

# Learning with minimal information in continuous games\*

Sebastian Bervoets<sup>†</sup> Mario Bravo<sup>‡</sup> and Mathieu Faure<sup>§</sup>

January 30, 2020

## Abstract

While payoff-based learning models are almost exclusively devised for finite action games, where players can test every action, it is harder to design such learning processes for continuous games. We construct a stochastic learning rule, designed for games with continuous action sets, which requires no sophistication from the players and is simple to implement: players update their actions according to variations in own payoff between current and previous action. We then analyze its behavior in several classes of continuous games and show that convergence to a stable Nash equilibrium is guaranteed in all games with strategic complements as well as in concave games, while convergence to Nash occurs in all locally ordinal potential games as soon as Nash equilibria are isolated.

---

\*We wish to thank Y. Bramoullé, P. Mertikopoulos, N. Allouch, F. Deroïan, M. Belhaj, and participants of the various seminars where we presented this paper. The authors also thank the French National Research Agency (ANR) for their financial support through the ANR programs SOSOSO (13 JSH1 0009 01) and CIGNE (ANR-15-CE38-0007-01). Mario Bravo gratefully acknowledges the support provided by FONDECYT 11151003 and 1991924.

<sup>†</sup>Aix-Marseille Univ., CNRS, EHESS, Centrale Marseille, AMSE, Marseille, France. Email: sebastian.bervoets@univ-amu.fr

<sup>‡</sup>Universidad de Santiago de Chile, Departamento de Matemática y Ciencia de la Computación. Email: mario.bravo.g@usach.cl

<sup>§</sup>Aix-Marseille Univ., CNRS, EHESS, Centrale Marseille, AMSE, Marseille, France. Email: mathieu.faure@univ-amu.fr

Keywords: Payoff-based learning, continuous games, stochastic approximation.

JEL Codes: C6, C72, D83

## 1 Introduction

In this paper we construct a stochastic learning rule which is designed for games with continuous action sets, requires no sophistication from the players and is simple to implement. We analyze its behavior in several classes of continuous games, in particular to establish whether it converges to Nash equilibria or not.

The question of convergence to Nash equilibria by agents repeatedly playing a game 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 e.g. Hart and Mas-Colell [2003], Hart and Mas-Colell [2006], Foster and Young [2006], Germano and Lugosi [2007], Babichenko [2012]). 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. Being designed for finite games, classical learning models cannot be adapted without serious complications, because they usually rely on assigning a positive probability to each choice of action. It is actually not easy to construct learning rules for continuous games which do not require players to have access to substantial amounts of information. This is what we do in this paper, by introducing a learning rule which we call the *dampened gradient approximation process* (DGAP). We also 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). They also 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 are 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. In order to make sure that they remain in the state space, these movements are dampened as the actions get close to the boundary; hence the name of our learning rule.

The direction chosen at the exploration stage being random, the DGAP is a stochastic process. 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 related to some underlying *deterministic dynamical system*. We thus start by showing that our process is well-defined - i.e. players' actions always remain non-negative - and that the deterministic system underlying 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 zeroes - of this dynamical system, although other points, on the boundary of the state space, may also be stationary. However, we prove (Proposition 2) that non-Nash stationary points are necessarily unstable.<sup>1</sup> This is done in Section 2, where we present the DGAP and provide the necessary definitions.

Our objective is to design an analytically tractable payoff-based process for continuous games. This paper should thus be seen as a contribution to learning theory for cases so far unexplored. Therefore, we wish to analyze its asymptotic behavior in several continuous games. The major difficulty is that 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

---

<sup>1</sup>Throughout the paper, several notions of stability will be used. They are all defined in Section 2.

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.

We first prove a general result (Theorem 1) which says that if the process converges, it necessarily converges to a Nash equilibrium; additionally, under the condition that the interactions between players do not form a *bipartite* graph, this Nash equilibrium cannot be unstable.

Next, in Section 3 we analyze games with *strategic complements* and show (Theorem 2) that the DGAP almost surely converges to a Nash equilibrium, and that this Nash equilibrium is stable. To the best of our knowledge, this is the first paper to prove convergence of a payoff-based learning procedure in this class of games.

In Section 4, we analyze a class of games that we call *locally ordinal potential games* which contains all the potential games. We establish two results (Theorems 3 and 4). First, the limit set of the DGAP is almost surely 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. Second, we characterize the set of stable sets (attractors) by proving that they are stable sets for another, unrelated dynamical system: the Best-Response dynamics.

In Section 5, we focus on *concave games*, as defined by Rosen [1965]. In that paper it is shown that these games have a unique Nash equilibrium and that a gradient system converges to the unique Nash equilibrium. We obtain the same results for our process, with convergence to the unique equilibrium with probability one.

Finally, in Section 6 we discuss which properties of the learning process are critical for our results to hold and how it can be generalized in several directions.

### ***Related Literature***

As mentioned earlier, the learning literature has essentially focused on finite action games. Many rules have been proposed and studied, but they cannot be adapted to the context of continuous games without major complications (see for instance Perkins and Leslie [2014], who adapt stochastic fictitious play and show that it converges in 2-player zero-sum games).

The literature on continuous games is sparse. In the context of non payoff-based procedures, 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. Vives [1990] proves that the Best-Response dynamics converges to a Nash equilibrium in games with strategic complements for almost all initial conditions, while Benaïm et al. [2005] prove that it converges to a connected set of Nash equilibria in continuous potential games.

In the context of limited information, the literature is both sparse and very recent. To the best of our knowledge, our paper is the first to consider a payoff-based learning process in games with strategic complements. In potential games, Tatarenko [2018] considers a learning process in which agents pick an action according to a Gaussian probability distribution, the mean parameter of which is updated as payoffs are realized. She proves convergence of the mean parameter of the distribution to a Nash equilibrium of the game. In contrast, in our process it is the actions of players that are updated and we get convergence results for the actual sequence of actions. Mertikopoulos and Zhou [2019] analyze procedures in a situation where agents receive some noisy information about their payoffs' gradients, mainly in games that enjoy a property called variational stability. Among other results, the authors establish almost sure convergence to the (convex) set of equilibria. Recently, using a similar approach, Bravo et al. [2018] study a subclass of those games in the payoff-based setting, obtaining almost sure convergence to the (in this case unique) Nash equilibrium. It is worth noting that their procedure and the DGAP share the convergence result for the games described in Section 5.

Two papers also address payoff-based procedures in specific games. Dindos and Mezzetti [2006] consider a stochastic adjustment process called the better-reply, in a specific class of games called aggregative games. At each step, agents are sequentially picked to play a strategy chosen at random, while the other players do not 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. The authors show that this process converges to Nash when actions are either substitutes or complements around the equilibrium. Huck et al. [2004] consider a learning process - called trial and error - and analyze it in the Cournot oligopoly game. 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 converges to the joint-profit maximizing profile and not to the (unique) Nash equilibrium of the game.

## 2 The model

### 2.1 Definitions and Assumption

Let  $\mathcal{N} = \{1, \dots, N\}$  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  $X = \times_{i=1, \dots, N} X_i$ . We denote by  $\partial X$  the boundary of  $X$ , i.e.  $\partial X := \{x \in X; x_i = 0 \text{ for some } i \in \mathcal{N}\}$ . And we let  $\text{Int}(X) := X \setminus \partial X$  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  $u = (u_i)_{i=1, \dots, N}$  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:

**Assumption 1** *For any  $i$ , the payoff map  $u_i$  is assumed to be  $C^1$  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  $x_i > M(x_{-i})$ .*

Assumption 1 implies that best responses are unique and  $\text{BR}_i(x_{-i}) = M(x_{-i})$ . 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  $\lim_{x_i \rightarrow +\infty} \frac{\partial u_i}{\partial x_i}(x_i, x_{-i}) < 0$ .

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:

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

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

where  $\text{sgn}(a) = 0$  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  $\text{sgn}\left(\frac{\partial u_i}{\partial x_j}(x)\right) = \text{sgn}\left(\frac{\partial u_i}{\partial x_j}(x')\right)$  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  $x = (x_1, \dots, x_N)$  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 Assumption 2 is symmetric.

In what follows, some of our results will hold for every interaction graph, except for bipartite graphs. For the sake of completeness we recall the definition:

**Definition 1** *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 said to be non bipartite.*

Finally, we will deal with games where Nash equilibria (henceforth,  $NE$ ) are not necessarily isolated. We will therefore consider *connected components* of  $NE$ :

**Definition 2** *Let  $\Lambda$  be a compact connected subset of  $NE$  and let  $N^\delta(\Lambda) := \{y \in X : d(y, \Lambda) < \delta\}$ . We say that  $\Lambda$  is a connected component of  $NE$  if there exists  $\delta > 0$  such that  $N^\delta(\Lambda) \cap NE = \Lambda$ .*

Obviously, an isolated equilibrium is a (trivial) connected component. In what follows we will refer to connected components only when dealing with non-trivial connected components.

## 2.2 The Learning Process

We consider a payoff-based learning process in which agents construct a partial approximation of the gradient of their own payoff functions 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 present the simplest version of the DGAP, while in fact our results hold for a family of learning rules. We discuss what features are essential, why they are essential, and what can be generalized in Section 6.

### The Dampened Gradient Approximation Process (DGAP)

- 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  $n$ ,  $\epsilon_n^i$  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  $\Delta u_{n+1}^i$  is positive, player  $i$  follows the direction that he just explored, while he goes in the opposite direction when  $\Delta u_{n+1}^i$  is negative.<sup>2</sup>

- 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  $\{e_1, \dots, e_{2n+1}\}$ . 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  $(x_n)_n$ .

---

<sup>2</sup>Note that the payoff difference is multiplied by  $e_{2n}^i$ . This is how actions are dampened close to the boundary, where variations can only be small. This specific dampening method is just one example out of many possibilities, which we discuss in Section 6. Note also that if one wishes to extend our analysis to compact subsets of  $\mathbb{R}_+$  instead of  $\mathbb{R}_+$ , it would be necessary to dampen movements when approaching the upper boundary. For simplicity, in this paper we only consider the lower boundary.



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

**Proposition 1** *Assume  $x_0^i > 1$  for all  $i$ . Then the iterative process is such that  $x_n^i > 0$  for all  $i$ .<sup>3</sup> It can be written as*

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

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.  $\mathbb{E}(U_{n+1} | \mathcal{F}_n) = 0$ ),
- (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  $\frac{1}{n+1}$ . 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)

$$\dot{x} = F(x) \quad (2)$$

or, in our case:

$$\dot{x}_i = x_i \frac{\partial u_i}{\partial x_i}(x)$$

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).<sup>4</sup>

---

<sup>3</sup>The assumption that  $x_0^i > 1$  is made just for convenience. We could in fact assume  $x_0^i$  to be arbitrary, in which case we must make sure that players stay in the positive orthant after the first exploration stage. Actions could be negative after the first step ( $e_1 < 0$ ), only because the first step is large ( $n = 1$ , so  $1/n = 1$ ). 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 with a step small enough, for instance  $n \geq \min_i \{E(1/x_0^i)\} + 1$ , where  $E(a)$  is the integer part of  $a$ . In any case, this is innocuous for what we do and guarantees that  $e_1 > 0$ .

<sup>4</sup>Stochastic approximation theory (see Benaïm [1996] or Benaïm [1999] for instance) states that,

## 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*.<sup>5</sup>

**Definition 3 (Limit set of  $(x_n)_n$ )** *Given a realization of the random process, we denote the limit set of  $(x_n)_n$  by*

$$\mathcal{L}((x_n)_n) := \{x \in X; \exists \text{ a subsequence } x_{n_k} \text{ such that; } x_{n_k} \rightarrow x\}. \quad (3)$$

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 set of candidates for  $\mathcal{L}((x_n)_n)$  (see Benaïm [1996] and Benaïm [1999] for an exact characterization). This set notably includes the zeroes of  $F$  and the  $\omega$ -limit set of any point  $x$ :

**Definition 4 ( $\omega$ -limit set of  $\dot{x} = F(x)$ )** *Let  $x \in X$ . Let  $\varphi(x, t)$  denote the flow of  $F(\cdot)$ , i.e. the position of the solution of (2) with initial condition  $x$ , at time  $t$ .<sup>6</sup> Then, the  $\omega$ -limit set of  $x$  is given by*

$$\omega(x) := \{z \in X; \lim_{k \rightarrow \infty} \varphi(x, t_k) = z \text{ for some } t_k \rightarrow \infty\}$$

However, several difficulties remain: first, there might be other candidates that are not  $\omega$ -limit sets of the underlying ODE. For instance, any continuum of equilibria is a candidate. Moreover, this theory does not provide general criteria to exclude any of these candidates, nor to confirm that they are indeed equal to  $\mathcal{L}((x_n)_n)$ . For a

---

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.

<sup>5</sup>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.

<sup>6</sup>Notice that by the regularity assumption on  $u(\cdot)$ ,  $F$  satisfies the Cauchy-Lipschitz condition that guarantees that, for all  $x \in X$ ,  $\varphi$  is well-defined and unique. We consider the restriction of  $\varphi$  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.

given game, making precise statements thus requires to determine the entire set of candidates, and to be able to exclude those which are not relevant.

The stationary points of the dynamical system (2) are particular  $\omega$ -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:  $OZ = \{x : F(x) = 0 \text{ and } \exists i \text{ s.t. } x_i = 0, \frac{\partial u_i}{\partial x_i}(x) > 0\}$ .

We have the following partition of F:

$$Z = NE \cup OZ. \tag{4}$$

Note that  $\partial X$  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  $\dot{x} = F(x)$ . We now recall the definitions of some stability notions that we will use.

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:  $\lim_{t \rightarrow +\infty} \sup_{x \in W} \|\varphi(x, t) - \hat{x}\| = 0$ , 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  $\lambda \in \text{Sp}(DF(\hat{x}))$  - where  $DF(\hat{x})$  is the Jacobian matrix of  $F$  evaluated at  $\hat{x}$  and  $\text{Sp}(M)$  is the spectrum of matrix  $M$  - we have  $\text{Re}(\lambda) < 0$ , where  $\text{Re}(a)$  is the real-part of number  $a$ ; and  $\hat{x}$  is *linearly unstable* (denoted by  $\hat{x} \in Z^{LU}$ ) if there exists  $\lambda \in \text{Sp}(DF(\hat{x}))$  such that  $\text{Re}(\lambda) > 0$ . Note that if  $\hat{x}$  is hyperbolic (that is  $\text{Re}(\lambda) \neq 0$  for any  $\lambda \in \text{Sp}(DF(\hat{x}))$ ) then it is either linearly stable or linearly unstable. Note also that linear stability and instability are defined on  $\text{Int}(X)$  only, that is  $Z^{LS}, Z^{LU} \subset Z^{int} := \text{Int}(X) \cap Z$ . In our terminology a zero is *stable* if

either it is asymptotically stable or if it is interior and not linearly unstable: let  $Z^S := (Z^{int} \setminus Z^{LU}) \cup Z^{AS}$  be the set of stable zeroes.

We have the following inclusions:

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

**Proposition 2** *Stable stationary points are necessarily Nash equilibria:  $Z^S \subset NE$ .*

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. Other zeroes are therefore discarded as stable stationary points. In the remainder, we will denote linearly unstable Nash equilibria by  $NE^{LU}$ , and in view of Proposition 2, we will use the following notations:

$$NE^{LS} := Z^{LS}, NE^{AS} := Z^{AS} \text{ and } NE^S := Z^S$$

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]). Let  $S$  be a compact subset of  $\mathbb{R}^N$ . Then  $S$  is *invariant for the flow*  $\varphi$  if it remains in  $S$  forever from initial conditions in  $S$ , and every point in  $S$  is attainable at any given time from another point in  $S$ . Formally, *i*)  $\forall x \in S, \forall t \in \mathbb{R}, \varphi(x, t) \in S$  and *ii*)  $\forall y \in S, \forall t \in \mathbb{R}$ , there exists  $x \in S$  such that  $\varphi(x, t) = y$ .

**Definition 5** *Let  $S \subset X$  be invariant for the flow  $\varphi$ . 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:*

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

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

**Remark 1** *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  $\dot{x} = F(x)$ .*

In the next sections we will establish results on the behavior of the limit set of  $(x_n)_n$  for large classes of games. However, we can always discard convergence to undesirable zeroes of the dynamics:

**Theorem 1** *Let  $\mathcal{G}$  be a game satisfying Assumption 1 and 2. Then:*

(i)  $\mathbb{P}(\lim_n x_n = \hat{x}) > 0$  implies  $\hat{x} \in NE$

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

If the process converges to a point, then only stationary points (i.e. points in  $Z$ ) are candidates. Therefore, the proof of point i) consists in showing that if  $\hat{x} \in OZ$ , then  $\mathbb{P}(\lim_n x_n = \hat{x}) = 0$ . We use a probabilistic argument. 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.

Notice that (i) does not imply that if  $\hat{x} \in OZ$ , then  $\hat{x} \notin \mathcal{L}((x_n)_n)$ , since  $\mathcal{L}((x_n)_n)$  may include connected components of zeroes which contain  $\hat{x}$ . Indeed, 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, theorem 1 says that when it does converge to a point, it is necessarily to a Nash equilibrium.

The proof of point (ii) uses dynamical systems arguments. It involves several ideas which we describe here, since they will be useful in the next sections. First we discuss how we can discard convergence to unstable Nash equilibria; second we discuss the non-bipartiteness condition. The point  $\hat{x}$  is a linearly unstable Nash equilibrium if there is some direction (associated with the positive eigenvalue) in which the dynamical system (2), if pushed that way, will "escape". Yet, although linearly unstable,  $\hat{x}$  is still a stationary point. This is where the noise component  $U_{n+1}$  of the random process (1) plays an important role. While the deterministic system could get stuck at  $\hat{x}$ , the random noise will push the system in random directions around  $\hat{x}$  and eventually in the unstable direction, allowing the system to escape. The details are provided in the appendix with the proof of theorem 1.

However, for this to happen, a sufficient condition is that the random noise is able to push the system in *every* direction around  $\hat{x}$ . This is where the non-bipartiteness condition, together with Assumption 2 (symmetric externalities), comes into play. As we detail in the proof, the random noise can always push the system in any direction, say  $v$ , except if every pair of connected agents moves in opposite directions from each

other (i.e.  $\text{sgn}(v_i) = -\text{sgn}(v_j)$ ). When that happens, the noise vanishes and the system might get stuck. However, this can only occur when the interaction graph  $\mathbf{G}$ , evaluated at  $\hat{x}$ , has no odd cycles (i.e. bipartite graphs). It cannot happen as long as the graph has one odd cycle. To see this, take the case of three agents linked together in a triangle. It is not possible to construct a vector  $v$  such that for every pair of players  $i$  and  $j$ ,  $\text{sgn}(v_i) = -\text{sgn}(v_j)$ .

Note that theorem 1 does not say that the process will converge to an unstable Nash equilibrium if the graph is bipartite. However, we provide an example in the appendix (Example A.1) in which we show that the noise vanishes on a bipartite network (a pair) in a game with symmetric externalities.

We now turn to the analysis of several classes of games.

### 3 Strategic complements

In this section, to avoid unnecessary complexity we assume that the interaction graph is constant (i.e.  $\forall x, g_{ij}(x) = 1$  or  $\forall x, g_{ij}(x) = 0$ ) and connected.

**Definition 6** *A game  $\mathcal{G}$  is a game with strategic complements if payoff functions are such that for all  $i, j \in N$ ,  $g_{ij} = 1 \iff \frac{\partial^2 u_i}{\partial x_i \partial x_j}(x) > 0$ .*

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 Assumption 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. In order to study convergence of the DGAP, we need to consider all possible trajectories and, to the best of our knowledge, 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.

Because we are interested in the behavior of the DGAP, in particular if it allows for convergence to Nash, we distinguish two cases. Either the game has some Nash equilibrium on the boundary, or all equilibria are interior. In the first case, theorem 1 already guarantees that if the process converges to a point on the boundary, that point is a Nash equilibrium. However, when it does not converge to a point but to a set, there is not much we can predict, except that convergence to Nash is not à priori excluded.

The second case is trickier. If all equilibria are interior, and since the boundary always contains other zeroes which are stationary points (and thus natural candidates for the limit set of the process), it could be the case that the process goes to the boundary and therefore never goes to the set of Nash equilibria. In fact, we prove that this is not the case, and even more, we prove that convergence to Nash occurs with probability 1.

This result is difficult to obtain, because of the following: 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  $\partial X$ . We show that this cannot happen.

In order to place ourselves in this second case, we make the mild assumption that the origin is repulsive.

**Assumption 3** *For any agent  $i$ ,*

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

Because of strategic complementarities, a direct consequence of assumption 3 is that all Nash equilibria are interior, since  $\frac{\partial u_i}{\partial x_i}(0, x_{-i}) > \frac{\partial u_i}{\partial x_i}(0, 0)$ .

The following proposition proves that the process will stay away from the boundary in the long run.

**Proposition 3** *Under Assumption 3, there exists a  $\delta > 0$  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 no general result along these lines in the literature.

The proof of Proposition 3 is 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.

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

**Theorem 2** *Consider a game of strategic complements satisfying assumptions 1, 2 and 3, and assume the interaction graph is non-bipartite on  $\text{Int}(X)$ . Then 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 theorem guarantees that the learning process will not only converge to Nash in most cases, it will additionally converge to a stable equilibrium. This result is very tight. It is also positive, since the hypotheses of the theorem are verified for most common economic models we can think of. In cases where the interaction graph is bipartite, we cannot guarantee that the process will not converge to general unstable sets.<sup>7</sup>

Note that the graph being bipartite does not imply that the process will not converge to an element of  $NE^S$ . However, we provide (Examples B.1 and B.2) in

---

<sup>7</sup>Linearly unstable equilibria are unstable sets, but unstable sets also include much more complex structures.



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.

However, we can still exclude convergence to linearly unstable equilibria. If the interaction graph is non-bipartite, then theorem 1 applies. With strategic complements we show that the non-bipartiteness condition is not necessary.

**Proposition 4** *Consider a game of strategic complements satisfying assumptions 1, 2 and 3. The learning process  $(x_n)$  cannot converge to an unstable Nash equilibrium:*

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

## 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  $P : X \rightarrow \mathbb{R}$  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 7** *A game  $\mathcal{G}$  is a locally ordinal potential game (LOPG) if there is a differentiable function  $P : X \rightarrow \mathbb{R}$  such that*

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

The class of *LOPG* is large, in the sense that  $PG \subset OPG \subset LOPG$  when  $P$  is differentiable, and it contains many games of economic interest.

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.

**Theorem 3** *Let  $\mathcal{G}$  be an *LOPG* satisfying Assumption 1 and 2, and assume  $P$  is sufficiently regular.<sup>8</sup> Then*

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

---

<sup>8</sup>We will use Sard's theorem in the proof. This theorem requires, in our setting, that  $P$  is  $C^N$  where  $N$  is the number of agents. Usually, potential functions in economics are  $C^\infty$ .

(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  $\omega$ -limit sets of the dynamical system, which are non-zeros, can be discarded (point (i)). 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 (point (ii)), however, convergence to Nash is proved by the conjunction of the first point and point (i) of Theorem 1. The addition to Theorem 1 here is that we can guarantee that the process will converge, while in Theorem 1 it was an assumption. The last statement is then a direct consequence of point (ii) of Theorem 1.

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  $\mathbb{P}(\mathcal{L}(x_n)_n \subset A) > 0$  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 3, we get the following important implication: if a connected set  $\Lambda$  is an attractor for  $\dot{x} = F(x)$ , then  $\Lambda$  is a connected component of  $Z$ .

However, when focusing on *LOPGs*, 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 8** Let  $\text{BR} : X \rightarrow X, x \mapsto \text{BR}(x) := (\text{BR}_1(x_{-1}), \dots, \text{BR}_n(x_{-n}))$ . The continuous-time Best-Response dynamics (thereafter, *BRD*) is defined as:

$$\dot{x} = -x + \text{BR}(x) \tag{5}$$

**Definition 9** Let  $P$  be a smooth map and  $\Lambda$  be a connected component of  $Z$ , we say that  $\Lambda$  is a local maximum of  $P$  if

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

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

We then have

**Theorem 4** *Assume  $\mathcal{G}$  is an LOPG and let  $\Lambda$  be a connected set. Then the following statements are equivalent*

- (i)  $\Lambda$  is an attractor for  $\dot{x} = F(x)$
- (ii)  $\Lambda$  is a local maximum of  $P$
- (iii)  $\Lambda \subset NE$  and  $\Lambda$  is an attractor for the Best-Response dynamics  $\dot{x} = -x + BR(x)$ .

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 4 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 dynamical system, 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 Concave games

Arrow and Hurwicz [1960] and Rosen [1965] have analyzed similar dynamical systems in concave games. The first paper 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 a continuous-time gradient system.

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 a weighted gradient system 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 theirs. 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.

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<sup>9</sup>  $g_i(x, r) = r_i \frac{\partial u_i}{\partial x_i}$ . A game is *diagonally strictly concave* if

$$\exists r \in (\mathbb{R}_+^*)^N \mid \forall x^0 \neq x^1 \in X \text{ we have } \langle x^1 - x^0, g(x^0, r) \rangle + \langle x^0 - x^1, g(x^1, r) \rangle > 0 \quad (6)$$

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 *strictly concave-convex*.

**Remark 2** *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. A proof is provided in the appendix.*

When the strategy space  $X$  is unbounded as in this paper, there might be a unique Nash equilibrium, or none. In the next result, we show that in the later case, the process would go to infinity, while when the Nash equilibrium is unique, we show that

---

<sup>9</sup>The dynamical system  $\dot{x} = g(x, r)$  is a weighted gradient system, and is significantly different from the system (2)

the set of other zeroes is finite, and therefore isolated. By Theorem 1, this guarantees that our process converges to the Nash equilibrium with probability 1.

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

(i)  *$Z$  is a finite set*

(ii) *If  $\mathcal{G}$  has a Nash equilibrium  $\bar{x}$ , it is unique and*

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

(iii) *If  $\mathcal{G}$  has no Nash equilibrium, then*

$$\mathbb{P}(\limsup_n \|x_n\| = +\infty) = 1.$$

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 Theorem 1.

## 6 Generalization of the process

We discuss how the DGAP can be extended or generalized along several directions. The DGAP described in section 2.2 has the following characteristics:

C1- Agents all move simultaneously at every period

C2- At the exploration stage players explore upwards and downwards with equal probability

C3- The step size at period  $n$  is  $\frac{1}{n+1}$

C4- At the updating stage the payoff difference  $\Delta u_i^{n+1}$  is multiplied by  $e_i^n$

For our proofs to work, C1 can be totally relaxed, C3 and C4 can be modified and generalized, but not relaxed, while C2 is necessary.

## 6.1 C1

Simultaneous exploration is not necessary. We can instead assume that at each period any arbitrary subgroup of players experiments and updates. Formally, let  $(I(n))_n$  be a sequence of random variables taking values in  $\mathcal{P}(N)$ , the power set of  $N$ . Consider the following modified learning process:

- At the beginning of round  $n$ ,  $I(n)$ , a given subset of players is drawn.
- *Exploration stage*

$$e_{2n+1}^i = \begin{cases} e_n^i + \frac{1}{n+1}\epsilon_n^i & \text{if } i \in I(n) \\ e_{2n}^i & \text{if } i \notin I(n) \end{cases}$$

- *Updating stage*

$$e_{2n+2}^i = \begin{cases} e_{2n}^i + \epsilon_n^i \Delta u_{n+1}^i e_{2n}^i & \text{if } i \in I(n) \\ e_{2n}^i & \text{if } i \notin I(n) \end{cases}$$

A sufficient condition for our results to hold is that  $(I(n))_n$  is an i.i.d. sequence, such that the events  $\{i \in I(n)\}_{i \in \mathcal{N}}$  are mutually independent, and

$$\mathbb{P}(i \in I(n)) = p_i > 0.$$

The case presented in the paper satisfies these conditions, with  $\mathbb{P}(i \in I(n)) = 1$  for all  $i$ .

Now we show that all the processes with players' selection device satisfying these conditions share the same features as the DGAP. Pick  $i \in N$  and let  $\chi_n^i = 1$  if  $i \in I(n)$ , and 0 otherwise. We have

$$x_{n+1}^i - x_n^i = x_n^i \tilde{\epsilon}_n^i \left[ u_i \left( x_n^i + \frac{1}{n+1} \tilde{\epsilon}_n^i, x_n^{-i} + \frac{1}{n+1} \tilde{\epsilon}_n^{-i} \right) - u_i(x_n^i, x_n^{-i}) \right],$$

where  $\tilde{\epsilon}_n^i := \chi_n^i \epsilon_n^i$ . The first order development now gives

$$x_{n+1}^i - x_n^i = \tilde{\epsilon}_n^i x_n^i \left( \frac{1}{n+1} \tilde{\epsilon}_n^i \frac{\partial u_i}{\partial x^i}(x_n) + \sum_{j \neq i} \frac{1}{n+1} \tilde{\epsilon}_n^j \frac{\partial u_i}{\partial x^j}(x_n) + \mathcal{O}(1/n^2) \right).$$

We have (taking the right filtration  $(\mathcal{F}_n)_n$ ):

$$\mathbb{E} [\tilde{\epsilon}_n^i \tilde{\epsilon}_n^j | \mathcal{F}_n] = 0; \quad \mathbb{E} [(\tilde{\epsilon}_n^i)^2 | \mathcal{F}_n] = \mathbb{P}(i \in I(n) | \mathcal{F}_n)$$

As a consequence

$$\mathbb{E} [x_{n+1}^i - x_n^i | \mathcal{F}_n] = \frac{\mathbb{P}(i \in I(n) | \mathcal{F}_n)}{n+1} x_n^i \frac{\partial u_i}{\partial x^i}(x_n)$$

We get the following continuous-time dynamical system as the mean dynamics of our new system:

$$\dot{x} = F(x) \text{ where } F_i(x) = p_i x^i \frac{\partial u_i}{\partial x^i}(x_n)$$

which shares the same asymptotic behavior as the system analyzed in the paper.

## 6.2 C2

At the exploration stage we assume that players increase their actions by  $\frac{1}{n+1}$  with probability  $\frac{1}{2}$  and decrease them by  $\frac{1}{n+1}$  with probability  $\frac{1}{2}$ . What matters in fact is that  $\mathbb{E}(U_{n+1} | \mathcal{F}_n) = 0$  and  $\sup_n \mathbb{E}(U_{n+1}^2 | \mathcal{F}_n) < +\infty$  in order to write the stochastic difference equation as the sum of a deterministic term and a random component whose first order term is null in average with bounded variance, which is critical to use stochastic approximation methods. As can be seen from the proof of Proposition 1, any probability distribution satisfying  $\mathbb{E}(\epsilon_n^i) = 0$  and  $\sup_n \mathbb{E}((\epsilon_n^i)^2) < +\infty$  for all  $i$  would work.

## 6.3 C3

The amplitude with which explorations are made by players is given by the sequence  $(\frac{1}{n})_n$ . In fact, it could be any sequence  $(\alpha_n)_n$  such that

$$\sum_k \alpha_k = \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} \alpha_n = 0$$

which  $(\frac{1}{n})_n$  naturally satisfies.

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.

We could also consider a family of sequences, differing for each individual, as long as they go to zero "at the same rate"; i.e. for all pair of players  $(i, j)$ ,  $\lim_{n \rightarrow \infty} \alpha_n^i / \alpha_n^j = k_{ij}$  where  $k_{ij} > 0$  is a constant.

## 6.4 C4

The variation in payoffs after an exploration stage ( $\Delta u_{n+1}^i$ ) is multiplied by  $e_i^{2n} = x_i^n$  at the updating stage. This results in the updating stage being written

$$x_i^{n+1} - x_i^n = \epsilon_i^n \Delta u_{n+1}^i x_i^n \quad (7)$$

leading to the dynamical system  $\dot{x}_i = x_i \frac{\partial u_i}{\partial x_i}$ , where the  $x_i$  term comes from this multiplication by  $x_i^n$ . Multiplying by  $x_i$  prevents the players from playing 0, by dampening the variations in their actions as they get closer to 0.

It is not necessary that every agent smooths their behaviors in that way. What we need is that the variation in actions between two steps are multiplied by some  $f_i(x_i)$ , potentially different for every agent, with  $f_i$  is a weakly decreasing Lipschitz function, strictly decreasing around 0 and such that  $f_i(0) = 0$ . This is to prevent players from playing 0, in a smooth way.

Choosing  $f_i(x_i) = x_i$  for any  $i$  is one possibility, but there are many others. For instance,  $f_i(x_i) = \min\{s, x_i\}$  where  $s > 0$  is one, where the increments' amplitudes depend on the current action only when close to 0 but does not depend on the current action otherwise.

Regardless of the choice for  $f_i(\cdot)$ , this term implies that other zeroes are stationary points of the dynamical system, which in turn implies that we need to take care of the set  $OZ$  in every section of the paper.

If we did not dampen actions around 0, the discrete-time system would be instead

$$x_i^{n+1} - x_i^n = \epsilon_i^n \Delta u_{n+1}^i$$

and actions could become negative at the updating stage. Thus one needs to make precise what the algorithm does when this happens. With gradient-type algorithms, there are essentially two ways to deal with this. The first one is to set actions to 0 when they should have become negative (i.e. actions are defined as  $\max\{x_i, 0\}$ ). The second way is to define the algorithm in such a way that actions remain positive. We followed this second option.

There are two reasons for choosing this second option. First, in order to use stochastic approximation theory, it is necessary that the first order term of the noise is mean zero. If players played 0, they could not explore downwards and exploration would no longer be mean zero. The following example illustrates why this might cause serious problems.



Consider the two players game with payoffs

$$\begin{cases} u_1(x_1, x_2) &= x_1 - \frac{1}{2}x_1^2 - 2x_2 \\ u_2(x_1, x_2) &= x_2 - \frac{1}{2}x_2^2 - 2x_1 \end{cases}$$

The unique Nash equilibrium is  $(x_1, x_2) = (1, 1)$ . Assume that players get to  $(0, 0)$  at period  $n$ , where they can only explore upwards. At the exploration stage they play  $e_1^{2n+1} = e_2^{2n+1} = \frac{1}{n+1}$ , and get  $u_1(\frac{1}{n+1}, \frac{1}{n+1}) = u_2(\frac{1}{n+1}, \frac{1}{n+1}) < 0$ . Thus  $x_i^{n+1}$  is set to 0, and players are trapped at  $(0, 0)$ .

The multiplication by  $f(x_i)$  guarantees that players will not play 0 and, although this does not artificially prevent them from converging to  $(0, 0)$ , it will help avoid such situations. Indeed, in this example our process will converge to the unique Nash equilibrium  $(1, 1)$  with probability one, instead of remaining trapped at  $(0, 0)$ . To see why players will converge to  $(1, 1)$ , observe that this game is a potential game, with  $P(x_1, x_2) = x_1 - x_1^2/2 + x_2 - x_2^2/2$ . There are exactly 4 zeroes with  $(0, 0)$ ,  $(1, 0)$  and  $(0, 1)$  being other zeroes. Theorem 3 (ii) proves convergence to  $(1, 1)$ .

The second for this choice is that projecting non admissible actions onto the feasible set by setting actions to  $\max\{x_i, 0\}$  models a somewhat discontinuous behavior. There are smoother ways of preventing actions from leaving the feasible set, one of which is to multiply actions by  $x_i$  as we do in equation (7). Interestingly, we can point out a parallel between our choice of multiplying by  $x_i$  and some well-known methods from the convex optimization literature<sup>10</sup>, as we explain now.

The standard (projected) gradient method is defined as:

$$\begin{cases} y_{n+1} &= x_n + \alpha_t \nabla f(x_n) \\ x_{n+1} &= \Pi_K(y_{n+1}) \end{cases}$$

where  $f$  is the concave function which has to be maximized,  $x_n$  is the vector of actions at step  $n$ , constrained to belong to some closed convex set  $K$  ( $\mathbb{R}^+$  in our case). Here  $y_n$  can be interpreted as the vector of unconstrained actions (the actions that players would choose if they had no constraints),  $\alpha_n$  is the step size, and  $\Pi_K(\cdot)$  is the projection operator onto  $K$ . Another idea is to consider the dual averaging algorithm based on the well-known *Mirror Descent* algorithm,<sup>11</sup> where a primal-dual procedure is considered:

$$\begin{cases} y_{n+1} &= y_n + \alpha_n \nabla f(x_n) \\ x_{n+1} &= Q_K(y_{n+1}), \end{cases}$$

<sup>10</sup>For modern applications of these procedures, see for instance Hazan [2016].

<sup>11</sup>See Nesterov [2009] and Mertikopoulos and Zhou [2019] for recent applications in game theory.

with  $Q_K(y) = \operatorname{argmax}\{\langle y, x \rangle - h(x); x \in K\}$ , where  $h$  is a strongly convex regularizing function on  $K$ . The map  $Q_K$  is called a *mirror map*, since it mirrors dual variables (gradients) onto primal variables.

Although our learning process is not related to these optimization procedures, the parallel is the following: our discrete learning algorithm induces an algorithm which is asymptotically equivalent to a (multi-agent) dual averaging procedure, associated to an entropic-like regularizing function  $h$  on  $K = \mathbb{R}_+^N$ . So multiplying by  $x_i$  is “as if” we were pseudo-projecting actions in a specific way.

# Appendix

## A Proof of results of Section 2

### A.1 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  $n$ .

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

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

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

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

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  $\epsilon_n^i$  and  $\epsilon_n^j$  are independent, so that

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

Let us now show that the process is well-defined. Notice that Assumption 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. Thus, for  $n \geq 0$ ,

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

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

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

Thus,  $x_1^i \geq 0$  and

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

■

## A.2 Proof of Proposition 2

Pick  $\hat{x} \in OZ$  and assume without loss of generality that  $\hat{x}_1 = 0$  with  $\frac{\partial u_1}{\partial x_1}(\hat{x}) = 2\alpha > 0$ . Then there exists  $\epsilon > 0$  such that, for any  $x \in B(\hat{x}, \epsilon)$  we have  $\frac{\partial u_1}{\partial x_1}(x) \geq \alpha$ . This implies that

$$\dot{x}_1 \geq \alpha x_1, \quad \forall x \in B(\hat{x}, \epsilon).$$

By standard arguments, this implies that  $\hat{x}$  is not asymptotically stable. ■

## A.3 Proof of Theorem 1

**Proof of Theorem 1 (i).** Pick an  $\hat{x} \in OZ$  and let us fix  $i \in \{1, \dots, N\}$  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  $\{\sup_{n \rightarrow +\infty} \|x_n\| < +\infty\}$  since, otherwise, there is nothing to prove.

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

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

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  $\mathbb{E}(\tilde{U}_{n+1}^i | \mathcal{F}_n) = 0$ , 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. ■

**Proof of Theorem 1 (ii).** Let us first 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  $\mu$  with positive real part, and call  $w$  the associated normalized eigenvector. We use a result of Pemantle [1990] for the case of isolated stationary points, adapted by Brandiere and Duflo [1996] to the case of connected components of stationary points. This result states that a sufficient condition for non convergence to  $\tilde{x}$  is that the noise is exciting in the unstable direction, i.e.:

$$\liminf_{n \rightarrow +\infty} \mathbb{E} (\langle U_{n+1}, w \rangle^2 \mid \mathcal{F}_n) > 0 \quad (8)$$

on the event  $\{\lim_n x_n = \tilde{x}\}$ . A sufficient condition for (8) to hold is

$$\mathbb{E} (\langle U_{n+1}, v \rangle^2 \mid x_n = \tilde{x}) = 0 \text{ if and only if } v = 0 \quad (9)$$

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

$$\mathbb{E} (\langle U_{n+1}, v \rangle^2 \mid x_n) = \sum_{i < j} \left( v_i x_n^i \frac{\partial u_i}{\partial x^j}(x_n) + v_j x_n^j \frac{\partial u_j}{\partial x^i}(x_n) \right)^2 \quad (10)$$

By equation (10), we see that  $\mathbb{E} (\langle U_{n+1}, v \rangle^2 \mid x_n = \tilde{x}) = 0$  if and only if

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

We now prove that under the assumption of symmetric externalities 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  $g_{ij} = g_{ik} = g_{jk} = 1$  and

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

We thus have  $\text{sgn}(v_i) = -\text{sgn}(v_j) = \text{sgn}(v_k) = -\text{sgn}(v_i)$  which implies 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, we must have  $v = 0$ , which concludes the proof. ■

The next examples illustrates the fact that things can go wrong (meaning that the noise condition might not hold) in games with symmetric externalities, when the graph is bipartite.

**Example A.1** Consider the following 2-player game

$$u_1(x) = -\frac{1}{2}x_1^2 - 2x_1x_2 + 3x_1$$

$$u_2(x) = -\frac{1}{2}x_2^2 - 2x_1x_2 + 3x_2$$

This game has symmetric externalities and its interaction graph is a pair, which is a bipartite graph.

One can check that the profile  $\hat{x} = (1, 1)$  is a Nash equilibrium. Recalling that  $F_i(x) = x_i \frac{\partial u_i}{\partial x_i}$ , the Jacobian matrix associated to  $\hat{x}$  is

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

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

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

and the noise condition (8) does not hold.

## 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

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

Since the game has strategic complements,

$$\frac{\partial u_i}{\partial x_i}(x_i, x_{-i}) > \alpha_i > 0, \forall x_i < \bar{x}_i, \forall x_{-i} \in X_{-i}. \quad (11)$$

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  $\partial X$ . Thus, by the aforementioned result of Benaïm (1999), we can conclude that

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

In what follows we will show that  $\mathbb{P}(\mathcal{L}((x_n)_n) \subset \partial X) = 0$ . 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  $\partial X$ ).

**Remark 3** *Three simple observations are in order.*

- (i) *Condition (11) 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  $i \in \{1, \dots, N\}$ ,  $\frac{\partial u_i}{\partial x^i}(x) \leq 0$  then  $x \in [a, +\infty[^N$ .*
- (iii) *Moreover, also from (11), 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  $\partial X$ , 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  $\partial X$  can be written as an increasing union of the form:

$$\partial X = \bigcup_{R=1}^{\infty} (\partial X \cap \mathcal{U}_R), \quad (12)$$

In order to show that  $\mathbb{P}(\mathcal{L}((x_n)_n) \subset \partial X) = 0$ , it is sufficient to prove that  $\mathbb{P}(\mathcal{L}((x_n)_n) \subset \partial X \cap \mathcal{U}_R) = 0$ , for all  $R > a$ . By contradiction, assume that there exists  $R > a$  such that

$\mathbb{P}(\mathcal{L}((x_n)_n) \subset \partial X \cap \mathcal{U}_R) > 0$  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

$$\mathbb{P}(\{\mathcal{L}((x_n)_n) \subset \partial X\} \cap \{x_n \in V_\varepsilon \cap \mathcal{U}_R, \text{ for all } n \geq n_*\}) > 0, \quad (13)$$

where

$$V_\varepsilon = \{x \in X; d(x, \partial X) \leq \varepsilon\}.$$

In what follows, we work on the event  $E$  defined by (13) 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  $\Phi$  verifies the well-known relation

$$\min_{i=1, \dots, N} x^i - \frac{\ln(N)}{\beta} \leq \Phi_\beta(x) \leq \min_{i=1, \dots, N} x^i. \quad (14)$$

From a straightforward calculation we have that, for all  $i \in \{1, \dots, N\}$ ,

$$\frac{\partial \Phi_\beta}{\partial x^i}(x) = \pi_i(x), \text{ where } \pi_i(x) = \frac{\exp(-\beta x^i)}{\sum_{j=1}^N \exp(-\beta x^j)}.$$

Also, for all  $i, j \in \{1, \dots, N\}$

$$\frac{\partial^2 \Phi_\beta}{\partial x^j \partial x^i}(x) = -\beta \pi_i(x) (\delta_{ij} - \pi_j(x)),$$

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

$$\pi_i(x) \leq \exp(-\beta(a - \varepsilon)),$$

using the fact that there exist some  $k$  such that  $x^k \leq \varepsilon$  and that  $x^i \geq a$  (c.f. Remark 3).

On the other hand, for  $k \in I_x$  such that  $x^k = \min_i x^i$ ,

$$\pi_k(x) = \frac{1}{1 + \sum_{j \neq k} \exp(\beta x^k - \beta x^j)} \geq \frac{1}{N}.$$

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  $\mathbb{E}(U_{n+1} | \mathcal{F}_n) = 0$  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 (14),

$$-\frac{\ln(N)}{\beta} \leq \min_{i=1,\dots,N} x_n^i - \frac{\ln(N)}{\beta} \leq z_n \leq \min_{i=1,\dots,N} x_n^i \leq \epsilon. \quad (15)$$

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,

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

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

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

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

$$\begin{aligned} \mathbb{E}(z_{n+1} | \mathcal{F}_n) &\geq z_n + \frac{1}{n+1} \sum_{j \in I_{x_n}} \pi_j(x_n) x_n^j \frac{\partial u_i}{\partial x_j}(x_n) + \frac{1}{n+1} \sum_{j \in J_{x_n}} \pi_j(x_n) x_n^j \frac{\partial u_i}{\partial x_j}(x_n), \\ &\geq z_n + \frac{z_n}{n+1} \frac{\alpha}{N} + \frac{1}{n+1} \sum_{j \in J_{x_n}} \pi_j(x_n) x_n^j \frac{\partial u_i}{\partial x_j}(x_n), \end{aligned}$$

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

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

Thus

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

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,

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

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

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

Let us call  $\rho_n = \mathbb{E}(\theta_n) - c(\beta)$ . Then, we want to analyse the recursion  $\rho_{n+1} \geq \rho_n(1 + \frac{1}{n+1}\alpha/N)$ , 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  $E$ . ■

## B.2 Proofs of Theorem 2 and Proposition 4

The proof of Theorem 2 involves several arguments. 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, and we use a result from 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*

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

where

- (i)  $F : X \rightarrow \mathbb{R}^N$  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

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

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 2.** We want to apply Theorem B.1. When the game has strategic complements, our dynamics  $\dot{x} = F(x)$  is *cooperative* because all non-diagonal entries of  $DF(x)$  are nonnegative. In addition, the interaction graph is connected, and assumption 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 externalities, then (ii) holds.

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

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

if and only if  $v = 0$ . We proved this in the proof of point (ii) of Theorem 1. This concludes the proof. ■

**Proof of Proposition 4.** 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 to the unstable direction of  $\tilde{x}$ , for the strictly positive eigenvalue  $\mu$ .

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 point (ii) of Theorem 1. 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

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

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 Assumption 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  $\mu$  is strictly positive and  $v$  is the corresponding normalized eigenvector, we should have  $\langle v DF(\tilde{x}), v \rangle = \mu \sum_i v_i^2 > 0$ , 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,

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

a contradiction. To see why this inequality holds, remember that the terms in the first sum are all negative by Assumption 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. ■

The following example illustrates that the symmetric externalities assumption cannot be removed in Proposition 4.

**Example B.1** 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

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

and the noise condition does not hold.

The following example illustrates that strategic complements are essential in our proof.

**Example B.2** Consider the following 4-player example with strategic substitutes.

$$\begin{aligned} u_1(x) &= -cx_1 + b(x_1 + x_2 + x_4), & u_2(x) &= -cx_2 + b(x_2 + x_1 + x_3), \\ u_3(x) &= -cx_3 + b(x_3 + x_2 + x_4), & u_4(x) &= -cx_4 + b(x_4 + x_1 + x_3), \end{aligned}$$

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.

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

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

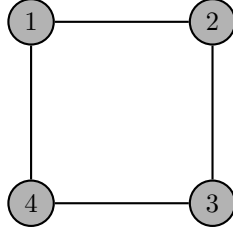


Figure 1: Interaction graph between 4 players. This graph is bipartite.

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

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

and the noise condition (8) does not hold.

## C Proof of results of Section 4.

### C.1 Proof of Theorem 3

Before proving Theorem 3, let us define the following dynamical concept:

**Definition C.1** Let  $P : X \rightarrow \mathbb{R}$  be continuously differentiable. We say that  $P$  is a strict<sup>12</sup> 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  $\dot{x} = -x + \text{BR}(x)$ .

<sup>12</sup>Generally,  $P$  is a Lyapunov function for  $\dot{x} = F(x)$  with respect to  $\Lambda$  if  $t \mapsto P(\varphi(x, t))$  is constant on  $\Lambda$  and strictly increasing for  $x \notin \Lambda$ ; when the component  $\Lambda$  coincides with the set of stationary points of the flow, then we say that  $P$  is strict.

(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

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

We need to check that, if  $x \notin NE$ , then this quantity is positive. Let  $i$  be such that  $x_i \neq \text{BR}_i(x)$ , say  $x_i < \text{BR}_i(x)$ . 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  $\langle DP(x), -x + \text{BR}(x) \rangle > 0$ .

(ii) We have

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

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. ■

**Lemma C.2** *Assume  $\mathcal{G}$  is an LOPG. Then, if  $P$  is  $\mathcal{C}^m$  for sufficiently large  $m$  then  $P(Z)$  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  $P^A : [0, 1]^A \rightarrow \mathbb{R}$  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 differentiable, 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 3 (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. ■

**Proof of Theorem 3 (ii).** By Lemma C.1 and Corollary 6.6 in Benaïm [1999], we have

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

Because convergence to the zeroes occurs almost surely, Theorem 1 gives us the result. ■

## C.2 Proof of Theorem 4

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  $\Lambda$ . Let  $v := P(\Lambda)$ . If  $\Lambda$  is not a local maximum of  $P$  then there exists a sequence  $x_n$  such that  $d(x_n, \Lambda) \rightarrow_n 0$  and  $P(x_n) > v$ . Since  $\Lambda$  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  $d(\varphi(x_n, t), \Lambda) \not\rightarrow 0$  and  $\Lambda$  is not an attractor.

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

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



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  $\Lambda$  is an attractor for the Best-Response dynamics.  $P$  is a strict Lyapunov function for the best-response dynamics<sup>13</sup> 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  $\Lambda$  is actually a strict local maximum of  $P$ : there exists an open (isolating) neighborhood  $U$  of  $\Lambda$  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<sup>14</sup>. This implies that  $\Lambda = \cap_{r>0} V_r$  contains an attractor  $A$ . The potential being constant on  $\Lambda$ ,  $A$  cannot be strictly contained in  $\Lambda$  and therefore  $\Lambda$  is an attractor.

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

## D Proof of results of Section 5

**Proof of Remark 2:** 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  $\mathcal{G}^{\text{Ros}}$  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  $\mathcal{G}^{\text{Arr}}$  are such that  $G(x, \mathbf{1}) + G'(x, \mathbf{1})$  is negative definite. Define the matrices  $A, B^k$  and  $C$  as follows:

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

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

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

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. ■

<sup>13</sup>Keep in mind that this means that it is a lyapunov function with respect to NE.

<sup>14</sup>We need to make sure that  $r$  is small enough so that  $\overline{V_r} = P^{-1}([v - r, v]) \subset U$

**Proof of Theorem 5:** Suppose first that there is a unique Nash equilibrium. Then note that under (6) we have, for any  $x \neq \bar{x}$ ,

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

because

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

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  $u_i^J(z) := u_i(z, 0^{|J|})$ .

**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  $J \subset I(\bar{x})$  then  $\tilde{x}^J = \bar{x}$ . 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  $J \subset I(\bar{x})$ . Then  $\bar{x}$  satisfies (i). Moreover for all  $i \notin J$ ,

$$\bar{x}_i = BR_i(\bar{x}_{-i}) = \operatorname{argmax}_{x_i} u_i(x_i, \bar{x}_{-i}) = \operatorname{argmax}_{z_i} u^J(z_i, \bar{z}_{-i}),$$

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

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

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

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

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

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

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

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

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

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

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

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

Now suppose that  $x \in \tilde{X}^k$  for a given  $k \in \{1, \dots, K\}$ . Note that  $I(x) \subset I(\tilde{x}^k)$ . We can then define  $\Phi^k : \tilde{X}^k \rightarrow \mathbb{R}$  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  $\Phi^k$  is again concave, with unique maximum in  $x = \tilde{x}^k$  on  $\tilde{X}^k$ . Let  $\phi(t) = \Phi^k(\mathbf{x}(t))$ , where  $\mathbf{x}(t)$  is a solution of  $\dot{x} = F(x)$ , with  $\mathbf{x}(0) \in \tilde{X}^k$ . We have

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

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

As a consequence every solution curve converges to a zero of  $F$ , i.e either  $\bar{x}$  or one of the  $\tilde{x}^k$ <sup>15</sup> More precisely,  $\bar{X}$  and  $\tilde{X}^k$  are invariant and  $\{\bar{x}\}$  (resp.  $\{\tilde{x}^k\}$ ) is a global

---

<sup>15</sup>This guarantees that the deterministic system is dissipative, i.e. admits a global attractor, which in turns shows that  $(x_n)_n$  is almost surely bounded. However, it is not enough to guarantee that our random process converges with probability one to one of the zeroes of the dynamics.

attractor for the flow  $\phi|_{\bar{X}}$  (resp.  $\phi|_{\tilde{X}^k}$ ); in particular, for any  $x_0 \in \bar{X}$  (resp.  $x_0 \in \tilde{X}^k$ ) then  $\lim_{t \rightarrow +\infty} \phi_t(x_0) = \bar{x}$  (resp.  $\lim_{t \rightarrow +\infty} \phi_t(x_0) = \tilde{x}^k$ ).

A set  $L$  is *internally chain transitive (ICT)* for the flow  $\phi_t$  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  $\bar{L} := \{\bar{x}\}$  are ICT.

**Theorem D.1 (Benaim, 1999)** *On the event  $\{\limsup_n \|x_n\| < +\infty\}$ , 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  $\bar{L}$  are the only internally chain transitive sets. This will conclude the proof because as we mentioned above  $(x_n)_n$  is almost surely bounded. Note that  $\bar{X}$  is an open set in  $X$ . To do so we first claim that it is always possible to relabel the family  $(\tilde{x}^k)_{k=1, \dots, K}$  such that  $\tilde{X}^k$  is an open set of  $\cup_{l=1}^k X^l$  for  $k = 2, \dots, K$ .

Let  $L$  be internally chain transitive. By previous result, if  $L$  intersects  $\bar{X}$  then  $L \subset \{\bar{x}\}$  because  $\bar{X}$  is the basin of attraction of  $\bar{x}$ . 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  $L = \{\bar{x}\}$  or  $L = \{\tilde{x}^k\}$  for some  $k$ .

Suppose now that there is no Nash equilibrium and assume by contradiction that  $\mathbb{P}(\limsup_n \|x_n\| < +\infty) > 0$ . By the same reasoning as above the only ICT sets are zeroes of the dynamics. Since there is no Nash equilibrium, on the event  $\{\limsup_n \|x_n\| < +\infty\}$ , we necessarily have  $\lim_n x_n = \hat{x}$ , where  $\hat{x}$  is some other zero (Theorem D.1 (Benaim, 1999)). This is a contradiction with Theorem 1. ■

## References

- Kenneth J Arrow and Leonid Hurwicz. Stability of the gradient process in n-person games. *Journal of the Society for Industrial and Applied Mathematics*, 8(2):280–294, 1960.
- Yakov Babichenko. Completely uncoupled dynamics and Nash equilibria. *Games and Economic Behavior*, 76(1):1–14, 2012.

- Michel Benaïm. A dynamical system approach to stochastic approximations. *SIAM Journal on Control and Optimization*, 34:437, 1996.
- Michel Benaïm. Dynamics of stochastic approximation algorithms. In *Séminaire de Probabilités, XXXIII*, volume 1709 of *Lecture Notes in Mathematics*, pages 1–68. Springer, Berlin, 1999.
- Michel Benaïm and Mathieu Faure. Stochastic approximations, cooperative dynamics and supermodular games. *Annals of Applied Probability*, 22(5):2133–2164, 2012.
- Michel Benaïm, Josef Hofbauer, and Sylvain Sorin. Stochastic approximations and differential inclusions. *SIAM Journal on Control and Optimization*, 44:328–348, 2005.
- Odile Brandiere and Marie Dufflo. Les algorithmes stochastiques contournent-ils les pièges? *Annales de l’I.H.P. Probabilités et Statistiques*, 32(3):395–427, 1996.
- Mario Bravo, David Leslie, and Panayotis Mertikopoulos. Bandit learning in concave n-person games. In *Advances in Neural Information Processing Systems*, pages 5661–5671, 2018.
- Martin Dindos and Claudio Mezzetti. Better-reply dynamics and global convergence to Nash equilibrium in aggregative games. *Games and Economic Behavior*, 54(2):261–292, 2006.
- Dean P Foster and H Peyton Young. Regret testing: Learning to play nash equilibrium without knowing you have an opponent. *Theoretical Economics*, 1(3):341–367, 2006.
- Fabrizio Germano and Gabor Lugosi. Global Nash convergence of Foster and Young’s regret testing. *Games and Economic Behavior*, 60(1):135–154, 2007.
- Sergiu Hart and Andreu Mas-Colell. Uncoupled dynamics do not lead to Nash equilibrium. *The American Economic Review*, 93(5):1830–1836, 2003.
- Sergiu Hart and Andreu Mas-Colell. Stochastic uncoupled dynamics and Nash equilibrium. *Games and Economic Behavior*, 57:286–303, 2006.
- Elad Hazan. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016.

- Steffen Huck, Hans-Theo Normann, and Jörg Oechssler. Through trial and error to collusion. *International Economic Review*, 45(1):205–224, 2004.
- Panayotis Mertikopoulos and Zhengyuan Zhou. Learning in games with continuous action spaces and unknown payoff functions. *Mathematical Programming, Series A*, pages 1–36, 2019.
- Yakov Nesterov. Primal-dual subgradient methods for convex problems. *Mathematical Programming*, 120(1):221–259, 2009.
- Robin Pemantle. Nonconvergence to unstable points in urn models and stochastic approximations. *Annals of Probability*, 18(2):698–712, 1990.
- Steven Perkins and David S Leslie. Stochastic fictitious play with continuous action sets. *Journal of Economic Theory*, 152:179–213, 2014.
- Martin Posch. Cycling in a stochastic learning algorithm for normal form games. *Journal of Evolutionary Economics*, 7(2):193–207, 1997.
- J Ben Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*., pages 520–534, 1965.
- David Ruelle. Small random perturbations of dynamical systems and the definition of attractors. *Communications in Mathematical Physics*, 82(1):137–151, 1981.
- Tatiana Tatarenko. Stochastic learning in potential games: Communication and payoff-based approaches. *Mimeo*, 04 2018.
- Xavier Vives. Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, 19(3):305–321, 1990.