ is a connected component of Z.

However, when focusing on $LOPG$s, more can be said, since we are able to relate attractors of the dynamics to the potential function P, and to another dynamical system, extensively used in economics: Best-Response Dynamics (BRD).

Definition 7 *Let* BR $: X \to X, x \mapsto \text{BR}(x) := (\text{BR}_1(x_{-1}), ..., \text{BR}_n(x_{-n}))$. *The* continuous-time Best-Response dynamics *(thereafter, BRD) is defined as:*

$$\dot{x} = -x + \text{BR}(x) \tag{7}$$

Definition 8 *Let P be a smooth map and Λ be a connected component of Z, we say that Λ is a local maximum of P if*

 (i) *P is constant on Λ: $P(x) = v$, $\forall x \in \Lambda$;*

 (ii) *there exists an open neighborhood U of Λ such that $P(y) \leq v$ $\forall y \in U$*

We then have

Theorem 3 *Assume $\mathcal{G}$ is an $LOPG$ and let Λ be a connected set. Then the following statements are equivalent*

 (i) *Λ is an attractor for $\dot{x} = F(x)$*

 (ii) *Λ is a local maximum of P*

(iii) $\Lambda \subset NE$ and Λ is an attractor for the best-response dynamics $\dot{x} = -x + \mathrm{BR}(x)$.

This result is positive and informative. First, it tells us that attractors are necessarily included in the set of Nash equilibria. Thus, although the process might converge to other zeros when stationary points are non-isolated, these points are unstable.

Second, Theorem 3 provides two methods of finding the attractors: one way is to look for local maxima of the potential function, which is very convenient when the function P is known; and the other is to look for attractors for another dynamics, possibly simpler to analyze, the BRD. Note that this second method establishes a relation between two dynamics that are conceptually unrelated. Indeed, the BRD assumes that agents are very sophisticated, as they know their exact payoff function, they observe their opponents' play and perform potentially complex computations. Solution curves may be very different, but surprisingly, both dynamics share the same set of attractors.

5   Isolated zeros

In the two previous sections we did not assume any specific structure on the set of zeros of the dynamical system. However, in most economics games with continuous action spaces, the set of zeros, and in particular the set of Nash equilibria, would be finite. In that case, zeros are isolated points. For instance, in the public good game of example 1, Bramoullé et al. (2014) show that the game has a finite number of equilibria for almost every value of substitutability between efforts. The same can be said about the games in example 2. In this section, we restrict our attention to these games.

We start with a useful remark:

Remark 3 If $\hat{x} \in NE^{AS}$, then

$$\mathbb{P}\left(\lim_n x_n = \hat{x}\right) > 0$$

on the event $\{x_0 \in \mathcal{B}(\hat{x})\}$.

This is just a consequence of the result in Benaïm (1999) mentioned earlier and the fact that $\hat{x} \in NE^{AS}$ is an attractor. It says that the process can converge to

desirable outcomes. We next turn to the hard part, i.e. excluding the convergence to undesirable zeros in every game with isolated zeros.

5.1 Non convergence to undesirable zeros

In games with continuum of equilibria, we cannot exclude the possibility of our learning process getting arbitrarily close to elements of the set of other zeros. More precisely, there is no a priori reason to believe that the learning process will converge (to a point) when zeros of the dynamical system are connected components. If it does not, then the process could come arbitrarily close to a continuum of NE that is connected to a continuum of OZ, and oscillate between the two. However, when zeros are isolated this cannot happen and we can discard convergence to other zeros. Further, we can almost always discard convergence to linearly unstable Nash equilibria.

Theorem 4 *Let $\mathcal{G}$ be a game with isolated zeros and assume that $\hat{x} \in Z$. Then:*

i) *If $\hat{x} \in OZ$, then $\mathbb{P}\left(\lim_n x_n = \hat{x}\right) = 0$.*

ii) *If $\hat{x} \in \text{Int}(X)$, $\hat{x} \in NE^{LU}$, and $\mathbf{G}(\hat{x})$ is non-bipartite, then $\mathbb{P}\left(\lim_n x_n = \hat{x}\right) = 0$*

The proof of point a) is a probabilistic proof. We show that in OZ, the players who are playing 0 although they have a strictly positive gradient will, in expectation, increase their action level as they approach the boundary. This is of course a contradiction.

5.2 Concave games

As mentioned in Remark 1, Arrow and Hurwicz (1960) and Rosen (1965) analyzed similar dynamical systems in concave games. The first investigates a subclass of all games with payoff functions that are concave in players' own actions and convex in other players' actions. These games include the well-known class of zero-sum games. The authors then prove global convergence of system (3).

Rosen (1965) deals with concave games, and provides sufficient conditions for the game to have a unique Nash equilibrium when the strategy space is compact and convex: if there are some positive weights such that the weighted sum of the payoff

functions is diagonally strictly concave, then the equilibrium of the game is unique. Under that assumption, the author proves that the weighted gradient system (4) globally converges to this unique equilibrium.

We are interested in determining whether the DGAP also converges in these games, but this raises several problems. First, we need to show that our deterministic system (2) has the same good convergence properties as (3) and (4). But this is not enough, since our process is stochastic, unlike theirs. Second therefore, we need to show that the limit set of the stochastic process (1) is included in the set of stationary points of the dynamical system (2) for these games. Last, the games considered in Arrow and Hurwicz (1960) sometimes have continua of equilibria. For instance, in zero-sum games, the set of equilibria is known to be convex. To avoid this issue, we maintain the concavity condition on the payoff functions but we require that at least one player's payoff function is strictly concave in own action. Under this assumption, we show that these games satisfy Rosen's (1965) condition - and thus have a unique Nash equilibrium. We next show that all games satisfying Rosen's condition have isolated zeros for our dynamical system. With this in hand, we prove that the DGAP converges to the unique Nash equilibrium with probability 1.

Suppose that u_i is concave in x_i for every i. Following Rosen (1965), given $r \in (\mathbb{R}_+^*)^N$ and $x \in X$, let $g(x,r) \in \mathbb{R}^N$ be given by $g_i(x,r) = r_i \frac{\partial u_i}{\partial x_i}$[14]. A game is *diagonally strictly concave* if

$$\exists r \in (\mathbb{R}_+^*)^N \mid \forall x^0 \neq x^1 \in X \text{ we have } \langle x^1 - x^0 \mid g(x^0, r) \rangle + \langle x^0 - x^1 \mid g(x^1, r) \rangle > 0 \tag{8}$$

Games having this property are denoted by $\mathcal{G}^{Ros}$. It is proved (Theorem 2 of Rosen (1965)) that games in $\mathcal{G}^{Ros}$ have a unique Nash equilibrium when the state space is compact. In our context, where the state space is unbounded, they may have none.

Games considered by Arrow and Hurwicz (1960) (which we call *concave-convex games*, and denote by $\mathcal{G}^{Arr}$) are as follow. Let S be a subset of N, the set of players, and define $f^S = \sum_{i \in S} u_i - \sum_{i \in N \setminus S} u_i$. A game is concave-convex if *a)* for each $S \subseteq N$, the function $f^S(x^S, x^{N \setminus S})$ is concave in x^S for each $x^{N \setminus S}$ and convex in $x^{N \setminus S}$ for each x^S, and *b)* for some $S^0 \subseteq N$, $f^{S^0}(x^{S^0}, x^{N \setminus S^0})$ is strictly concave in x^{S^0} for each $x^{N \setminus S^0}$. If in addition u_i is strictly concave in x_i, then we say that the game is

---

[14]The dynamical system $\dot{x} = g(x, r)$ is a weighted gradient system, and is significantly different from the system (2)

strictly concave-convex.

Remark 4 *Stricly concave-convex games are diagonally strictly concave, i.e. $\mathcal{G}^{Arr} \subset \mathcal{G}^{Ros}$. Thus all properties of the later apply to the former.*

For simplicity, in the remainder of this section we will place ourselves in the setting of Rosen (1965), i.e. we assume that the strategy space X is a compact set. This guarantees that the Nash equilibrium is unique. When the set X is unbounded, the game could have no equilibrium at all and if that happened, the process would go to infinity. Because this introduces unnecessary complexities in the proof, we restrict our attention to compact sets.

The fact that the Nash equilibrium is unique is convenient for the study of dynamics where the Nash equilibria are the only stationary points. However, the system (2) also has other zeros, since $Z = NE \cup OZ$. In the following theorem, we show that there is a finite number of other zeros, and thus all the stationary points are isolated. We also state our convergence result.

Theorem 5 *Let $\mathcal{G} \in \mathcal{G}^{Ros}$. Then,*

(i) Z *is a finite set*

(ii) *There is a unique Nash equilibrium $\bar{x}$ and*

$$\mathbb{P}(\lim_n x_n = \overline{x}) = 1.$$

The proof of the first point goes as follows: we prove that games in $\mathcal{G}^{Ros}$ are such that, after removing a subset of players playing 0, the remaining subgame is also in $\mathcal{G}^{Ros}$. Thus there is at most one Nash equilibrium for any combination of agents playing 0. The number of such potential combinations is finite, so the result follows.

In order to prove the second point of Theorem 5, we show that the zeros of (2) are the only candidates for limit points of our process. We cannot do this in general games with isolated zeros, but in diagonally strictly concave games we can, by decomposing the state space into several subspaces (respectively, the interior of the space and every face) and constructing appropriate Lyapunov functions for each subspace. As a consequence, we prove that every solution of (2) converges to one of the zeros. Since zeros are the only candidates, we get the desired conclusion by using point i) of Theorem 4.

Appendix

A   Proof of results of Section 2

Proof of Proposition 1.   We first prove that the process can be written as in equation (1). Second we prove that the process is well-defined, i.e. $x_n^i > 0$ for all i and all n.

1- We have, for any $i \in \mathcal{N}$,

$$e_{2n+2}^i - e_{2n}^i = e_{2n}^i \epsilon_n^i \Delta u_{n+1}^i$$

A first order development gives

$$
\begin{aligned}
\epsilon_n^i \Delta u_{n+1}^i &= \epsilon_n^i \left(u_i \left(e_{2n}^i + \frac{1}{n+1}\epsilon_n^i, e_{2n}^{-i} + \frac{1}{n+1}\epsilon_n^{-i} \right) - u_i(e_{2n}^i, e_{2n}^{-i}) \right) \\
&= \frac{1}{n+1}(\epsilon_n^i)^2 \frac{\partial u_i}{\partial x^i}(e_{2n}) + \frac{1}{n+1}\epsilon_n^i \sum_{j \neq i} \epsilon_n^j \frac{\partial u_i}{\partial x^j}(e_{2n}) + \mathcal{O}(\frac{1}{n^2})
\end{aligned}
$$

Because $(\epsilon_n^i)^2 = 1$ and $x_n = e_{2n}$, we have

$$x_{n+1}^i - x_n^i = \frac{1}{n+1}x_n^i \frac{\partial u_i}{\partial x^i}(x_n) + \frac{1}{n+1}\epsilon_n^i x_n^i \sum_{j \neq i} \epsilon_n^j \frac{\partial u_i}{\partial x^j}(x_n) + \mathcal{O}(\frac{1}{n^2})$$

By setting $U_{n+1}^i = \epsilon_n^i x_n^i \sum_{j \neq i} \epsilon_n^j \frac{\partial u_i}{\partial x^j}(x_n)$, we get equation (1). Finally, note that $\mathbb{E}(\epsilon_n^j) = 0$ for all j, and that ϵ_n^i and ϵ_n^j are independent, so that

$$\mathbb{E}(U_{n+1} \mid \mathcal{F}_n) = 0. \quad \blacksquare$$

2- Let us now show that the process is well-defined. Notice that Hypothesis 1 implies that Du_i is bounded everywhere. For simplicity and without loss of generality, we will assume that $|u_i(x) - u_i(x')| < \|x - x'\|_\infty$. This is just for simplicity, the proof can easily be accommodated otherwise. Let $n \geq 0$. By assumption, $|u_i(x) - u_i(x')| \leq \|x - x'\|_\infty$ for all i. Thus,

$$\frac{x_{n+1}^i}{x_n^i} \geq (1 - \|e_{2n+1} - x_n\|_\infty),$$

and $|e_{2n+1}^i - x_n^i| \leq \frac{1}{n+1}$ for all i. As a consequence,

$$\frac{x_{n+1}^i}{x_n^i} \geq (1 - \frac{1}{n+1}).$$

25

Thus, $x_1^i \geq 0$ and

$$x_n^i \geq x_1^i \prod_{k=1}^{n-1} \left(1 - \frac{1}{k+1}\right) = \frac{1}{n+1} x_1^i \geq 0.$$

Note that at the beginning of the process, steps are large. Thus in case x_0^i is close to 0, the exploration phase might take players to the negative orthant ($e_1 \leq 0$). This can only happen because the first steps are large. In order to avoid that, we can either assume that $x_0^i > 1$ (i.e. players start far enough from the boundary), or that the process begins at step $n \geq \min_i\{E(1/x_0^i)\} + 1$, where $E(a)$ is the integer part of a (i.e. the first steps are not too large). In any case, this is totally innocuous for what we do and guarantees that $e_1 \geq 0$. ∎

Proof of Proposition 2. Pick $\hat{x} \in OZ$ and assume without loss of generality that $\hat{x}_1 = 0$ with $\frac{\partial u_1}{\partial x_1}(\hat{x}) > 0$. Then

$$\frac{\partial F_1}{\partial x_1}(\hat{x}) = \frac{\partial u_1}{\partial x_1}(\hat{x}), \quad \text{and} \quad \frac{\partial F_1}{\partial x_j}(\hat{x}) = 0 \ \text{ for } \ j \neq 1.$$

Hence $\frac{\partial u_1}{\partial x_1}(\hat{x}) \in \text{Sp}(DF(\hat{x}))$, and the associated eigenvector is $v = (1, 0, ..., 0)$ which points inwards (i.e. $\hat{x} + v \in X$). Thus, necessarily $OZ \subset Z^{LU}$.

Next, $Z^S \subset NE$ is a consequence of $Z = NE \cup OZ$, $Z^S = Z \backslash Z^{LU}$ and $OZ \subset Z^{LU}$.
∎

B Proof of results of Section 3

B.1 Proof of Proposition 3

Under Assumption 3, for any i, there exists $\bar{x}_i > 0$ such that

$$\frac{\partial u_i}{\partial x_i}(x_i, 0) > \alpha_i > 0, \ \forall x_i \leq \bar{x}_i.$$

Since the game has strategic complements,

$$\frac{\partial u_i}{\partial x_i}(x_i, x_{-i}) > \alpha_i > 0, \ \forall x_i < \bar{x}_i, \ \forall x_{-i} \in X_{-i}. \tag{9}$$

As a consequence, any solution trajectory with initial condition in the set $\{x \in X : x_i \in]0, \bar{x}_i[\}$ is in the set $\{x \in X : x_i > \bar{x}_i\}$ after some finite time $t > 0$. Let $a = \min_i \bar{x}_i$. Therefore any invariant set is contained either in $[a, \infty[^N$ or in ∂X. Thus, by the aforementioned result of Benaïm (1999), we can conclude that

$$\mathbb{P}\left(\mathcal{L}((x_n)_n) \subset [a, \infty[^N\right) + \mathbb{P}\left(\mathcal{L}((x_n)_n) \subset \partial X\right) = 1.$$

In what follows we will show that $\mathbb{P}\left(\mathcal{L}((x_n)_n) \subset \partial X\right) = 0$. The main idea is to exploit the fact that the strategic complementarity condition implies that, if $x \in \partial X$ and for some coordinate $x_i = 0$ then $\frac{\partial u_i}{\partial x^i}(x)$ must be strictly positive (there is no Nash equilibria on ∂X).

Remark 5 *Three simple observations are in order.*

(i) *Condition (9) implies that if $\frac{\partial u_i}{\partial x^i}(x) \leq 0$, then $x^i \geq a$.*

(ii) *If $x \in X \setminus [a, +\infty[^N$, the set of coordinates for which $\frac{\partial u_i}{\partial x^i}(x) > 0$, I_x, is always nonempty. This is because if for all $i \in \{1, \ldots, N\}$, $\frac{\partial u_i}{\partial x^i}(x) \leq 0$ then $x \in [a, +\infty[^N$.*

(iii) *Moreover, also from (9), the coordinate k achieving the minimum value of a vector $x \in X \setminus [a, +\infty[^N$ verifies that $\frac{\partial u_k}{\partial x^k}(x) > \alpha$, where $\alpha = \min_i \alpha_i > 0$. Therefore this particular k belongs to the set I_x.*

Let $d(x, \partial X)$ be distance for the infinity norm of x to ∂X, i.e. $d(x, \partial X) = \min_i x_i$. Let us take $R > a$ and consider the sets $\mathcal{U}_R$:

$$\text{and } \mathcal{U}_R = \left\{ x \in X \,;\, \frac{\partial u_i}{\partial x^i}(x) < 0 \Rightarrow -x_i \frac{\partial u_i}{\partial x^i}(x) \leq R \right\}.$$

Observe that ∂X can be written as an increasing union of the form:

$$\partial X = \bigcup_{R=1}^{\infty} \left(\partial X \cap \mathcal{U}_R\right), \tag{10}$$

In order to show that $\mathbb{P}\left(\mathcal{L}((x_n)_n) \subset \partial X\right) = 0$, it is sufficient to prove that $\mathbb{P}\left(\mathcal{L}((x_n)_n) \subset \partial X \cap \mathcal{U}_R\right) = 0$, for all $R > a$. By contradiction, assume that there exists $R > a$ such that $\mathbb{P}\left(\mathcal{L}((x_n)_n) \subset \partial X \cap \mathcal{U}_R\right) > 0$ and let $0 < \varepsilon < a$. On the event $\{\mathcal{L}((x_n)_n) \subset \partial X \cap \mathcal{U}_R\}$, there exists a (random) $n_* \in \mathbb{N}$ such that

$$\mathbb{P}\left(\{\mathcal{L}((x_n)_n) \subset \partial X\} \cap \{x_n \in V_\varepsilon \cap \mathcal{U}_R, \text{ for all } n \geq n_*\}\right) > 0, \tag{11}$$

where
$$\mathcal{V}_\varepsilon = \{x \in X \,;\, d(x, \partial X) \leq \varepsilon\}\,.$$

In what follows, we work on the event E defined by (11) and we assume that $n \geq n_*$.

For $\beta > 0$, let the function
$$\Phi_\beta(x) = -\frac{1}{\beta} \ln \left(\sum_{i=1}^{N} \exp(-\beta x^i) \right),$$

which is concave if extended as $-\infty$ to $\mathbb{R}^N$. The function Φ verifies the well-known relation
$$\min_{i=1,\ldots,N} x^i - \frac{\ln(N)}{\beta} \leq \Phi_\beta(x) \leq \min_{i=1,\ldots,N} x^i. \tag{12}$$

From a straightforward calculation we have that, for all $i \in \{1, \ldots, N\}$,
$$\frac{\partial \Phi_\beta}{\partial x^i}(x) = \pi_i(x), \text{ where } \pi_i(x) = \frac{\exp(-\beta x^i)}{\sum_{j=1}^{N} \exp(-\beta x^j)}.$$

Also, for all $i, j \in \{1, \ldots, N\}$
$$\frac{\partial^2 \Phi_\beta}{\partial x^j \partial x^i}(x) = -\beta \pi_i(x)(\delta_{ij} - \pi_j(x)),$$

where $\delta_{ij} = 1$ if $i = j$, and 0 otherwise. This implies that $\nabla \Phi_\beta$ is L-Lipschitz. In fact, $L \leq 2\beta$ for the infinity norm.

Observe that if $x \in V_\varepsilon \cap \mathcal{U}_R$ and if $\frac{\partial u_i}{\partial x^i}(x) \leq 0$ for some coordinate i, we have that
$$\pi_i(x) \leq \exp\left(-\beta(a - \varepsilon)\right),$$

using the fact that it exists some k such that $x^k \leq \varepsilon$ and that $x^i \geq a$ (c.f. Remark 5).

On the other hand, for $k \in I_x$ such that $x^k = \min_i x^i$,
$$\pi_k(x) = \frac{1}{1 + \sum_{j \neq k} \exp\left(\beta x^k - \beta x^j\right)} \geq \frac{1}{N}.$$

Recall that the variable x_n follows the recursion
$$x_{n+1}^i = x_n^i + \frac{1}{n+1}\left(x_n^i \frac{\partial u_i}{\partial x_i}(x_n) + U_{n+1}^i + \xi_{n+1}^i\right),$$

where $\mathbb{E}\left(U_{n+1} \mid \mathcal{F}_n\right) = 0$ and $|\xi_n^i| \leq C/n$, for a deterministic constant C.

Let us define $z_n = \Phi_\beta(x_n)$. Note first that, from equation (12),

$$-\frac{\ln(N)}{\beta} \leq \min_{i=1,\dots,N} x_n^i - \frac{\ln(N)}{\beta} \leq z_n \leq \min_{i=1,\dots,N} x_n^i \leq \epsilon. \tag{13}$$

Consequently, $\mathcal{L}((z_n)_n) \subset [-\ln(N)/\beta, 0]$ almost surely on E.

On the other hand, since the function $-\Phi_\beta$ is convex with L-Lipschitz gradient, we have that

$$-\Phi_\beta(x_{n+1}) \leq -\Phi_\beta(x_n) + \langle -\nabla\Phi_\beta(x_n), \, x_{n+1} - x_n \rangle + \frac{L}{2} \|x_{n+1} - x_n\|^2.$$

Equivalently,

$$z_{n+1} \geq z_n + \sum_{j=1}^{N} \pi_j(x_n)(x_{n+1}^j - x_n^j) - \frac{L}{2} \|x_{n+1} - x_n\|^2,$$

$$= z_n + \frac{1}{n+1} \sum_{j=1}^{N} \pi_j(x_n) \left(x_n^j \frac{\partial u_i}{\partial x_j}(x_n) + U_{n+1}^j + \xi_{n+1}^j \right) - \frac{L}{2} \|x_{n+1} - x_n\|^2,$$

$$\geq z_n + \frac{1}{n+1} \sum_{j=1}^{N} \pi_j(x_n) x_n^j \frac{\partial u_i}{\partial x_j}(x_n) + \frac{1}{n+1} \sum_{j=1}^{N} \pi_j(x_n) U_{n+1}^j - \frac{c}{(n+1)^2},$$

for some deterministic constant $c \geq 0$. Therefore, taking conditional expectation and omitting the quadratic term,

$$\mathbb{E}\left(z_{n+1} \mid \mathcal{F}_n\right) \geq z_n + \frac{1}{n+1} \sum_{j=1}^{N} \pi_j(x_n) x_n^j \frac{\partial u_i}{\partial x_j}(x_n).$$

Recall that I_{x_n} is the set of indices such that $\frac{\partial u_i}{\partial x^i}(x_n) > 0$ and that k_n the coordinate giving the minimum of x_n is in I_{x_n} and verifies moreover that $\frac{\partial u_i}{\partial x^i}(x_n) > \alpha$. Let J_{x_n} the set of indices such that $\frac{\partial u_i}{\partial x^i}(x_n) \leq 0$.

For all $n \geq n_*$ we have

$$\mathbb{E}\left(z_{n+1} \mid \mathcal{F}_n\right) \geq z_n + \frac{1}{n+1} \sum_{j \in I_{x_n}} \pi_j(x_n) x_n^j \frac{\partial u_i}{\partial x_j}(x_n) + \frac{1}{n+1} \sum_{j \in J_{x_n}} \pi_j(x_n) x_n^j \frac{\partial u_i}{\partial x_j}(x_n),$$

$$\geq z_n + \frac{z_n}{n+1} \frac{\alpha}{N} + \frac{1}{n+1} \sum_{j \in J_{x_n}} \pi_j(x_n) x_n^j \frac{\partial u_i}{\partial x_j}(x_n),$$

using that $x_n^{k_n} \geq z_n$ and that $\pi_{k_n}(x_n) \geq 1/N$. On the other hand, using the definition of $\mathcal{U}_R$, we obtain

$$\sum_{j \in J_{x_n}} \pi_j(x_n) x_n^j \frac{\partial u_i}{\partial x_j}(x_n) \geq -|J_{x_n}| R \exp\left(-\beta(a - \varepsilon)\right) \geq -NR \exp\left(-\beta(a - \varepsilon)\right).$$

29

Thus

$$\mathbb{E}\left(z_{n+1} \mid \mathcal{F}_n\right) \geq z_n + \frac{1}{n+1}\left(\frac{\alpha}{N}z_n - NR\exp\left(-\beta(a-\varepsilon)\right)\right).$$

Let us consider the change of variables

$$\theta_n = \left(z_n + \frac{\ln(N)}{\beta}\right) \geq \min_i x_n^i \geq 0.$$

Then,

$$\mathbb{E}(\theta_{n+1}) \geq \mathbb{E}(\theta_n) + \frac{\alpha}{N}\frac{1}{n+1}\left(\mathbb{E}(\theta_n) - \underbrace{\left\{\frac{N^2}{\alpha}R\exp\left(-\beta(a-\varepsilon)\right) + \frac{\ln(N)}{\beta}\right\}}_{c(\beta)}\right).$$

Let us note that $\mathbb{E}(\theta_{n_*}) > 0$ since $\min_i x_{n_*}^i \geq 1/(n_*+1)$ almost surely. Now, we can fix $\beta > 0$ sufficiently large such that $0 < c(\beta) < \mathbb{E}(\theta_{n_*})$. So that

$$\mathbb{E}(\theta_{n+1}) \geq \mathbb{E}(\theta_n) + \frac{\alpha}{N}\frac{1}{n+1}\left(\mathbb{E}(\theta_n) - c(\beta)\right).$$

Let us call $\rho_n = \mathbb{E}(\theta_n) - c(\beta)$. Then, we want to analyse the recursion $\rho_{n+1} \geq \rho_n(1 + \frac{\alpha}{N}\frac{1}{n+1})$, with $\rho_{n_*} > 0$. Hence, for $n \geq n_*$,

$$\rho_{n+1} \geq \rho_{n_*}\prod_{i=n_*}^{n}\left(1 + \frac{\alpha}{N}\frac{1}{i+1}\right),$$

where the right-hand-side goes to infinity. Finally, we can conclude that $\mathbb{E}(z_n)$ goes to infinity, which is a contradiction with the fact that $z_n \in [-\ln(N)/\beta, \epsilon]$ almost surely on the event E. $\blacksquare$

B.2   Proof of Theorem 1

The proof of Theorem 1 involves several arguments. For point (i), we use a result from Pemantle (1990), while for point (ii) we adapt a result from Benaïm and Faure (2012). For both points we use Proposition 3, i.e the fact that the limit set of the process cannot include points on the boundary of the state space.

For the first point of Theorem 1, let us recall some results on non convergence. Let $\tilde{x}$ be a linearly unstable equilibrium. Assume without loss of generality that the unstable space at $\tilde{x}$ is one-dimensional, that is $DF(\tilde{x})$ has only one eigenvalue μ with

30

positive real part, and we call v the associated normalized eigenvector. We use a result from Pemantle (1990), more precisely in the settings of Brandiere and Duflo (1996), which states that a sufficient condition for non convergence to $\tilde{x}$ is that the noise is exciting in the unstable direction, i.e.:

$$\liminf_{n \to +\infty} \mathbb{E}\left(\langle U_{n+1}, v \rangle^2 \mid \mathcal{F}_n\right) > 0 \tag{14}$$

on the event $\{\lim_n x_n = \tilde{x}\}$.

Consider any x_n and any vector v. Then

$$\langle U_{n+1}, v \rangle^2 = \left(\sum_{i<j} \epsilon_n^i \epsilon_n^j \left(v_i x_n^i \frac{\partial u_i}{\partial x^j}(x_n) + v_j x_n^j \frac{\partial u_j}{\partial x^i}(x_n)\right)\right)^2$$

Using $\mathbb{E}\left(\epsilon_n^i \epsilon_n^j\right) = 0$ if $i \neq j$ and $(\epsilon_n^i)^2 = 1$, we get

$$\mathbb{E}\left(\langle U_{n+1}, v \rangle^2 \mid x_n\right) = \sum_{i<j} \left(v_i x_n^i \frac{\partial u_i}{\partial x^j}(x_n) + v_j x_n^j \frac{\partial u_j}{\partial x^i}(x_n)\right)^2 \tag{15}$$

Proof of Theorem 1 (i). Let $\tilde{x}$ be a linearly unstable equilibrium. We want to show that

$$\sum_{i<j} \left(v_i \tilde{x}^i \frac{\partial u_i}{\partial x^j}(\tilde{x}) + v_j \tilde{x}^j \frac{\partial u_j}{\partial x^i}(\tilde{x})\right)^2 \neq 0.$$

where v is the normalized eigenvector associated associated to the unstable direction of $\tilde{x}$.

Note that $\tilde{x} \notin \partial X$ by Proposition 3, and that when $\tilde{x} \in \text{Int}(X)$ and the interaction graph is non-bipartite, then the result is a direct implication of Theorem 4 in Section 5. Thus, here we assume that the interaction graph is connected and bipartite, and that $\tilde{x} \in \text{Int}(X)$. This implies that there exists a partition (A, B) of N such that if $a \in A$ and $\frac{\partial u_a}{\partial x_b}(\tilde{x}) \frac{\partial u_b}{\partial x_a}(\tilde{x}) > 0$ then $b \in B$.

Using the computations just developed, we need to show that

$$\sum_{a<b} \left(v_a x^a \frac{\partial u_a}{\partial x^b}(\tilde{x}) + v_b x^b \frac{\partial u_b}{\partial x^a}(\tilde{x})\right)^2 \neq 0. \tag{16}$$

Assume the contrary. Then we must have $v_a x^a \frac{\partial u_a}{\partial x^b}(\tilde{x}) + v_b x^b \frac{\partial u_b}{\partial x^a}(\tilde{x}) = 0$ for all $a \in A$ and all $b \in B$. Because $x^i > 0$ for all i, and by Hypothesis 2 (symmetric externalities),

it must be that $\text{sgn}(v_a) = -\text{sgn}(v_b)$ for any $a \in A$ and any $b \in B$. Since the interaction graph is connected, we may assume without loss of generality that $v_a > 0 \, \forall \, a \in A$ and $v_b < 0 \, \forall \, b \in B$.

Because μ is strictly positive and v is the corresponding normalized eigenvector, we should have $\langle vDF(\tilde{x}), v \rangle = \mu \sum_i v_i^2 > 0$, since $v \neq 0$. However, we will show that this can only be true if equation (16) holds. By a simple rearrangement of the indexes, the Jacobian matrix at $\tilde{x}$ can be written as follows:

$$DF(\tilde{x}) = \begin{pmatrix} D_A & M \\ N & D_B \end{pmatrix},$$

where D_A is diagonal and the diagonal terms are equal to $x^a \partial^2 u_a / \partial (x^a)^2 (\tilde{x}) \leq 0$ with $a \in A$; and similarly for D_B. M and N are non-negative matrices, as $x^i \partial^2 u_i / \partial x^i \partial x^j \geq 0 \, \forall i \neq j$.

Thus,

$$\langle vDF(\tilde{x}), v \rangle = \sum_i v_i^2 x^i \partial^2 u_i / \partial (x^i)^2 + \sum_{a \in A, b \in B} v_a v_b \left(x^a \frac{\partial^2 u_a}{\partial x^a \partial x^b} + x^b \frac{\partial^2 u_b}{\partial x^a \partial x^b} \right) \leq 0,$$

a contradiction. To see why this inequality holds, remember that the terms in the first sum are all negative by Hypothesis 1 and the fact that $\tilde{x}$ is a Nash equilibrium. The terms in the second sum are also all negative since $v_a . v_b < 0$ and by strategic complements. ∎

For the proof of point (ii), we use the following theorem of Benaïm and Faure (2012), conveniently adapted to our setting.

Theorem B.1 (Benaïm and Faure, 2012) *Let $(x_n)_n \in X$ be a random process that can be written as*

$$x_{n+1} = x_n + \frac{1}{n+1} \left(F(x_n) + U_{n+1} + \xi_{n+1} \right)$$

where

  (i) $F : X \to \mathbb{R}^N$ is a smooth map, that is cooperative and irreducible in $\text{Int}(X)$,

 (ii) U_{n+1} is a bounded martingale difference and is uniformly exciting, i.e. the matrix

$$\mathbb{E} \left(U_{n+1} U_{n+1}^T \mid x_n = x \right)$$

  is positive definite for any $x \in \text{Int}(X)$,

(iii) $\xi_n = \mathcal{O}(1/n)$, *and*

(iv) *there exists $a > 0$ such that $\mathcal{L}(x_n)_n \subset [a, +\infty[$ almost surely.*

Then

$$\mathbb{P}\left(\exists\, x^* \in Z^S : \lim_n x_n = x^*\right) = 1$$

on the event $\{\limsup_n \|x_n\| < +\infty\}$.

Proof of Theorem 1 (ii).   We want to apply Theorem B.1.

When the game is of strategic complements, our dynamics $\dot{x} = F(x)$ is *cooperative* because all non-diagonal entries of $DF(x)$ are nonnegative. In addition, Hypothesis 2 guarantees that the interaction graph is strongly connected. Thus the matrix $DF(x)$ is *irreducible* for any x in the interior of X. These two facts provide point (i). Points (iii) and (iv) follow from Propositions 1 and 3, respectively. To prove point (ii), we prove that if a network is non-bipartite and the game exhibits symmetric interactions, then the noise is uniformly exciting.

Since for any $v \in \mathbb{R}^N$ we have that $v^T U_{n+1} U_{n+1}^T v = \langle U_{n+1}, v \rangle^2 \geq 0$, (ii) amounts to showing that, for any $x \in \mathrm{Int}(X)$, we have

$$\mathbb{E}\left(\langle U_{n+1}, v \rangle^2 \mid x_n = x\right) = 0$$

if and only if $v = 0$. By equation (15), we see that the condition is verified in v if and only if

$$\forall i < j, \ v_i \tilde{x}^i \frac{\partial u_i}{\partial x^j}(\tilde{x}) + v_j \tilde{x}^j \frac{\partial u_j}{\partial x^i}(\tilde{x}) = 0.$$

We now prove that under the assumption of symmetric interactions and non-bipartiteness of the graph, this quantity is positive.

Since the interaction graph is non bipartite in $\tilde{x}$, there is at least one odd cycle. Let us assume, for simplicity but without loss of generality, that this cycle is of length 3: there exist i, j, k such that

$$\frac{\partial u_i}{\partial x_j}(\tilde{x})\frac{\partial u_j}{\partial x_i}(\tilde{x}) > 0, \frac{\partial u_j}{\partial x_k}(\tilde{x})\frac{\partial u_k}{\partial x_j}(\tilde{x}) > 0, \frac{\partial u_i}{\partial x_k}(\tilde{x})\frac{\partial u_k}{\partial x_i}(\tilde{x}) > 0.$$

We thus have $\mathrm{sgn}(v_i) = -\mathrm{sgn}(v_j) = \mathrm{sgn}(v_k) = -\mathrm{sgn}(v_i)$ which implies, since $\tilde{x}$ is interior, that $v_i = v_j = v_k = 0$. As a consequence, for every agent l linked to i, j or k, we must have $v_l = 0$. Recursively, since the interaction graph is connected, we must have $v = 0$, which concludes the proof.   ∎

33

C Proof of results of Section 4.

C.1 Proof of Theorem 2

Before proving Theorem 2, let us define the following dynamical concept:

Definition C.1 *Let $P : X \to \mathbb{R}$ be continuously differentiable. We say that P is a strict[15] Lyapunov function for $\dot{x} = F(x)$ if*

- *for $x \in Z$ the map $t \mapsto P(\varphi(x,t))$ is constant;*

- *for $x \notin Z$ the map $t \mapsto P(\varphi(x,t))$ is strictly increasing.*

Lemma C.1 *Assume that $\mathcal{G}$ is an LOPG with continuously differentiable potential P. Then*

(i) *P is a strict Lyapunov function for $\dot{x} = -x + \mathrm{BR}(x)$.*

(ii) *P is a strict Lyapunov function for $\dot{x} = F(x)$ (where $F_i(x) = x_i \frac{\partial u_i}{\partial x_i}(x)$).*

Proof. By assumption,

$$\forall x, \forall i, \ \frac{\partial u_i}{\partial x_i}(x) > 0 \Rightarrow \frac{\partial P}{\partial x_i}(x) > 0 \text{ and } \frac{\partial u_i}{\partial x_i}(x) < 0 \Rightarrow \frac{\partial P}{\partial x_i}(x) < 0.$$

(i) We have

$$\langle DP(x), -x + \mathrm{BR}(x) \rangle = \sum_i \frac{\partial P}{\partial x_i}(x)(-x_i + \mathrm{BR}_i(x)).$$

We need to check that, if $x \notin NE$, then this quantity is positive. Let i be such that $x_i \neq \mathrm{BR}_i(x)$, say $x_i < \mathrm{BR}_i(x_{-i})$. Then by strict concavity of u_i we have $\frac{\partial u_i}{\partial x_i}(x) > 0$. Thus $\frac{\partial P}{\partial x_i}(x) > 0$ and $\langle DP(x), -x + \mathrm{BR}(x) \rangle > 0$.

(ii) We have

$$\langle DP(x), F(x) \rangle = \sum_i x_i \frac{\partial u_i}{\partial x_i}(x) \frac{\partial P}{\partial x_i}(x).$$

---

[15]Generally, P is a Lyapunov function for $\dot{x} = F(x)$ with respect to Λ if $t \mapsto P(\varphi(x,t))$ is constant on Λ and strictly increasing for $x \notin \Lambda$; when the component Λ coincides with the set of stationary points of the flow, then we say that P is strict.

We need to check that, if $x \notin Z$, then this quantity is positive. Let i be such that $F_i(x) \neq 0$. Then $x_i > 0$ and $\frac{\partial u_i}{\partial x_i}(x) \neq 0$, which implies that

$$x_i \frac{\partial u_i}{\partial x_i}(x) \frac{\partial P}{\partial x_i}(x) > 0.$$

and the proof is complete. ∎

Proof of Theorem 2 (i). For this part, we use the general result given by Proposition 6.4 in Benaïm (1999), which asserts that if P is a strict Lyapunov function with respect to Z and $P(Z)$ has empty interior, then $\mathcal{L}(x_n) \subset Z$ almost surely. Therefore, the following lemma finishes the proof. ∎

Lemma C.2 *Assume $\mathcal{G}$ is an LOPG. Then, if P is sufficiently regular, $P(Z)$ has an empty interior.*

Proof. We decompose the set of zeroes of F as a finite union of sets on which we can use Sard's Theorem.

Let A be any subset of agents and Z_A be the set

$$\left\{ x \in Z : \ x_i = 0 \ \forall i \notin A, \ \frac{\partial u_i}{\partial x_i} = 0 \ \forall i \in A \right\}.$$

It is not hard to see that Z_A is closed. Moreover $Z = \cup_{A \in \mathcal{P}(\{1,\dots,N\})} Z_A$.

We now prove that P is constant on Z_A. Let $P^A : [0,1]^A \to \mathbb{R}$ be defined as

$$P^A(z) := P(z, 0).$$

For $x \in Z_A$, denote by $x^A = (x_i)_{i \in A}$. We then have $P^A(x^A) = P(x)$. Moreover, for $i \in A$,

$$\frac{\partial P^A}{\partial x_i} = 0$$

by definition of Z_A and the additional assumption we made on P. Hence

$$\{x^A : \ x \in Z_A\} \subset \{z \in [0,1]^A : \ \nabla_z P^A = 0\}.$$

Now P is sufficiently differenciable, so is P^A, and by Sard's Theorem, $P^A(\{x^A : \ x \in Z_A\})$ has empty interior in $\mathbb{R}^A$. As an immediate consequence, P^A is constant on $\{x^A : \ x \in Z_A\}$, which directly implies that $P(Z_A)$ has empty interior. Since Z is a finite union of such sets, $P(Z)$ has empty interior. ∎

Proof of Theorem 2 (ii). This proof relies on Theorem 4 in Section 5, and on the following: by Lemma C.1 and Corollary 6.6 in Benaïm (1999), we have

$$\mathbb{P}\left(\exists \hat{x} \in Z \text{ such that } \lim_n x_n = \hat{x}\right) = 1.$$

Because convergence to the zeroes is guaranteed, Theorem 4 gives us the result. Point a) of Theorem 4 gives us the first point of Theorem 2 (ii), and the second point of Theorem 2 (ii) is a consequence of point b) of Theorem 4. ∎

C.2 Proof of Theorem 3

First we prove that (i) implies (ii). Since $\mathcal{G}$ is an LOPG, $P(Z)$ has empty interior (see Lemma C.2 above). Moreover, we have $\Lambda \subset Z$. Thus P is constant on Λ. Let $v := P(\Lambda)$. If Λ is not a local maximum of P then there exists a sequence x_n such that $d(x_n, \Lambda) \to_n 0$ and $P(x_n) > v$. Since Λ is isolated we have $x_n \in X \setminus Z$ and $P(\varphi(x_n, t)) > P(x_n) > v$ for any $t > 0$ hence $d(\varphi(x_n, t), \Lambda) \not\to 0$ and Λ is not an attractor.

Let us now prove that (ii) implies (iii). First we show that Λ is contained in NE. Suppose that there exists $\hat{x} \in \Lambda \setminus NE$. Without loss of generality, we suppose that

$$\hat{x}_1 = 0, \ \frac{\partial u_1}{\partial x_1}(\hat{x}) > 0.$$

Since $\frac{\partial u_1}{\partial x_1}(\hat{x}) > 0$, we also have $\frac{\partial P}{\partial x_1}(\hat{x}) > 0$, by definition of an LOPG. As a consequence, $\hat{x}$ is not a local maximum of P.

We now prove that Λ is an attractor for the Best-Response dynamics. P is a strict Lyapunov function for the best-response dynamics[16] and $\Lambda \subset NE$. The statement we want to prove is then a consequence of Proposition 3.25 in Benaïm et al. (2005). We adapt the proof in our context for convenience. First of all observe that Λ is actually a strict local maximum of P: there exists an open (isolating) neighborhood U of Λ such that $P(x) < v = P(\Lambda), \ \forall x \in U \setminus \Lambda$. This is a simple consequence of the fact that P is strictly increasing along any solution curve with initial conditions in $U \setminus \Lambda$. Now let $V_r := \{x \in U : P(x) > v - r\}$. Clearly $\cap_r V_r = \Lambda$. Also $\varphi(\overline{V_r}, t) \subset V_r$, for $t > 0$, r small enough[17]. This implies that $\Lambda = \cap_{r>0} V_r$ contains an attractor A.

---

[16]Keep in mind that this means that it is a lyapunov function with respect to NE.

[17]We need to make sure that r is small enough so that $\overline{V_r} = P^{-1}([v - r, v]) \subset U$

The potential being constant on Λ, A cannot be strictly contained in Λ and therefore Λ is an attractor.

Now clearly (iii) implies (i): $\Lambda = \omega_{\mathrm{BR}}(U)$ for some open neighborhood U of Λ. Since $U \cap Z \subset NE$, $\omega_F(U) = \omega_{\mathrm{BR}}(U)$ and the proof is complete. $\blacksquare$

D   Proof of results of Section 5

D.1   Proof of Theorem 4.

Proof of point i). Pick an $\hat{x} \in OZ$ and let us fix $i \in \{1, \ldots, N\}$ such that $\hat{x}^i = 0$ and $\frac{\partial u^i}{\partial x^i}(\hat{x}) > 0$. Observe first that we can work on the event $\{\sup_{n \to +\infty} \|x_n\| < +\infty\}$ since, otherwise, there is nothing to prove. We proceed following a similar argument as in Posch (1997).

Let us assume by contradiction that $\mathbb{P}(\lim_n x_n = \hat{x}) > 0$. By continuity and from the fact that $\hat{x}$ is an isolated point in OZ, there exists a neighborhood $\mathcal{V}$ of $\hat{x}$ such that $\frac{\partial u^i}{\partial x^i} \geq \eta > 0$ for all $x \in \mathcal{V}$ and we can choose $k_* \in \mathbb{N}$ such that

$$\mathbb{P}\left(\{\lim_n x_n = \hat{x}\} \cap \{x_n \in \mathcal{V}, \text{ for all } n \geq k_*\}\right) > 0.$$

Let $\tilde{U}_{n+1}^i = \epsilon_n^i \sum_j \epsilon_n^j \frac{\partial u_i}{\partial x_j}(x_n)$, so that

$$x_{n+1}^i = x_n^i \left(1 + \frac{1}{n+1}\left(\frac{\partial u_i}{\partial x_i}(x_n) + \tilde{U}_{n+1}^i + \frac{\xi_{n+1}^i}{x_n^i}\right)\right).$$

Using a Taylor expansion and the fact that $\xi_n^i = \mathcal{O}(\frac{1}{n})$ and $x_n^i \geq 1/(n+1)$ for n sufficiently large, we obtain that

$$\frac{1}{x_{n+1}^i} = \frac{1}{x_n^i}\left(1 - \frac{1}{n+1}\left(\frac{\partial u_i}{\partial x_i}(x_n) + \tilde{U}_{n+1}^i\right) + o\left(\frac{1}{n}\right)\right).$$

Using that, for $n \geq k_*$, $\frac{\partial u_i}{\partial x_i}(x_n) \geq \eta$ and $\mathbb{E}(\tilde{U}_{n+1}^i \mid \mathcal{F}_n) = 0$, we obtain

$$\mathbb{E}\left(\frac{1}{x_{n+1}^i} - \frac{1}{x_n^i} \mid \mathcal{F}_n\right) \leq -\frac{1}{x_n^i} \cdot \frac{1}{n+1} \cdot \frac{\eta}{2} \leq 0.$$

Therefore, the random sequence $(1/x_n^i)_n$ is a positive supermartingale. It then converges almost surely to some random variable Y. However, on the event $\{\lim_n x_n = \hat{x}\}$, we have that x_n^i tends to zero almost surely. These two convergence properties are in contradiction and the conclusion follows. $\blacksquare$

Proof of point ii**).** As in the proof of Theorem 1, we need to show that the condition (14) is verified in the unstable direction. However, we have proved in Theorem 1 part (ii) that the quantity (15) is always strictly positive, unless $v = 0$. ∎

D.2   Proof of results in Section 5.2

Proof of Remark 4: Following Rosen (1965), we define $G(x, r)$ as the Jacobian matrix of $g(x, r)$, with $r_i \geq 0$. A sufficient condition for a game to belong to $\mathcal{G}^{Ros}$ is that $G(x, r) + G'(x, r)$ is negative definite, where G' is the transpose of G. For simplicity, we set $r = \mathbf{1}$, so that $g_i(x, \mathbf{1}) = \frac{\partial u_i}{\partial x_i}$ and $G_{ij}(x, \mathbf{1}) = \frac{\partial^2 u_i}{\partial x_i \partial x_j}$, and show that games in $\mathcal{G}^{Arr}$ are such that $G(x, \mathbf{1}) + G'(x, \mathbf{1})$ is negative definite. Define the matrices A, B^k and C as follows:

$A_{ii} = \frac{\partial^2 u_i}{\partial x_i^2}$ and $A_{ij} = 0$ if $i \neq j$

$B_{ij}^k = 0$ if $i = k$ or $j = k$ and $B_{ij}^k = \frac{\partial^2 u_k}{\partial x_i \partial x_j}$ if $i \neq k$ and $j \neq k$

$C_{ij} = \sum_k \frac{\partial^2 u_k}{\partial x_i \partial x_j}$

Then $G(x, \mathbf{1}) + G'(x, \mathbf{1}) = A(x) - \sum_k B^k(x) + C(x)$. By concavity of u_i in x_i, A is negative semi-definite and is negative definite as soon as one u_i is strictly concave in x_i. Every B^k is positive semi-definite by convexity of u_i in x_{-i}. Finally, strictly concave-convex games are such that $\sum_k u_k(x)$ is concave in x, by taking $S = N$ in the definition of strictly concave-convex games. Thus C is negative semi-definite. This proves that $G(x, \mathbf{1}) + G'(x, \mathbf{1})$ is negative definite. ∎

Proof of Theorem 5: Suppose first that there is a unique Nash equilibrium. Then note that under (8) we have, for any $x \neq \overline{x}$,

$$\langle \overline{x} - x, g(x, r) \rangle > 0,$$

because

$$\langle x - \overline{x} \mid g(\overline{x}, r) \rangle = \sum_{i : \overline{x}_i = 0} r_i x_i \frac{\partial u_i}{\partial x_i}(\overline{x}) \leq 0.$$

Given an element $x \in X$, let $I(x) := \left\{ i \in \mathcal{N} : x_i = 0 \text{ and } \frac{\partial u_i}{\partial x_i}(x) > 0 \right\}$. Given $J \subset N$, we call $\mathcal{G}^J$ the $N - |J|$-player game where the set of players is $N \setminus J$ and, for any strategy profile $z \in [0, +\infty[^{N-|J|}$, the payoff function of player $i \in \mathcal{N} \setminus J$ is $u_i^J(z) := u_i(z, 0^{|J|})$.

38

Lemma D.1 *Let $J \subset N$. There exists a unique profile $\tilde{x}^J$ with the following properties:*

(i) $J \subset I(\tilde{x}^J)$,

(ii) $\tilde{z} := (\tilde{x}_i^J)_{i \notin J}$ is a Nash equilibrium of $\mathcal{G}^J$.

and $\tilde{x}^J \in Z(F)$. Moreover, if $J \subset I(\overline{x})$ then $\tilde{x}^J = \overline{x}$. If not then $\tilde{x}^J$ belongs to OZ.

Proof. Fix $J \subset N$. The associated game $\mathcal{G}^J$ is also strictly diagonally concave. Thus it admits a unique Nash equilibrium $\tilde{z} \in [0, +\infty[^{N-|J|}$. Note that $J \subset I(\tilde{z})$ but is not necessarily equal. Now let $\tilde{x}^J := (\tilde{z}, 0^J)$. Clearly $\tilde{x}^J$ is the only element of X satisfying both (i) and (ii). Let $i \notin J$. We have $\tilde{x}_i^J = \tilde{z}_i = BR_i^J(\tilde{z}_{-i}) = BR_i(\tilde{z}_{-i}, 0^J) = BR_i(\tilde{x}_{-i}^J)$. This proves that $\tilde{x}^J$ belongs to $Z(F)$.

Now suppose that $J \subset I(\overline{x})$. Then $\overline{x}$ satisfies (i). Moreover for all $i \notin J$,

$$\overline{x}_i = \mathrm{BR}_i(\overline{x}_{-i}) = \mathrm{argmax}_{x_i} u_i(x_i, \overline{x}_{-i}) = \mathrm{argmax}_{z_i} u^J(z_i, \overline{z}_{-i}),$$

by definition of u^J and the fact that $\overline{x}_j = 0$ for any $j \in J$. Thus $(\overline{x}_i)_{i \notin J}$ is a Nash equilibrium of $\mathcal{G}^J$ and $\tilde{x}^J = \overline{x}$. Finally if J is not contained in $I(\hat{x})$ then $\tilde{x}^J \neq \overline{x}$ because $\overline{x}$ does not satisfy (i). ∎

As a consequence, $\{\tilde{x}^J, J \subset I\}$ can be written as $\{\overline{x}, \tilde{x}^1, ..., \tilde{x}^K\}$ where all elements are distinct, and there is a natural partition of X:

$$X = \left(\cup_{k=1}^K \tilde{X}^k \right) \cup \overline{X}, \quad \text{where } \tilde{X}^k := \{x \in X : \hat{x}^{I(x)} = \hat{x}^k\} \text{ and } \overline{X} := \{x \in X : \hat{x}^{I(x)} = \overline{x}\}.$$

Note that $\overline{X} = \{x \in X : I(x) \subset I(\overline{x})\}$ and the sets $\overline{X}, \tilde{X}^k, k = 1, ..., K$ are convex. More accurately every $\overline{X}, \tilde{X}^k$ is a union of faces of X: there exist $\overline{\mathcal{J}}$ and a family $(\mathcal{J}_k)_{k=1,...,K}$ of subsets of N such that:

$$\overline{X} = \cup_{J \in \overline{\mathcal{J}}} \{x \in X : I(x) = J\} \tilde{X}^k = \cup_{J \in \mathcal{J}^k} \{x \in X : I(x) = J\}.$$

Now we are ready to prove the theorem, i.e. when a game is diagonally strictly concave with unique Nash equilibrium $\overline{x}$, necessarily

$$\mathbb{P}(\lim_n x_n = \overline{x}) = 1$$

First let $x \in \overline{X}$, which amounts to having $I(x) \subset I(\overline{x})$ and define, for $x \in \overline{X}$,

$$\Phi(x) = \sum_{i \in I(\overline{x})} r_i x_i + \sum_{i \notin I(\overline{x})} r_i (x_i - \overline{x}_i \log(x_i)).$$

39

Then Φ is concave on $\overline{X}$ and achieves its minimum in $\overline{x}$. Let $\phi(t) = \Phi(\mathbf{x}(t))$, where $\mathbf{x}(t)$ is a solution of $\dot{x} = F(x)$, with $\mathbf{x}(0) \in \overline{X}$. We have

$$\frac{d}{dt}\phi(t) = \sum_{i \in \mathcal{N}} r_i(x_i(t) - \overline{x}_i)\frac{\partial u_i}{\partial x_i}(\mathbf{x}(t)) \leq 0,$$

with equality if and only if $x = \overline{x}$ and $\overline{x}$ is a global attractor the flow $\Phi_{|\overline{X}}$.

Now suppose that $x \in \tilde{X}^k$ for a given $k \in \{1, ..., K\}$. Note that $I(x) \subset I(\tilde{x}^k)$. We can then define $\Phi^k : \tilde{X}^k \to \mathbb{R}$ as the following:

$$\Phi^k(x) = \sum_{i \in I(\tilde{x}^k)} r_i x_i + \sum_{i \notin I(\tilde{x}^k)} r_i(x_i - \tilde{x}_i^k \log(x_i)).$$

Then Φ^k is again concave, with unique maximum in $x = \tilde{x}^k$ on $\tilde{X}^k$. Let $\phi(t) = \Phi(\mathbf{x}(t))$, where $\mathbf{x}(t)$ is a solution of $\dot{x} = F(x)$, with $\mathbf{x}(0) \in \tilde{X}^k$. We have

$$\frac{d}{dt}\phi(t) = \sum_{i \in \mathcal{N}} r_i(x_i(t) - \overline{x}_i)\frac{\partial u_i}{\partial x_i}(\mathbf{x}(t)) \leq 0,$$

with equality if and only if $x = \tilde{x}^k$. Thus $\tilde{x}^k$ is a global attractor the flow $\Phi_{|\tilde{X}^k}$.

As a consequence every solution curve converges to a zero of F, i.e either $\overline{x}$ or one of the $\tilde{x}^k$.[18] More precisely, $\overline{X}$ and $\tilde{X}^k$ are invariant and $\{\overline{x}\}$ (resp. $\tilde{x}^k$) is a global attractor for the flow $\phi_{|\overline{X}}$ (resp. $\phi_{|\tilde{X}^k}$); in particular, for any $x_0 \in \overline{X}$ (resp. $x_0 \in \tilde{X}^k$) then $\lim_{t \to +\infty} \phi_t(x_0) = \overline{x}$ (resp. $\lim_{t \to +\infty} \phi_t(x_0) = \tilde{x}^k$).

A set L is *internally chain transitive (ICT)* for the flow ϕ_t if it is compact, invariant and the restriction of the flow $\phi_{|L}$ admits no proper attractor. Of course $L_k := \{\tilde{x}_k\}$, as well as $\overline{L} := \{\overline{x}\}$ are ICT.

Theorem D.1 (Benaim, 1999) *The limit set of $(x_n)_n$ is almost surely internally chain transitive. Moreover let L be an internally chain transitive set for a flow $(\phi_t)_t$ and A be an attractor with basin of attraction $\mathcal{B}(A)$. If $L \cap \mathcal{B}(A) \neq \emptyset$ then $L \subset A$.*

We now prove that the sets L_k and $\overline{L}$ are the only internally chain transitive sets. This will conclude the proof. Note that $\overline{X}$ is an open set in X. To do so we first claim that it is always possible to relabel the family $(\tilde{x}^k)_{k=1,...,K}$ such that $\tilde{X}^k$ is an open set of $\cup_{l=1}^k X^l$ for $k = 2, ..., K$.

---

[18]This is not enough to guarantee that our random process converges with probability one to one of the zeroes of the dynamics.

Let L be internally chain transitive. By previous result, if L intersects $\overline{X}$ then $L \subset \{\overline{x}\}$ because $\overline{X}$ is the basin of attraction of $\overline{x}$. Suppose that it is not the case. then $L \subset \cup_{k=1}^{K} X^k$. Since X_K is open in $\cup_{k=1}^{K} X^k$, $\tilde{x}^K$ is an attractor of the flow restricted to $\cup_{k=1}^{K} X^k$, with basin of attraction $\tilde{X}^k$. Hence if $L \cap \tilde{X}^k \neq \emptyset$ then $L = \{\tilde{x}^k\}$. By a recursive argument, either $L = \{\overline{x}\}$ or $L = \{\tilde{x}^k\}$ for some k. ■

E   Examples

In this section, we illustrate through two examples the importance of two conditions we have used in this paper. The first example illustrate why bipartite interaction graphs might cause some trouble, while the second shows why the assumption of symmetric externalities matters.

In every proof of non convergence, the key argument we used relied on the noise condition (14).

Example E.1 *Consider the following 4-player example with strategic substitutes.*

$$u_1(x) = -cx_1 + b(x_1 + x_2 + x_4), u_2(x) = -cx_2 + b(x_2 + x_1 + x_3),$$
$$u_3(x) = -cx_3 + b(x_3 + x_2 + x_4), u_4(x) = -cx_4 + b(x_4 + x_1 + x_3),$$

with b strictly concave and such that $b'(1) = c$. This is a game of strategic substitutes, with an interaction graph represented by the square in Figure 1.
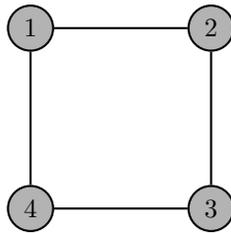

Figure 1: Different isolated Nash equilibria

One can check that the profile $\hat{x} = (1/3, 1/3, 1/3, 1/3)$ is a Nash equilibrium. Choosing b such that $b''(1) = -3$ for simplicity, the Jacobian matrix associated to $\hat{x}$

is

$$DF(\hat{x}) = \begin{pmatrix} -1 & -1 & 0 & -1 \\ -1 & -1 & -1 & 0 \\ 0 & -1 & -1 & -1 \\ -1 & 0 & -1 & -1 \end{pmatrix},$$

which eigenvalues are $-3, -1, -1, 1$. Thus this Nash equilibrium is linearly unstable. However, the eigenspace associated to the positive eigenvalue is generated by $v = (1, -1, 1, -1)$ so that, on the event $\{\lim_n x_n = \hat{x}\}$, we have

$$
\begin{aligned}
\lim_{n \to +\infty} \mathbb{E}\left(\langle U_{n+1}, v \rangle^2 \mid \mathcal{F}_n \right) &= \left(\frac{\partial u_1}{\partial x_2}(\hat{x}) - \frac{\partial u_2}{\partial x_1}(\hat{x}) \right)^2 \\
&+ \left(-\frac{\partial u_2}{\partial x_3}(\hat{x}) + \frac{\partial u_3}{\partial x_2}(\hat{x}) \right)^2 + \left(\frac{\partial u_3}{\partial x_4}(\hat{x}) - \frac{\partial u_4}{\partial x_3}(\hat{x}) \right)^2 = 0
\end{aligned}
$$

and the noise condition (14) does not hold.

Example E.2 *Consider the following 2-player game with strategic complements.*

$$u_1(x_1, x_2) = -\frac{x_1^2}{2} + 2x_1 - x_1(2 - x_2)^2; \quad u_2(x_1, x_2) = -\frac{x_2^2}{2} - x_1^2(2 - x_2).$$

This game has anti-symmetric externalities, since $\frac{\partial u_2}{\partial x_1}(x) = -\frac{\partial u_1}{\partial x_2}(x)$. Now, the profile $(1, 1)$ is a Nash equilibrium, and

$$\frac{\partial^2 u_i}{\partial x_i \partial x_j}(\hat{x}) = 2, \quad i = 1, 2.$$

As a consequence the Jacobian matrix associated to the dynamics F is simply

$$DF(\hat{x}) = \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix},$$

which eigenvalues are -3 and 1. Thus this Nash equilibrium is linearly unstable. The eigenspace associated to the positive eigenvalue is generated by $v = (1, 1)$. Thus, on the event $\{\lim_n x_n = \hat{x}\}$, we have

$$\lim_{n \to +\infty} \mathbb{E}\left(\langle U_{n+1}, v \rangle^2 \mid \mathcal{F}_n \right) = \left(\frac{\partial u_1}{\partial x_2}(\hat{x}) + \frac{\partial u_2}{\partial x_1}(\hat{x}) \right)^2 = 0$$

and the noise condition does not hold.

References

Arrow, K. J. and Hurwicz, L. (1960). Stability of the gradient process in n-person games. *Journal of the Society for Industrial and Applied Mathematics*, 8(2):280–294.

Babichenko, Y. (2012). Completely uncoupled dynamics and Nash equilibria. *Games and Economic Behavior*, 76(1):1–14.

Ballester, C., Calvó-Armengol, A., and Zenou, Y. (2006). Who's who in networks. Wanted: the key player. *Econometrica*, 74(5):1403–1417.

Beggs, A. (2005). On the convergence of reinforcement learning. *Journal of Economic Theory*, 122(1):1–36.

Benaïm, M. (1996). A dynamical system approach to stochastic approximations. *SIAM Journal on Control and Optimization*, 34:437.

Benaïm, M. (1999). Dynamics of stochastic approximation algorithms. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Mathematics*, pages 1–68. Springer, Berlin.

Benaïm, M. and Faure, M. (2012). Stochastic approximations, cooperative dynamics and supermodular games. *Annals of Applied Probability*, 22(5):2133–2164.

Benaïm, M. and Hirsch, M. (1999). Mixed equilibria and dynamical systems arising from fictitious play in perturbed games. *Games and Economic Behavior*, 29(1-2):36–72.

Benaïm, M., Hofbauer, J., and Sorin, S. (2005). Stochastic approximations and differential inclusions. *SIAM Journal on Control and Optimization*, 44:328–348.

Berger, U. (2005). Fictitious play in 2×n games. *Journal of Economic Theory*, 120(2):139–154.

Bervoets, S. and Faure, M. (2018). Best-reponse in public good games on networks. *Mimeo*.

Bomze, I. M. (1990). Dynamical aspects of evolutionary stability. *Monatshefte für Mathematik*, 110(3-4):189–206.

Börgers, T. and Sarin, R. (1997). Learning through reinforcement and replicator dynamics. *Journal of Economic Theory*, 77(1):1–14.

Bramoullé, Y. and Kranton, R. (2007). Public goods in networks. *Journal of Economic Theory*, 135(1):478–494.

Bramoullé, Y., Kranton, R., and D'amours, M. (2014). Strategic interaction and networks. *The American Economic Review*, 104(3):898–930.

Brandiere, O. and Duflo, M. (1996). Les algorithmes stochastiques contournent-ils les pièges? *Annales de l'I.H.P. Probabilités et Statistiques*, 32(3):395–427.

Bravo, M. (2016). An adjusted payoff-based procedure for normal form games. *Mathematics of Operations Research*, 41(4):1469–1483.

Bravo, M. and Faure, M. (2015). Reinforcement learning with restrictions on the action set. *SIAM Journal on Control and Optimization*, 53(1):287–312.

Brown, G. W. (1951). Iterative solution of games by fictitious play. In *Activity Analysis of Production and Allocation*, pages 374–376. Wiley, New York.

Cheung, M.-W. (2014). Pairwise comparison dynamics for games with continuous strategy space. *Journal of Economic Theory*, 153:344–375.

Cheung, M.-W. (2016). Imitative dynamics for games with continuous strategy space. *Games and Economic Behavior*, 99:206–223.

Cominetti, R., Melo, E., and Sorin, S. (2010). A payoff-based learning procedure and its application to traffic games. *Games and Economic Behavior*, 70(1):71–83.

Cressman, R. (2005). Stability of the replicator equation with continuous strategy space. *Mathematical Social Sciences*, 50(2):127–147.

Dindos, M. and Mezzetti, C. (2006). Better-reply dynamics and global convergence to Nash equilibrium in aggregative games. *Games and Economic Behavior*, 54(2):261–292.

Erev, I. and Roth, A. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, 88(4):848–881.

Foster, D. P. and Young, H. P. (2003). Learning, hypothesis testing, and Nash equilibrium. *Games and Economic Behavior*, 45(1):73–96.

Foster, D. P. and Young, H. P. (2006). Regret testing: Learning to play nash equilibrium without knowing you have an opponent. *Theoretical Economics*, 1(3):341–367.

Fudenberg, D. and Kreps, D. (1993). Learning mixed equilibria. *Games and Economic Behavior*, 5(3):320–367.

Fudenberg, D. and Tirole, J. (1991). Game theory. Technical report, The MIT press.

Germano, F. and Lugosi, G. (2007). Global Nash convergence of Foster and Young's regret testing. *Games and Economic Behavior*, 60(1):135–154.

Hart, S. and Mas-Colell, A. (2003). Uncoupled dynamics do not lead to Nash equilibrium. *The American Economic Review*, 93(5):1830–1836.

Hart, S. and Mas-Colell, A. (2006). Stochastic uncoupled dynamics and Nash equilibrium. *Games and Economic Behavior*, 57:286–303.

Hirsch, M. (1999). Chain transitive sets for smooth strongly monotone dynamical systems. *Dynamics of Continuous Discrete and Impulsive Systems*, 5:529–544.

Hofbauer, J., Oechssler, J., and Riedel, F. (2009). Brown–von Neumann–Nash dynamics: the continuous strategy case. *Games and Economic Behavior*, 65(2):406–429.

Hofbauer, J. and Sandholm, W. (2002). On the global convergence of stochastic fictitious play. *Econometrica*, 70(6):2265–2294.

Hopkins, E. and Posch, M. (2005). Attainability of boundary points under reinforcement learning. *Games and Economic Behavior*, 53(1):110–125.

Huck, S., Normann, H.-T., and Oechssler, J. (2004). Through trial and error to collusion. *International Economic Review*, 45(1):205–224.

Kakade, S. M. and Foster, D. P. (2008). Deterministic calibration and Nash equilibrium. *Journal of Computer and System Sciences*, 74(1):115–130.

Karandikar, R., Mookherjee, D., Ray, D., and Vega-Redondo, F. (1998). Evolving aspirations and cooperation. *Journal of Economic Theory*, 80(2):292–331.

Lahkar, R. and Riedel, F. (2015). The logit dynamic for games with continuous strategy sets. *Games and Economic Behavior*, 91:268–282.

Laslier, J., Topol, R., and Walliser, B. (2001). A behavioral learning process in games. *Games and Economic Behavior*, 37(2):340–366.

Leslie, D. S. and Collins, E. J. (2005). Individual Q-learning in normal form games. *SIAM Journal on Control and Optimization*, 44(2):495–514.

Marden, J. R., Young, H. P., Arslan, G., and Shamma, J. S. (2009). Payoff-based dynamics for multiplayer weakly acyclic games. *SIAM Journal on Control and Optimization*, 48(1):373–396.

Mertikopoulos, P. (2018). Learning in concave games with imperfect information. *Mimeo*.

Miyazawa, K. (1961). On the convergence of the learning process in a 2×2 non-zero-sum two-person game. Econometric Research Program, Princeton University. *Research Memorandum*.

Monderer, D. and Shapley, L. (1996). Fictitious Play Property for Games with Identical Interests. *Journal of Economic Theory*, 68(1):258–265.

Oechssler, J. and Riedel, F. (2001). Evolutionary dynamics on infinite strategy spaces. *Economic Theory*, 17(1):141–162.

Oechssler, J. and Riedel, F. (2002). On the dynamic foundation of evolutionary stability in continuous models. *Journal of Economic Theory*, 107(2):223–252.

Pemantle, R. (1990). Nonconvergence to unstable points in urn models and stochastic approximations. *Annals of Probability*, 18(2):698–712.

Perkins, S. and Leslie, D. S. (2014). Stochastic fictitious play with continuous action sets. *Journal of Economic Theory*, 152:179–213.

Posch, M. (1997). Cycling in a stochastic learning algorithm for normal form games. *Journal of Evolutionary Economics*, 7(2):193–207.

Robinson, J. (1951). An iterative method of solving a game. *Annals of Mathematics*, 54(2):296–301.

Rosen, J. B. (1965). Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica:*, pages 520–534.

Ruelle, D. (1981). Small random perturbations of dynamical systems and the definition of attractors. *Communications in Mathematical Physics*, 82(1):137–151.

Shapley, L. (1964). Some topics in two-person games. *Advances in Game Theory*, 52:1–28.

Vives, X. (1990). Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, 19(3):305–321.

Young, H. P. (2009). Learning by trial and error. *Games and Economic Behavior*, 65(2):626–643.