Theoretical Economics 15 (2020), 1471–1508 1555-7561/20201471

Learning with minimal information in continuous games

Sebastian Bervoets
Aix-Marseille Univ, CNRS, and AMSE

Mario Bravo
Departamento de Matemática y Ciencia de la Computación, Universidad de Santiago de Chile

Mathieu Faure
Aix-Marseille Univ, CNRS, and AMSE

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 equilibrium occurs in all locally ordinal potential games as soon as Nash equilibria are isolated.

Keywords. Payoff-based learning, continuous games, stochastic approximation.
JEL classification. C6, C72, D83.

1. Introduction

In this paper we construct a stochastic learning rule that 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.

Sebastian Bervoets: sebastian.bervoets@univ-amu.fr
Mario Bravo: mario.bravo.g@usach.cl
Mathieu Faure: mathieu.faure@univ-amu.fr
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. This work was supported by French National Research Agency Grants ANR-17-EURE-0020. The authors also thank the French National Research Agency for their financial support through the programs SOSOSO (Grant ANR-13-JSH1-0009-01) and CIGNE (Grant ANR-15-CE38-0007-01). Mario Bravo gratefully acknowledges the support provided by FONDECYT Grants 11151003 and 1991924.

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—that 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, and time allocation, are nonnegative 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 that 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 that 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 or 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 to maximize their payoffs by choosing an action. If players know 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. So as to ensure 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 nonnegative) 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.[1] 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, 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) that 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* that 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 1. 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 1.

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 two-player zero-sum games).

[1]Throughout the paper, several notions of stability are used. They are all defined in Section 2.

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 Benaim 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 equilibrium 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 ∂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 or about the set of opponents. In this paper, we examine several classes of underlying games, each class being defined by different properties on the functions u_i. However, we 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 (BR) 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 $\lim_{x_i \to +\infty} \frac{\partial u_i}{\partial x_i}(x_i, x_{-i}) < 0$.

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

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 $\text{sgn}(\frac{\partial u_i}{\partial x_j}(x)) = \text{sgn}(\frac{\partial u_i}{\partial x_j}(x'))$ for $x \neq x'$.

Some of our results 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 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

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

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

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

Obviously, an isolated equilibrium is a (trivial) connected component. In subsequent text, we refer to connected components only when dealing with nontrivial 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 choose 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. In Section 6, we discuss what features are essential, why they are essential, and what can be generalized.

The Dampened Gradient Approximation Process.

- 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 independent asnd identically distributed (i.i.d.) random variables such that $\mathbb{P}(\epsilon_n^i = 1) = \mathbb{P}(\epsilon_n^i = -1) = 1/2$. At period n, ϵ_n^i is drawn and player i plays

$$e_{2n+1}^i := e_{2n}^i + \frac{1}{n+1}\epsilon_n^i.$$

- *Updating stage.* Player i observes his new payoff and computes

$$\Delta u_{n+1}^i := u_i\big(e_{2n+1}^i, e_{2n+1}^{-i}\big) - u_i\big(e_{2n}^i, e_{2n}^{-i}\big).$$

  This quantity provides i with an approximation of his payoff function's gradient. Using this information, player i updates his action by playing

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

Thus, when Δu^i_{n+1} is positive, player i follows the direction that he just explored, while he goes in the opposite direction when Δu^i_{n+1} is negative.[2]

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

Let $x_n = e_{2n}$ and let $\mathcal{F}_n$ 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 $(x_n)_n$.

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 $x^i_0 > 1$ for all i. Then the iterative process is such that $x^i_n > 0$ for all i.*[3] *It can be written as*

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

where

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

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

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

All our proofs are provided 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 be written

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

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

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

[2]Note that the payoff difference is multiplied by e^i_{2n}. 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 consider only the lower boundary.

[3]The assumption that $x^i_0 > 1$ is made just for convenience. We could, in fact, assume x^i_0 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 $1/n = 1$). To avoid that, we can either assume that $x^i_0 > 1$ (i.e., players start far enough from the boundary) or that the process begins with a step small enough, for instance, $n \geq \min_i\{E(1/x^i_0)\} + 1$, where $E(a)$ is the integer part of a. In any case, this is innocuous for what we do and guarantees that $e_1 > 0$.

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) informs us on the asymptotic behavior of (1).[4]

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*.[5]

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

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

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 1999 for an exact characterization). This set notably includes the zeroes of F and the ω-limit set of any point x.

Definition 4 (ω-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.[6] Then the ω-limit set of x is given by

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

However, several difficulties remain: first, there might be other candidates that are not ω-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 or to confirm that they are indeed equal to $\mathcal{L}((x_n)_n)$. For a given game, making precise statements thus requires determining the entire set of candidates and the ability to exclude those that are not relevant.

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

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

[6] Notice that by the regularity assumption on $u(\cdot)$, F satisfies the Cauchy–Lipschitz condition that guarantees that, for all $x \in X$, φ is well defined and unique. We consider the restriction of φ on $X (= \mathbb{R}_+^N)$, since $X = \mathbb{R}_+^N$ is invariant for its flow, and our random process (1) always remains in the positive orthant.

The stationary points of the dynamical system (2) are particular ω-limit sets that are of interest to us, as they contain all the Nash equilibria of the underlying game. The set of stationary points, denoted $Z(F)$, is 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,

$$x \in Z \quad \Longleftrightarrow \quad \forall i \in \mathcal{N}, \quad \left(x_i = 0 \text{ or } \frac{\partial u_i}{\partial x_i}(x_i, x_{-i}) = 0\right)$$

while $x \in \text{NE} \Longleftrightarrow \forall i \in \mathcal{N}(\frac{\partial u_i}{\partial x_i}(x_i, x_{-i}) = 0,$ or $x_i = 0$ and $\frac{\partial u_i}{\partial x_i}(x_i, x_{-i}) \leq 0)$. 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 $x \in Z \setminus \text{NE}$ an *other zero* (OZ) of the dynamical system: $\text{OZ} = \{x : F(x) = 0 \text{ and } \exists i \text{ such that} x_i = 0, \frac{\partial u_i}{\partial x_i}(x) > 0\}$.

We have the following partition of F:

$$Z = \text{NE} \cup \text{OZ}.$$

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

Convergence or non-convergence of our random process to a given point or set sometimes depends 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 use.

Let $\hat{x} \in Z$. The point $\hat{x}$ is *asymptotically stable* (denoted by $\hat{x} \in Z^{\text{AS}}$) if it uniformly attracts an open neighborhood W of itself: $\lim_{t\to+\infty} \sup_{x\in W} \|\varphi(x,t) - \hat{x}\| = 0$, where $\varphi(x,t)$ denotes the flow of $F(\cdot)$. The point $\hat{x}$ is *linearly stable* (denoted by $\hat{x} \in Z^{\text{LS}}$) 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; the point $\hat{x}$ is *linearly unstable* (denoted by $\hat{x} \in Z^{\text{LU}}$) 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, $Z^{\text{LS}}, Z^{\text{LU}} \subset Z^{\text{int}} := \text{Int}(X) \cap Z$. In our terminology a zero is *stable* if either it is asymptotically stable or it is interior and not linearly unstable: let $Z^{\text{S}} := (Z^{\text{int}} \setminus Z^{\text{LU}}) \cup Z^{\text{AS}}$ be the set of stable zeroes.

We have the inclusions:

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

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

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 denote linearly unstable Nash equilibria by NE^{LU}, and in view of Proposition 2, we use the notations

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

As mentioned earlier, we sometimes are dealing with connected components of NE instead of isolated points. We, 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* φ 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) $\forall x \in \mathrm{S}$, $\forall t \in \mathbb{R}$, $\varphi(x, t) \in \mathrm{S}$ and (ii) $\forall y \in \mathrm{S}$, $\forall t \in \mathbb{R}$, there exists $x \in \mathrm{S}$ such that $\varphi(x, t) = y$.

Definition 5. Let $\mathrm{S} \subset X$ be invariant for the flow φ. Then a set $A \subset \mathrm{S}$ is an attractor for $\dot{x} = F(x)$ if the following statements hold:

(i) The set A is compact and invariant;

(ii) There exists an open neighborhood U of A with the property

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

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 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* $\hat{x} \in \mathrm{NE}$

(ii) if $\hat{x} \in \mathrm{NE}^{\mathrm{LU}}$ *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 $\hat{x} \in \mathrm{OZ}$, 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, 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 $\hat{x} \in \mathrm{OZ}$, then $\hat{x} \notin \mathcal{L}((x_n)_n)$, since $\mathcal{L}((x_n)_n)$ may include connected components of zeroes that contain $\hat{x}$. Indeed, in games with a 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 converges (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 that we describe here, since they are 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, "escapes." Yet, although it is 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 pushes 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 **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 converges to an unstable Nash equilibrium if the graph is bipartite. However, we provide an example in the Appendix (Example 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 equilibrium 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. To study convergence of the DGAP, we need to consider all possible trajectories and, to the best of our knowledge, we 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 two players, each having two 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 equilibrium is neither guaranteed nor easy to show when it occurs.

Because we are interested in the behavior of the DGAP, in particular, whether it allows for convergence to Nash equilibrium, 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 equilibrium is not a priori excluded.

The second case is trickier. If all equilibria are interior, and since the boundary always contains other zeroes that 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 equilibrium occurs with probability 1.

This result is difficult to obtain because of the following scenario: assume players start close to the boundary. Then, at the exploration stage, some decrease their efforts while others increase theirs. Although complementarities imply that the players who decreased their efforts would have been better off if they had instead increased them, they could still end up with a better payoff than before the exploration, and, thus, continue decreasing at the updating stage, getting closer to ∂X. We show that this cannot happen.

So as 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 stays away from the boundary in the long run.

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

From the mathematical point of view, the major problem to obtain Proposition 3 is to show that a stochastic approximation algorithm like that 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 increases 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 that 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* $\mathrm{Int}(X)$. *Then the learning process (x_n) almost surely converges to a stable Nash equilibrium:*

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

This theorem guarantees that the learning process not only converges to Nash equilibrium in most cases, it additionally converges 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 does not converge to general unstable sets.[7]

Note that the graph being bipartite does not imply that the process does not converge to an element of NE^{S}. However, we provide (Examples 2 and 3) in which we show that the noise can vanish on bipartite networks in games that have either no strategic complements or no symmetric externalities. In our examples, the noise vanishes at unstable equilibria.

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:*

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

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

4. Locally ordinal potential games

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

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

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.

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

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

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

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

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

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

For any LOPG, the only set to which the stochastic learning process can converge is the set of zeros of F. Complex ω-limit sets of the dynamical system, which are nonzeros, can be discarded (point (i)). We cannot, however, be sure that the process does 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 equilibrium 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 converges, 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 converges 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 $\{x_0 \in$

[8] We 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^∞.

$\mathcal{B}(A)\}$ 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 Λ is an attractor for $\dot{x} = F(x)$, then Λ 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.

Definition 8. Let $\mathrm{BR} : X \to X, x \mapsto \mathrm{BR}(x) := (\mathrm{BR}_1(x_{-1}), \dots, \mathrm{BR}_n(x_{-n}))$. The *continuous-time best-response dynamics* (hereafter, BRD) is defined as

$$\dot{x} = -x + \mathrm{BR}(x).$$

Definition 9. Let P be a smooth map and let Λ be a connected component of Z. We say that Λ is a local maximum of P if

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

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

We then have the following theorem.

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

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

(ii) The set Λ is a local maximum of P.

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

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

Second, Theorem 4 provides two methods for 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; 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 they 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[9] $g_i(x,r) = r_i \frac{\partial u_i}{\partial x_i}$. A game is *diagonally strictly concave* if

$$\exists r \in \left(\mathbb{R}_+^*\right)^N \mid \forall x^0 \neq x^1 \in X \quad \text{we have } \langle x^1 - x^0, g(x^0, r)\rangle + \langle x^0 - x^1, g(x^1, r)\rangle > 0. \tag{3}$$

Games that have this property are denoted by $\mathcal{G}^{\text{Ros}}$. It is proved (Theorem 2 of Rosen 1965) that games in $\mathcal{G}^{\text{Ros}}$ 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 $\mathcal{G}^{\text{Arr}}$) are as follows. Let S be a subset of N, the set of players, and define $f^{\text{S}} = \sum_{i\in \text{S}} u_i - \sum_{i\in N\setminus \text{S}} u_i$. A game is concave–convex if (a) for each $\text{S} \subseteq N$, the function $f^{\text{S}}(x^{\text{S}}, x^{N\setminus \text{S}})$ is concave in x^{S} for each $x^{N\setminus \text{S}}$ and convex in $x^{N\setminus \text{S}}$ for each x^{S}, and (b) for some $\text{S}^0 \subseteq N$, $f^{\text{S}^0}(x^{\text{S}^0}, x^{N\setminus \text{S}^0})$ is strictly concave in x^{S^0} for each $x^{N\setminus \text{S}^0}$. If, in addition, u_i is strictly concave in x_i, then we say that the game is *strictly concave–convex*.

Remark 2. Strictly concave–convex games are diagonally strictly concave, i.e., $\mathcal{G}^{\text{Arr}} \subset \mathcal{G}^{\text{Ros}}$. Thus, all properties of the later apply to the former. A proof is provided in the Appendix.

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

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 latter case, the process would go to infinity, while when the Nash equilibrium is unique, we show that 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 $\mathcal{G} \in \mathcal{G}^{\mathrm{Ros}}$. Then the following statements hold:*

(i) The variable Z is a finite set.

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

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

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

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

The proof of the first point is as follows: we prove that games in $\mathcal{G}^{\mathrm{Ros}}$ are such that after removing a subset of players playing 0, the remaining subgame is also in $\mathcal{G}^{\mathrm{Ros}}$. 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.

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 in 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 upward and downward with equal probability.

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

C4. At the updating stage, the payoff difference Δu_i^{n+1} is multiplied by e_i^n.

For our proofs to work, C1 can be totally relaxed, and 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:*

$$e^i_{2n+1} = \begin{cases} e^i_n + \dfrac{1}{n+1}\epsilon^i_n & \text{if } i \in I(n), \\ e^i_{2n} & \text{if } i \notin I(n). \end{cases}$$

- *Updating stage:*

$$e^i_{2n+2} = \begin{cases} e^i_{2n} + \epsilon^i_n \Delta u^i_{n+1} e^i_{2n} & \text{if } i \in I(n), \\ e^i_{2n} & \text{if } i \notin I(n). \end{cases}$$

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

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

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 $\chi^i_n = 1$ if $i \in I(n)$ and 0 otherwise. We have

$$x^i_{n+1} - x^i_n = x^i_n \tilde{\epsilon}^i_n \left[u_i\left(x^i_n + \frac{1}{n+1}\tilde{\epsilon}^i_n, x^{-i}_n + \frac{1}{n+1}\tilde{\epsilon}^{-i}_n \right) - u_i\big(x^i_n, x^{-i}_n\big) \right],$$

where $\tilde{\epsilon}^i_n := \chi^i_n \epsilon^i_n$. The first order development now gives

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

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

$$\mathbb{E}\big[\tilde{\epsilon}^i_n \tilde{\epsilon}^j_n \mid \mathcal{F}_n\big] = 0; \qquad \mathbb{E}\big[\big(\tilde{\epsilon}^i_n\big)^2 \mid \mathcal{F}_n\big] = \mathbb{P}\big(i \in I(n) \mid \mathcal{F}_n\big).$$

As a consequence,

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

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

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

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 $\sup_n \mathbb{E}(U_{n+1}^2|\mathcal{F}_n) < +\infty$ so as to write the stochastic difference equation as the sum of a deterministic term and a random component whose first order term is null on average with bounded variance, which is critical to using stochastic approximation methods. As can be seen from the proof of Proposition 1, any probability distribution that satisfies $\mathbb{E}(\epsilon_n^i) = 0$ 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

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

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

It is important that the sum diverges so as 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), $\lim_{n\to\infty} \alpha_n^i/\alpha_n^j = k_{ij}$, where $k_{ij} > 0$ is a constant.

6.4 *C4*

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

$$x_i^{n+1} - x_i^n = \epsilon_i^n \Delta u_{n+1}^i x_i^n, \tag{4}$$

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 his/her 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, where 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, so as 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 downward 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:

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

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 upward. 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 do not play 0 and, although this does not artificially prevent them from converging to $(0, 0)$, it helps avoid such situations. Indeed, in this example, our process converges to the unique Nash equilibrium $(1, 1)$ with probability 1 instead of remaining trapped at $(0, 0)$. To see why players 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 four 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 nonadmissible actions onto the feasible set by setting actions to $\max\{x_i, 0\}$ models a somewhat discontinuous behavior. There are smoother ways to prevent actions from leaving the feasible set, one of which is to multiply actions by x_i, as we do in (4). 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,[10] as we explain now.

The standard (projected) gradient method is defined as

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

[10]For modern applications of these procedures, see, for instance, Hazan (2016).

where f is the concave function that has to be maximized and 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), α_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,[11] where a primal-dual procedure is considered,

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

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 that 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 (1). Second, we prove that the process is well defined, i.e., $x^i_n > 0$ for all i and all n.

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

$$e^i_{2n+2} - e^i_{2n} = e^i_{2n}\epsilon^i_n \Delta u^i_{n+1}.$$

A first order development gives

$$\begin{aligned} \epsilon^i_n \Delta u^i_{n+1} &= \epsilon^i_n\left(u_i\left(e^i_{2n} + \frac{1}{n+1}\epsilon^i_n, e^{-i}_{2n} + \frac{1}{n+1}\epsilon^{-i}_n\right) - u_i\left(e^i_{2n}, e^{-i}_{2n}\right)\right) \\ &= \frac{1}{n+1}\left(\epsilon^i_n\right)^2 \frac{\partial u_i}{\partial x^i}(e_{2n}) + \frac{1}{n+1}\epsilon^i_n \sum_{j\neq i} \epsilon^j_n \frac{\partial u_i}{\partial x^j}(e_{2n}) + \mathcal{O}\left(\frac{1}{n^2}\right). \end{aligned}$$

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

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

By setting $U^i_{n+1} = \epsilon^i_n x^i_n \sum_{j\neq i} \epsilon^j_n \frac{\partial u_i}{\partial x^j}(x_n)$, we get (1). Finally, note that $\mathbb{E}(\epsilon^j_n) = 0$ for all j, and that ϵ^i_n and ϵ^j_n are independent, so that

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

[11] See Nesterov (2009) and Mertikopoulos and Zhou (2019) for recent applications in game theory.

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 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 $n \geq 0$,

$$\frac{x^i_{n+1}}{x^i_n} \geq \left(1 - \|e_{2n+1} - x_n\|_\infty\right)$$

and $|e^i_{2n+1} - x^i_n| \leq \frac{1}{n+1}$ for all i. As a consequence,

$$\frac{x^i_{n+1}}{x^i_n} \geq \left(1 - \frac{1}{n+1}\right).$$

Thus, $x^i_1 \geq 0$ and

$$x^i_n \geq x^i_1 \prod_{k=1}^{n-1}\left(1 - \frac{1}{k+1}\right) = \frac{1}{n+1} x^i_1 \geq 0.$$

A.2 *Proof of Proposition 2*

Pick $\hat{x} \in \mathrm{OZ}$ 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 *Part* (i). Pick an $\hat{x} \in \mathrm{OZ}$ and let us fix $i \in \{1, \ldots, N\}$ such that $\hat{x}^i = 0$ and $\frac{\partial u^i}{\partial x^i}(\hat{x}) > 0$. Observe first that we can work on the event $\{\sup_{n\to+\infty} \|x_n\| < +\infty\}$ since, otherwise, there is nothing to prove.

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

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

Let $\tilde{U}^i_{n+1} = \epsilon^i_n \sum_j \epsilon^j_n \frac{\partial u_i}{\partial x_j}(x_n)$, so that

$$x^i_{n+1} = x^i_n\left(1 + \frac{1}{n+1}\left(\frac{\partial u_i}{\partial x_i}(x_n) + \tilde{U}^i_{n+1} + \frac{\xi^i_{n+1}}{x^i_n}\right)\right).$$

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

$$\frac{1}{x^i_{n+1}} = \frac{1}{x^i_n}\left(1 - \frac{1}{n+1}\left(\frac{\partial u_i}{\partial x_i}(x_n) + \tilde{U}^i_{n+1}\right) + o\left(\frac{1}{n}\right)\right).$$

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

$$\mathbb{E}\left(\frac{1}{x^i_{n+1}} - \frac{1}{x^i_n} \,\middle|\, \mathcal{F}_n\right) \leq -\frac{1}{x^i_n} \cdot \frac{1}{n+1} \cdot \frac{\eta}{2} \leq 0.$$

Therefore, the random sequence $(1/x^i_n)_n$ 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^i_n tends to zero almost surely. These two convergence properties are contradictory and the conclusion follows.

Part (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 μ 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.,

$$\liminf_{n \to +\infty} \mathbb{E}\big(\langle U_{n+1}, w\rangle^2 \mid \mathcal{F}_n\big) > 0, \tag{5}$$

on the event $\{\lim_n x_n = \tilde{x}\}$. A sufficient condition for (5) to hold is

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

Consider any x_n and any vector v. Then

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

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

$$\mathbb{E}\big(\langle U_{n+1}, v\rangle^2 \mid x_n\big) = \sum_{i<j} \left(v_i x^i_n \frac{\partial u_i}{\partial x^j}(x_n) + v_j x^j_n \frac{\partial u_j}{\partial x^i}(x_n)\right)^2. \tag{6}$$

By (6), 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, and k such that $g_{ij} = g_{ik} = g_{jk} = 1$ and

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

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 example 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 1. Consider the two-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

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

the eigenvalues of which 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

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

and the noise condition (5) does not hold. ◊

Appendix B: Proof of results of Section 3

B.1 *Proof of Proposition 3*

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

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

Since the game has strategic complements,

$$\frac{\partial u_i}{\partial x_i}(x_i, x_{-i}) > \alpha_i > 0 \quad \forall x_i < \overline{x}_i, \forall x_{-i} \in X_{-i}. \tag{7}$$

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

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

In what follows we 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 equilibrium on ∂X).

Remark 3. Three simple observations are in order.

(i) Condition (7) 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 (7), that the coordinate k achieves the minimum value of a vector $x \in X \setminus [a,+\infty[^N$ verifies that $\frac{\partial u_k}{\partial x^k}(x) > \alpha$, where $\alpha = \min_i \alpha_i > 0$. Therefore, this particular k belongs to the set I_x.

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

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

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

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

So as 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

$$\mathbb{P}\big(\big\{\mathcal{L}\big((x_n)_n\big) \subset \partial X\big\} \cap \{x_n \in V_\varepsilon \cap \mathcal{U}_R,\ \text{for all } n \geq n_*\}\big) > 0, \tag{8}$$

where

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

In what follows, we work on the event E defined by (8) and we assume that $n \geq n_*$.

For $\beta > 0$, let the function

$$\Phi_\beta(x) = -\frac{1}{\beta} \ln\left(\sum_{i=1}^{N} \exp\big(-\beta x^i\big)\right),$$

which is concave if extended as $-\infty$ to $\mathbb{R}^N$. The function Φ verifies the well known relation

$$\min_{i=1,\dots,N} x^i - \frac{\ln(N)}{\beta} \le \Phi_\beta(x) \le \min_{i=1,\dots,N} x^i. \tag{9}$$

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

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

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

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

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

Observe that if $x \in V_\varepsilon \cap \mathcal{U}_R$ and if $\frac{\partial u_i}{\partial x^i}(x) \le 0$ for some coordinate i, we have that

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

using the fact that there exists some k such that $x^k \le \varepsilon$ and that $x^i \ge a$ (cf. Remark 3).

Alternatively, for $k \in I_x$ such that $x^k = \min_i x^i$,

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

Recall that the variable x_n follows the recursion

$$x^i_{n+1} = x^i_n + \frac{1}{n+1}\left(x^i_n \frac{\partial u_i}{\partial x_i}(x_n) + U^i_{n+1} + \xi^i_{n+1}\right),$$

where $\mathbb{E}(U_{n+1} \mid \mathcal{F}_n) = 0$ and $|\xi^i_n| \le C/n$ for a deterministic constant C.

Let us define $z_n = \Phi_\beta(x_n)$. Note first that, from (9),

$$-\frac{\ln(N)}{\beta} \le \min_{i=1,\dots,N} x^i_n - \frac{\ln(N)}{\beta} \le z_n \le \min_{i=1,\dots,N} x^i_n \le \epsilon.$$

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

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

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

Equivalently,

$$z_{n+1} \ge z_n + \sum_{j=1}^N \pi_j(x_n)\big(x^j_{n+1} - x^j_n\big) - \frac{L}{2}\|x_{n+1} - x_n\|^2$$

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

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

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

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

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} be the set of indices such that $\frac{\partial u_i}{\partial x^i}(x_n) \leq 0$.

For all $n \geq n_*$, we have

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

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

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

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

Thus,

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

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

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

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

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

Let us call $\rho_n = \mathbb{E}(\theta_n) - c(\beta)$. Then we want to analyze 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 to the fact that $z_n \in [-\ln(N)/\beta, \epsilon]$ almost surely on the event E.

B.2 *Proof 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 6 (Benaïm and Faure 2012). *Let $(x_n)_n \in X$ be a random process that can be written as*

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

where the following statements hold:

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

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

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

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

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

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

Then

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

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

Proof of Theorem 2 We want to apply Theorem 6. 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 $x \in \operatorname{Int}(X)$, we have

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

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 μ.

Note that $\tilde{x} \notin \partial X$ by Proposition 3, and that when $\tilde{x} \in \operatorname{Int}(X)$ 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 $\tilde{x} \in \operatorname{Int}(X)$. This implies that there exists a partition (A, B) of N such that if $a \in A$ and $\frac{\partial u_a}{\partial x_b}(\tilde{x}) \frac{\partial u_b}{\partial x_a}(\tilde{x}) > 0$, then $b \in B$.

Using the computations just developed, we need to show that

$$\sum_{a<b}\left(v_a x^a \frac{\partial u_a}{\partial x^b}(\tilde{x}) + v_b x^b \frac{\partial u_b}{\partial x^a}(\tilde{x})\right)^2 \neq 0. \tag{10}$$

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, by Assumption 2 (symmetric externalities), it must be that $\operatorname{sgn}(v_a) = -\operatorname{sgn}(v_b)$ for any $a \in A$ and any $b \in B$. Since the interaction graph is connected, we may assume without loss of generality that $v_a > 0\ \forall a \in A$ and $v_b < 0\ \forall b \in B$.

Because μ is strictly positive and v is the corresponding normalized eigenvector, we should have $\langle v DF(\tilde{x}), v\rangle = \mu \sum_i v_i^2 > 0$, since $v \neq 0$. However, we show that this can only be true if (10) holds. By a simple rearrangement of the indices, the Jacobian matrix at $\tilde{x}$ can be written as

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

where D_A is diagonal and the diagonal terms are equal to $x^a\partial^2 u_a/\partial(x^a)^2(\tilde{x}) \le 0$ with $a \in A$, and similarly for D_B. The variables M and N are nonnegative matrices, as $x^i\partial^2 u_i/\partial x^i\partial x^j \ge 0\ \forall i \ne j$.

Thus,

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

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 2. Consider the two-player game with strategic complements:

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

This game has antisymmetric 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

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

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

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

and the noise condition does not hold. ◊

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

Example 3. Consider the four-player example with strategic substitutes,

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

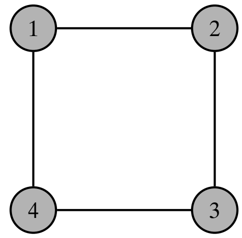

Figure 1. Interaction graph between four players. This graph is bipartite.

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

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

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

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

and the noise condition (5) does not hold. ◊

Appendix 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 10. Let $P : X \to \mathbb{R}$ be continuously differentiable. We say that P is a strict[12] Lyapunov function for $\dot{x} = F(x)$ if

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

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

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

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

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

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

Proof. By assumption,

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

(i) We have

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

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

(ii) We have

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

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 2. *Assume $\mathcal{G}$ is an* LOPG. *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 let 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, $Z = \bigcup_{A \in \mathcal{P}(\{1,\dots,N\})} Z_A$.

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

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

For $x \in Z_A$, denote $x^A = (x_i)_{i \in A}$. 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, as is P^A, and by Sard's theorem, $P^A(\{x^A : x \in Z_A\})$ has an 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 an empty interior. Since Z is a finite union of such sets, $P(Z)$ has an empty interior.

Part (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 if $P(Z)$ has an empty interior, then $\mathcal{L}(x_n) \subset Z$ almost surely.

Part (ii). By Lemma 1 and Corollary 6.6 in Benaïm (1999), we have

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

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 2 above). Moreover, we have $\Lambda \subset Z$. Thus, P is constant on Λ. Let $v := P(\Lambda)$. If Λ is not a local maximum of P, then there exists a sequence x_n such that $d(x_n, \Lambda) \to_n 0$ and $P(x_n) > v$. Since Λ is isolated, we have $x_n \in X \setminus Z$ and $P(\varphi(x_n, t)) > P(x_n) > v$ for any $t > 0$; hence, $d(\varphi(x_n, t), \Lambda) \nrightarrow 0$ and Λ is not an attractor.

Let us now prove that (ii) implies (iii). First we show that Λ is contained in NE. Suppose that there exists $\hat{x} \in \Lambda \setminus \text{NE}$. 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 Λ is an attractor for the best-response dynamics, P is a strict Lyapunov function for the best-response dynamics,[13] and $\Lambda \subset \text{NE}$. The statement we want to prove is then a consequence of Proposition 3.25 in Benaim et al. (2005). We adapt the proof in our context for convenience. First of all, observe that Λ is actually a strict local maximum of P: there exists an open (isolating) neighborhood U of Λ such that $P(x) < v = P(\Lambda)\ \forall x \in U \setminus \Lambda$. This is a simple consequence of the fact that P is strictly increasing along any solution curve with initial conditions in $U \setminus \Lambda$. Now let $V_r := \{x \in U : P(x) > v - r\}$. Clearly $\bigcap_r V_r = \Lambda$. Also $\varphi(\overline{V_r}, t) \subset V_r$, for $t > 0$, r small enough.[14] This implies that $\Lambda = \bigcap_{r>0} V_r$ contains an attractor A. The potential being constant on Λ, A cannot be strictly contained in Λ and, therefore, Λ is an attractor.

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

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

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

Appendix 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 we 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

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

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

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

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 $\mathrm{S} = N$ in the definition of strictly concave–convex games. Thus, C is negative semi-definite. This proves that $G(x,\mathbf{1}) + G'(x,\mathbf{1})$ is negative definite. □

Proof of Theorem 5. Suppose first that there is a unique Nash equilibrium. Then note that under (3), we have, for any $x \neq \overline{x}$,

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

because

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

Given an element $x \in X$, let $I(x) := \{i \in \mathcal{N} : x_i = 0 \text{ and } \frac{\partial u_i}{\partial x_i}(x) > 0\}$. 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 3. *Let $J \subset N$. There exists a unique profile $\tilde{x}^J$ with the properties*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 $\Phi^k : \tilde{X}^k \to \mathbb{R}$ as

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

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

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

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

As a consequence, every solution curve converges to a zero of F, i.e., either $\overline{x}$ or one of the $\tilde{x}^k$.[15] More precisely, $\overline{X}$ and $\tilde{X}^k$ are invariant, and $\{\overline{x}\}$ (resp. $\tilde{x}^k$) is a global attractor for the flow $\phi_{|\overline{X}}$ (resp. $\phi_{|\tilde{X}^k}$); in particular, for any $x_0 \in \overline{X}$ (resp. $x_0 \in \tilde{X}^k$), then $\lim_{t\to+\infty} \phi_t(x_0) = \overline{x}$ (resp. $\lim_{t\to+\infty} \phi_t(x_0) = \tilde{x}^k$).

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

Theorem 7 (Benaïm (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 let* A *be an attractor with basin of attraction* $\mathcal{B}(A)$. *If* $L \cap \mathcal{B}(A) \neq \varnothing$, *then* $L \subset A$.

We now prove that the sets L_k and $\overline{L}$ are the only internally chain transitive sets. This concludes the proof because, as we mentioned above, $(x_n)_n$ is almost surely bounded. Note that $\overline{X}$ is an open set in X. To do so, we first claim that it is always possible to relabel the family $(\tilde{x}^k)_{k=1,\dots,K}$ such that $\tilde{X}^k$ is an open set of $\bigcup_{l=1}^{k} X^l$ for $k = 2, \dots, K$.

Let L be internally chain transitive. By previous results, if L intersects $\overline{X}$, then $L \subset \{\overline{x}\}$ because $\overline{X}$ is the basin of attraction of $\overline{x}$. Suppose that this is not the case; then $L \subset \bigcup_{k=1}^{K} X^k$. Since X_K is open in $\bigcup_{k=1}^{K} X^k$, $\tilde{x}^K$ is an attractor of the flow restricted to $\bigcup_{k=1}^{K} X^k$, with basin of attraction $\tilde{X}^k$. Hence, if $L \cap \tilde{X}^k \neq \varnothing$, then $L = \{\tilde{x}^k\}$. By a recursive argument, either $L = \{\overline{x}\}$ 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 (Benaïm 1999)). This is a contradiction to Theorem 1.

References

Arrow, Kenneth J. and Leonid Hurwicz (1960), "Stability of the gradient process in n-person games." *Journal of the Society for Industrial and Applied Mathematics*, 8, 280–294. [1472, 1474, 1486]

Babichenko, Yakov (2012), "Completely uncoupled dynamics and Nash equilibria." *Games and Economic Behavior*, 76, 1–14. [1472]

Benaïm, Michel (1996), "A dynamical system approach to stochastic approximations." *SIAM Journal on Control and Optimization*, 34, 437–472. [1478]

Benaïm, Michel (1999), "Dynamics of stochastic approximation algorithms." In *Séminaire de Probabilités, XXXIII*, volume 1709 of Lecture Notes in Mathematics, 1–68, Springer, Berlin. [1478, 1484, 1503, 1506]

[15] 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 1 to one of the zeroes of the dynamics.

Benaïm, Michel and Mathieu Faure (2012), "Stochastic approximations, cooperative dynamics and supermodular games." *Annals of Applied Probability*, 22, 2133–2164. [1473, 1498]

Benaim, Michel, Josef Hofbauer, and Sylvain Sorin (2005), "Stochastic approximations and differential inclusions." *SIAM Journal on Control and Optimization*, 44, 328–348. [1474, 1503]

Brandiere, Odile and Marie Duflo (1996), "Les algorithmes stochastiques contournent-ils les pièges?" *Annales de l'I.H.P. Probabilités et Statistiques*, 32, 395–427. [1493]

Bravo, Mario, David Leslie, and Panayotis Mertikopoulos (2018), "Bandit learning in concave N-person games." In *Advances in Neural Information Processing Systems 31*, 5661–5671. [1474]

Dindos, Martin and Claudio Mezzetti (2006), "Better-reply dynamics and global convergence to Nash equilibrium in aggregative games." *Games and Economic Behavior*, 54, 261–292. [1474]

Foster, Dean P. and H. Peyton Young (2006), "Regret testing: Learning to play Nash equilibrium without knowing you have an opponent." *Theoretical Economics*, 1, 341–367. [1472]

Germano, Fabrizio and Gábor Lugosi (2007), "Global Nash convergence of Foster and Young's regret testing." *Games and Economic Behavior*, 60, 135–154. [1472]

Hart, Sergiu and Andreu Mas-Colell (2003), "Uncoupled dynamics do not lead to Nash equilibrium." *American Economic Review*, 93, 1830–1836. [1472]

Hart, Sergiu and Andreu Mas-Colell (2006), "Stochastic uncoupled dynamics and Nash equilibrium." *Games and Economic Behavior*, 57, 286–303. [1472]

Hazan, Elad (2016), "Introduction to online convex optimization." *Foundations and Trends in Optimization*, 2, 157–325. [1490]

Huck, Steffen, Hans-Theo Normann, and Jörg Oechssler (2004), "Through trial and error to collusion." *International Economic Review*, 45, 205–224. [1474]

Mertikopoulos, Panayotis and Zhengyuan Zhou (2019), "Learning in games with continuous action spaces and unknown payoff functions." *Mathematical Programming*, 173, 465–507. [1474, 1491]

Nesterov, Yurii (2009), "Primal-dual subgradient methods for convex problems." *Mathematical Programming*, 120, 221–259. [1491]

Pemantle, Robin (1990), "Nonconvergence to unstable points in urn models and stochastic approximations." *Annals of Probability*, 18, 698–712. [1493]

Perkins, Steven and David S. Leslie (2014), "Stochastic fictitious play with continuous action sets." *Journal of Economic Theory*, 152, 179–213. [1473]

Posch, Martin (1997), "Cycling in a stochastic learning algorithm for normal form games." *Journal of Evolutionary Economics*, 7, 193–207. [1482]

Rosen, J. Ben (1965), "Existence and uniqueness of equilibrium points for concave n-person games." *Econometrica*, 33, 520–534. [1473, 1474, 1486, 1504]

Ruelle, David (1981), "Small random perturbations of dynamical systems and the definition of attractors." *Communications in Mathematical Physics*, 82, 137–151. [1480]

Tatarenko, Tatiana (2018), "Stochastic learning in potential games: Communication and payoff-based approaches." Unpublished paper, arXiv:1804.04403. [1474]

Vives, Xavier (1990), "Nash equilibrium with strategic complementarities." *Journal of Mathematical Economics*, 19, 305–321. [1474, 1481]

Co-editor Simon Board handled this manuscript.

Manuscript received 6 September, 2018; final version accepted 10 January, 2020; available online 6 February, 2020.