Fast Convergence in Population Games

Itai Arieli Peyton Young*

September 25, 2011

Abstract

A stochastic learning dynamic exhibits fast convergence in a population game if the expected waiting time until the process comes near a Nash equilibrium is bounded above for all sufficiently large populations. We propose a novel family of learning dynamics that exhibits fast convergence for a large class of population games that includes coordination games, potential games, and supermodular games as special cases. These games have the property that, from any initial state, there exists a continuous better-reply path to a Nash equilibrium that is locally stable.

1 Overview

Evolutionary game theory is concerned with the dynamical behaviour of large populations of players who are engaged in repeated interactions. Players adapt their behaviour according to local or partial information but not always in a fully rational manner. Two questions present themselves: Are

*This research was supported by AFOSR grant #FA9550-09-1-0538.

there natural adaptive rules that lead to Nash equilibrium in general classes of games? If so, how long does it take to reach equilibrium from out-of-equilibrium conditions?

The answer to the first question depends on whether the adaptive dynamics are deterministic or stochastic. In the former case the results are largely negative: there exist no natural deterministic dynamics that converge to Nash equilibrium in general games (Hofbauer and Swinkels [14, 1991]; Hart and Mas-Colell [10, 2003]; Hofbauer and Sandholm [13, 2008]). The basic difficulty is that given virtually any deterministic dynamic one can choose the payoffs so that the process gets trapped in a cycle. By contrast, the answer to the question is positive for certain classes of *stochastic* dynamics. Indeed there exist simple stochastic learning rules that come close to Nash equilibrium with high probability in large classes of normal-form games (Foster and Young, [6, 2003],[7, 2006]; Hart and Mas-Colell, [11, 2006]; Germano and Lugosi, [8, 2007]; Young, [26, 2009]). Furthermore there are rules with this property that are completely uncoupled, that is, a players updating procedure depends only on his own realized payoffs, and not on the actions or payoffs of anyone else.

The answer to the second question – how long does it take to reach Nash equilibrium – has only recently been studied in any depth. The key result to date is due to Hart and Mansour [9, 2010]. They consider learning rules that are *uncoupled*, that is, the updating procedure depends on a player's own payoffs and possibly on the actions of others, but not on others' payoffs. Using methods from communication theory, they show that when everyone uses an uncoupled learning rule, there exist N-person normal form games such that the number of periods it takes to reach a Nash equilibrium is exponential in N.[1] These results are proved by constructing payoff functions that are difficult to learn by design, that is, the payoffs constitute a worst-case scenario.[2]

The results of Hart and Mansour leave open the question of whether Nash equilibrium, or something close to Nash equilibrium, can be learned reasonably quickly in games that have a natural payoff structure. The purpose of this paper is to show that this is indeed the case. In particular, we shall show that there exists a family of stochastic learning rules such that behaviors come close to Nash equilibrium with high probability in bounded time (not merely polynomial time) provided that the population game is weakly acyclic'. Such games have the property that, from any initial state, there exists a Lipschitz continuous path to a neighbourhood of some Nash equilibrium such that: i) it is a better-reply path, that is, the direction of motion at each point on the path represents a payoff-improving change by one or more subpopulations, and ii) for any $\epsilon > 0$ there is a $\delta > 0$ such that, once the process is within δ of the target Nash equilibrium there is no better-reply path that moves more than ϵ away from it.[3]

A special case of weakly acyclic population games are potential games in which every Nash equilibrium is a strict local maximum of the potential function and there are a finite number of local maxima. Another example are supermodular games (see section 5). The key feature of a weakly acyclic population game is that, although it may contain better-reply cycles, there

[1]By contrast, correlated equilibria can be learned in a polynomial number of time periods (Papadimitriou, [17, 2005]; Hart and Mansour, [9, 2010]).

[2]For related work on the speed of convergence see Babichenko ([1, 2010a], [2, 2010b]).

[3]This definition extends the concept of weak acyclicity in normal-form games. A finite normal form game G is *weakly acyclic* if from each strategy-tuple there exists a better-reply path – one player moving at a time – that leads to a pure-strategy Nash equilibrium (Young, [24, 1993], [25, 1998]).

always exists a better-reply path that leads away from such a cycle and toward a Nash equilibrium (A potential game is much more special because there can exist no better-reply cycles).

A novel feature of our set-up is that the learning dynamics involve stochastic perturbations of two different types. If an agent is currently playing strategy i and he learns that strategy j has a higher payoff, he switches from i to j with positive probability. In the simplest case, the switching probability is a linear function of the payoff difference, a specification known as the Smith dynamic (Smith [23, 1984]).[4] In addition, however, we shall assume that there is a positive probability that the environment prevents an i-player from learning about the payoffs of the j-players. In other words, at each point in time there is a chance that the information about the payoffs of the j-players is inaccessible to the i-players due to some form of interference.

We assume that these interference probabilities evolve over time according to a random walk, and that the updating opportunities for individual agents are governed by i.i.d. Poisson arrival processes. Taken together these two types of perturbations define a stochastic dynamical system that describes how the behaviors in the population evolve. The main contribution of this paper is to show that, when the game is weakly acyclic, the expected time it takes to reach a neighbourhood of a Nash equilibrium with high probability is bounded above for all sufficiently large populations N. We also show that this result fails in the absence of environmental interference, that is, when players get information with certainty and they update according to standard revision protocols such as the Smith dynamic. Stochasticity in the environment combined with stochasticity in the players' learning rules is

[4] More generally, switching rules of this type are called *pairwise comparison protocols* (Sandholm, [20, 2010]).

needed for fast convergence in this class of games.

2 Learning Dynamics in Population Games

We begin by defining a population game (Hofbauer and Sandholm [12, 2007]). Let $\mathcal{P} = \{1, \ldots, m\}$ be a set of m populations, and let N be the size of each population. Members of population $p \in \mathcal{P}$ choose strategies from the set $S^p = \{1, \ldots, n^p\}$. Let $X^p = \Delta^p$ be the $n^p - 1$ dimensional simplex and let $X = \prod_{p \in \mathcal{P}} X^p$ be the overall strategy distribution. Let $n = n^1 + \ldots + n^m$.

The payoff function of strategy $i \in \mathcal{S}^p$ is denoted by $F_i^p : X \to \mathbb{R}$ and is assumed to be Lipschitz continuous.

Definition 1. A population state x is a *Nash equilibrium* if, for every population $p \in \mathcal{P}$ and every strategy $i \in S^p$ such that $x_i^p > 0$,

$$\forall j \in S^p \; F_i^p(x) \geq F_j^p(x).$$

Definition 2. Given a population game, a *revision protocol* (Bjornerstedt and Weibull [4, 1996]) is a set of matrices $(R^1, \ldots, R^m)$ one for each population p, and R^p is of size $n^p \times n^p$. Each element, $\rho_{ij}^p \in R^p$ corresponds to a Lipschitz function $\rho^p : \mathbb{R}^{n^p} \times X^p \to \mathbb{R}_+^{n^p \times n^p}$. The value $\rho_{ij}^p(\pi^p, x^p)$, represents the *rate* at which members of population p who are playing strategy i switch to strategy j, when the state is x and the payoff vector π^p.

Let N be the number of agents in each population, and let $\chi^N = \{x \in X : Nx \in \mathbb{Z}^n\}$ be the set of feasible states. Each state $x \in \chi^N$ may be interpreted as a vector $x = (x^1, \ldots, x^m)$, such that each x^p represents the distribution of strategies in population p. In other words, $N \cdot x_i^p$ is the number of agents in population p who are playing strategy $i \in S^p$. Let $e_i^p \in \mathbb{R}^n$ be the unit vector with 1 in the ith coordinate.

The behaviour of the population can be described by a continuous time pure jump Markov process $\{X^N(t)\}_{t\geq 0}$ as follows. Let

$$\lambda = \max_{p,i}\{\sum_j \rho^p_{ij}(x) : x \in X\}. \tag{1}$$

Assume that, revision opportunities arrive at the rate $\lambda_N = m \cdot N \cdot \lambda$. Each time such an opportunity occurs and the population is in state x, a new state $x + z$ is determined according to the following transition probabilities:

$$\mathbb{P}^N_{x,x+z} = \begin{cases} \frac{x^p_i \rho^p_{ij}(x,F(x))}{m\lambda} & \text{for} \quad z = \frac{1}{N}(e^p_j - e^p_i) \\ 1 - \sum_p \sum_{i\in S^p} \sum_{j\in S^p} \frac{x^p_i \rho^p_{ij}(x,F(x))}{m\lambda} & \text{if} \quad z = 0. \end{cases} \tag{2}$$

Based on Lemma 1 in Benaïm and Weibull [3, 2003], we know that when the size of the population grows, the behaviour of $\{X^N(t)\}_{t\geq 0}$ can be approximated by the following *mean field* differential equation:[5]

$$\dot{z}^p_i = \sum_{j\in S^p} z^p_j \rho^p_{ji}(F^p(z), z^p) - z^p_i \sum_{j\in S^p} \rho^p_{ij}(F^p(z), z^p). \tag{3}$$

Example 1 (The Replicator Dynamic)**.**
By letting $\rho^p_{ij}(\pi^p, x^p) = x^p_j[\pi^p_i - \pi^p_j]_+$ and $F^p(z) = \sum_{j\in S^p} x^p_j F^p_j(z)$, a simple calculation reveals that (3) takes following form:

$$\dot{z}^p_i = z^p_i(F^p_i(z) - F^p(z)),$$

This is the *replicator dynamic.*

Example 2 (Smith Dynamic)**.**
If we let $\rho^p_{ij}(\pi) = [\pi^p_j - \pi^p_i]_+$, the resulting mean field dynamic is

$$\dot{z}^p_i = \sum_{j\in S^p} z^p_j[F^p_i(z) - F^p_j(z)]_+ - z^p_i[F^p_j(z) - F^p_i(z)]_+.$$

[5]Henceforth, to simplify notation, we omit the dependence on t.

Definition 3. A *pairwise comparison revision protocol* ρ has the property that, for each population p and every pair of strategies $i, j \in S^p$,

- $\rho_{ij}^p(\pi, x) = \rho_{ij}^p(\pi_i^p, \pi_j^p)$.[6]
- $\text{sign}(\rho_{ij}^p(\pi)) = \text{sign}(\pi_j - \pi_i)$.

In words, a pairwise comparison revision protocol depends on the state of the population only through the payoffs, and for every two strategies $i, j \in S^p$ there exists a positive switching rate from i to j iff $\pi_j > \pi_i$. The Smith revision protocol (Example 2) has this property, but not the protocol defining the replicator-dynamic (Example 1).

Henceforth we shall only consider pairwise comparison revision protocols, and therefore we shall omit the term 'pairwise comparison' in the interest of brevity.

3 Weakly Acyclic Games

Our first task is to generalize the definition of weakly acyclic games to the population game set-up. For every $x \in X$ the *tangent space* at x is defined as follows:

$$TX(x) = \{y \in \mathbb{R}^n : z = \alpha(y - x) \text{ for some } y \in X \text{ and } \alpha \geq 0\}.$$

Let $M > 0$ be some positive constant. For every $p \in \mathcal{P}$ and $i \in S^p$, let

$$V_i^p(x) = \{\sum_{j \in S^p} x_j^p \theta_{ji}^p [F_i^p(x) - F_j^p(x)]_+ - x_i^p \theta_{ij}^p [F_j^p(x) - F_i^p(x)]_+ : 0 \leq \|\theta\|_\infty \leq M\}. \tag{4}$$

Further, let $V = ((V_i^p)_{i \in S^p})_{1 \leq p \leq m}$.

[6]For simplicity write $\rho_{ij}^p(\pi^p)$.

Definition 4. A Lipschitz path $z : [0,T] \to X$ is a *better-reply* path if for almost every time $t \in [0,T]$,

$$\dot{z}(t) \in V(z(t)). \tag{5}$$

Equivalently, for every population p, strategy $i \in S^p$, and almost every time $t \in [0,T]$, $\forall p \in \mathcal{P},\ i \in S^p$

$$\dot{z}_i^p \in \{\sum_{j \in S^p} z_j^p \theta_{ji}^p [F_i^p - F_j^p]_+ - \sum_{j \in S^p} z_i^p \theta_{ij}^p [F_j^p - F_i^p]_+ : 0 \le \|\theta^p\|_\infty \le M\}.$$

A better-reply path has the property that, at each point along the path, a positive change from j to i in population p occurs only if strategy i yields a strictly higher payoff than does j.

Remark 1. Definition 5 can be generalized by replacing $[F_i^p - F_j^p]_+$ with a monotone Lipschitz function $g(F_i^p - F_j^p)$, such that $g(x) = 0$ for $x \le 0$ and $g(x) > 0$ for $x > 0$.

Let $\Phi(t,x)$ be the *semi-flow* of the differential inclusion (5), that is, $\Phi(t,x)$ is the set of points that are on some better-reply path starting from x:

$$\Phi(t,x) = \{z(t) : z(0) = x \text{ and } z(s)_{\{0 \le s \le t\}} \text{ is a better-reply path}\}.$$

Definition 5. Let x be a Nash equilibrium. The *ϵ-basin of x* is defined as follows:

$$\mathcal{BA}_\epsilon(x) = \{y \in X : \forall t > 0,\ \Phi(t,y) \subseteq \mathrm{B}_\epsilon(x)\}.$$

Definition 6. A population game Γ is *weakly acyclic* if for every $\epsilon > 0$ and $y \in X$ there exists a better-reply path $z : [0,T] \to X$ such that $z(T) \in \mathrm{int}(\mathcal{BA}_\epsilon(x))$ for some equilibrium x. (In particular we assume that $\mathcal{BA}_\epsilon(x)$ has a non empty interior.)

Recall that a normal form game $G = (I, (S_i)_{i \in I}, (U_i)_{i \in I})$, is *weakly acyclic* if for every pure strategy there exists a strict better-reply path to some pure Nash equilibrium s^* (Young [24, 1993]). That is, for every $s \in S$, there exists a sequence of strategy profiles $s^1, \dots, s^k$ and a corresponding sequence of players $i_1, \dots, i_k$ such that s^k is a Nash equilibrium, and for every $1 \leq l \leq k-1$, $s^l_{-i_l} = s^{l+1}_{-i_l}$ and $U_{i_l}(s_{i_l}) < U_{i_l}(s_{i_{l+1}})$.

A normal form game defines a population game Γ as follows. For each player $i \in I$, create a population in Γ with strategy space S_i. A state in Γ specifies the proportion of members in each population playing each available strategy, and the payoffs are defined as if these proportions were mixed strategies in G.

We establish the following connection between the two definitions for weakly acyclicity.

Proposition 1. If Γ is the population game derived from a generic normal form weakly acyclic game G, then Γ is weakly acyclic.

Proof. Assume for the moment that G is a weakly acyclic *two-player* game and let Γ be the population game on $X = X^1 \times X^2$ that is derived from G. Let $y \in X$ be an initial state that is not a Nash equilibrium. Then one of the players, say player 1, can improve his payoff. By genericity we may assume that i_1 is a unique best reply for player 1.

Define a better-reply path $z(t)$ such that, over some initial time interval $[0, T_1]$, most of population 1 switches to playing strategy i_1. (For this purpose let $\theta^1_{ji_1} = M$ in Equation 4 for every $j \neq i_1$.) Note that the rate of change from j to i_1 is proportional to the fraction of j-players in population 1, namely x^1_j, hence after any finite time there will be some residue of the population that is still playing j. However, this residue can be made as small as we like by choosing T_1 to be large.

By genericity, there exists a strategy i_2 for player 2 such that i_2 is the unique best reply for player 2 in state $z(T_1)$. Beginning at time T_1 we let $\theta^2_{ji_2} = M$ for all $t \in (T_1, T_2]$, where T_2 is chosen to be sufficiently large that by time T_2 most of the population is concentrated on (i_1, i_2). By assumption, there exists a better-reply path in G from (i_1, i_2) to some strict Nash equilibrium e. One can extend the better-reply path in Γ in a way that very closely mimics this better reply path in G, and that ends in the ϵ-basin of e. A similar argument applies if G is an n-player weakly acyclic game with $n > 2$. $\square$

We now demonstrate two important families of normal form games which are weakly acyclic.

Definition 7. A normal form game G with a strategy set S is a *potential game* if there exists a function $P : S \to \mathbb{R}$ such that, for every player $i \in I$, a strategy profile $s_{-i} \in S_{-i}$, and pure strategies $a, b \in S_i$,

$$U_i(a, s_{-i}) - U_i(b, s_{-i}) = P(a, s_{-i}) - P(b, s_{-i}).$$

Every generic potential game is weakly acyclic because the potential strictly increases along any better-reply path, hence cycles are impossible.

A second family of weakly acyclic games consists of two player coordination games (which are not necessarily potential games.) A third family consists of supermodular games, which we shall consider in more detail in section 5.

Definition 8. A two player game G is a *coordination game* if both players have the same strategy set $S = S_1 = S_2$, and the profile (s, s) is a strict pure Nash equilibrium for every $s \in S$.

4 Population Potential Games

Potential games can be generalized to the population game set-up as follows.

Definition 9. Let F be the payoff function of a population game, and assume that we can extend F continuously to a function $F : R^n_+ \to \mathbb{R}$. Say that F is a *potential game* if there exists a continuously differentiable function $f : \mathbb{R}^n_+ \to \mathbb{R}$ such that,

$$\nabla f(x) = F(x) \text{ for all } x \in \mathbb{R}^n_+. \tag{6}$$

This condition is equivalent to:[7]

$$\frac{\partial f}{\partial x_i^p}(x) = F_i^p(x) \text{ for all } i \in S^p, p \in \mathcal{P},\ x \in \mathbb{R}^n_+.$$

A natural question is whether population potential games are weakly acyclic as in the normal form case. It turns out that under some additional regularity conditions the answer is affirmative.

Definition 10. Let Γ be a population potential game and let e be a Nash equilibrium. A Nash equilibrium is *strict* if it uniquely maximizes the potential function f locally in X. That is, there exists a neighbourhood U_e of e in in X such that, $e = \text{argmax}_{x \in U_e} f(x)$.

Definition 11. A Nash equilibrium e is *escapable* if there exists a neighbourhood of e, U_e, such that for every $x \in U_e \setminus \{e\}$ there exists a better-reply path starting at x that goes outside of U_e.

Definition 12. A population potential game Γ is *regular* if their are finitely many Nash equilibria and each Nash equilibrium is either strict or escapable.

[7] In a recent paper Sandholm [18, 2009] provides an alternative notion of potential game that does not rely on the extension of F to all R^n_+. In this paper, for the sake of simplicity, we will use the definition above.

Proposition 2. Every regular potential game is weakly acyclic.

Proof. Let F be a potential game with a potential function f. Let E be the (finite) set of Nash equilibria and let E^c be its complement. Let E^ϵ be the set of points within ϵ of some point in E. We shall show that, for every $x \in (E)^c$ and $\epsilon > 0$, there is a better-reply path to the *interior* of some ϵ-basin of some strict equilibrium.

By Lemma 7.1.1 and Lemma 5.6.5 in Sandholm [20, 2010], the potential function increases along every better-reply path. Let $o(x)$ be the orbit of the point x, that is,

$$o(x) = \cup_{t \geq 0} \Phi(t, x).$$

Let $\gamma = \sup\{f(y) : y \in o(x)\}$ and let ϵ be small enough that,

$$\begin{aligned} &0 < \epsilon < \frac{1}{2}\min\{\mathrm{d}(y, y') : e, e' \in E,\ \text{and } e \neq e'\}. \\ &\forall\ e \in E,\ \mathrm{B}_\epsilon(x) \subset U_e. \end{aligned}$$

By the continuity of F, there exists δ such that for every $y' \in (E^\epsilon)^c$, there exists a population p and strategies $i, j \in S^p$ such that,

$$\forall y,\ \|y - y'\| \leq \delta \Rightarrow F_i^p(y) \geq F_j^p(y) + \delta \text{ and } y_j^p \geq \delta. \tag{7}$$

(If this were false we could construct a sequence of points $\{y^n\} \subset (E^\epsilon)^c$ that converges to a Nash equilibrium, which contradicts the definition of E^ϵ.)

Now let $z : [0, T] \to X$ be a better-reply path starting at x such that $f(z(T)) > \gamma - \delta^4$. We claim first that $z(T) \in U_e$ for some Nash equilibrium e.

Assume by way of contradiction that $z(T) \notin U_e$ for every $e \in E$. In particular, $z(T) \notin E^\epsilon$. Let p be a population and $i, j \in S^p$ two strategies such that Equation (7) holds for $y' = z(T)$. Extend $z(T)$ by letting $\theta_{ji}^p(t) = 1$

in Equation 4 for $t \in [T, T+\delta]$. By definition of population potential game for every $t \in [T, T+\delta]$,

$$\dot{f}(z(t)) = \nabla f(z(t)) \cdot \dot{z}(t) = z_j^p (F_i^p(z(t)) - F_j^p(z(t)))^2.$$

For all $t \in [T, T+\delta]$ the path $z(t)$ stays within δ of $z(T)$, hence by (7) $\dot{f}(z(t)) \geq \delta^3$. It follows that $f(z(T+\delta)) \geq f(z(T)) + \delta^4$, which contradicts the definition of γ.

We can now deduce that $z(T) \in U_e$ for some *strict* Nash equilibrium e. If this were not the case, then e is escapable, hence we can find a better reply path from $z(T)$ to some $y \in (E^\epsilon)^c$. Since the potential increases along any better-reply path we have $f(y) \geq f(z(t)) \geq \gamma - \delta^4$. As before this leads to a contradiction. This completes the proof of the lemma.

$\square$

5 Supermodular Games

Let Γ be a population game. For each population p define $T^p : X^p \to \mathbb{R}^{n^p - 1}$ by

$$T^p(x^p)_i = \sum_{j=i+1}^{n^p} x_j^p.$$

Define $T : X \to \mathbb{R}^{n-m}$ by $T(x) = (T^1(x^1), \ldots, T^m(x^m))$.

Definition 13. A population game Γ is *ingroup neutral* if the payoff to the members of each population depends only on the behaviors of members in the *other* populations, that is,

$$\forall p,\ x \in X,\ j \in S^p,\ F_j^p(x) = F_j^p(x^{-p}).$$

Definition 14. An ingroup neutral population game is *supermodular* if for every p, $i < n^p$, and $x^{-p}, y^{-p} \in X^{-p}$:

$$T^{-p}(y^{-p}) \leq T^{-p}(x^{-p}) \Rightarrow F^p_{i+1}(y^{-p}) - F^p_i(y^{-p}) \leq F^p_{i+1}(x^{-p}) - F^p_i(x^{-p}). \quad (8)$$

It is *strictly* supermodular if the inequality in (8) holds strictly.

Remark 2. A supermodular game has the following property. Suppose that $(i^1, \dots, i^m) \in S$ is a pure strategy profile and that s' is obtained from s by some population p switching to $j^p \neq i^p$. If $j^p > i^p$ then for any population $q \neq p$, q's best reply to s' is nondecreasing relative to q's best reply to s. Similarly if $j^p < i^p$ then q's best-reply to s' is nonincreasing relative to q's best reply to s. This fact will be used repeatedly in the proof of the following proposition.

Proposition 3. Let Γ be an ingroup neutral, strictly supermodular population game with at least two populations such that at every pure strategy profile, each population has a unique best reply. Then Γ is weakly acyclic.

Proof. Let $x \in X$ be an initial state that is not a Nash equilibrium. By ingroup neutrality and strict supermodularity of Γ, one can show that there exists a better-reply path to a state where almost all of the mass of every population is concentrated on a pure strategy profile s. To be specific, let population 1 flow to a state in which almost everyone is playing a pure best reply to x_{-1}, holding the other populations constant. (Note that this construction assumes ingroup neutrality, for otherwise the best reply of population 1 may shift as the distribution of behaviors in this population changes.) Then repeat for each population in succession until the whole population is concentrated on some pure strategy profile $s = (j^1, \dots, j^m) \in S$.

We shall show first that there exists a pure best reply path from s to a strict pure Nash equilibrium. In particular, there exists a sequence of pure strategy profiles $s_1, \ldots, s_k$ such that $s_h = (j_h^1, \ldots, j_h^m)$ for each $1 \leq h \leq k$, and a corresponding sequence of populations $p_1, \ldots, p_{k-1}$ such that for every $1 \leq h \leq k-1$, $s_h^{-p_h} = s_{h+1}^{-p_h}$ and $j_{h+1}^{p_h} = \text{argmax}_i \ F_i^{p_h}(s_h^{-p_h})$. Furthermore, $s^1 = s$ and s^k is a strict Nash equilibrium.

To see this, start with s, and for every population p let $j^p(s)$ be the unique best reply to s. If for every population p, $j^p(s) \geq i^p$, call s a *local minimum*.

If $j^p(s) < i^p$ for some population p let $p_1 = p$ and $s_2 = (j^p(s), s^{-p})$. Since there are finitely many pure strategy profiles, we can repeat this process until we reach a local minimum $r = (i^1, \ldots, i^m)$. If r is a Nash equilibrium then by assumption it is a strict Nash equilibrium and we are done. If r is not Nash equilibrium, one can construct an increasing best reply sequence $r \leq r_2 \leq \ldots \leq r_k$ as follows. Since r is a local minimum, there exists a population p such that $j^p(r) > j^p$. Let $r_2 = (j, r^{-p})$. If r^2 is an equilibrium we are done. Otherwise there exists a population $q \neq p$ such that $j^q(r_2) \neq j^q$. By remark 2, $j^q < j^q(r^2)$. Let $r_3 = (i, r_2^{-i})$. Continuing in this fashion we obtain a monotone sequence $r \leq r_2 \leq r_3 \leq \ldots$. Since the set of pure strategy profiles is finite, this increasing sequence must end in a strict Nash equilibrium.

From this point on the proof is the same as the proof of Proposition 1. Namely, one can construct a better-reply path in Γ that closely mimics any given pure best reply path in G. This completes the proof of Proposition 3. $\square$

6 Fast Convergence

Fix a population game Γ and revision protocol ρ. Let $X^N(t)$ denote the process defined by (2) when the population size is N. Let $E \subseteq X$ be the set of Nash equilibria of Γ, and for each $\epsilon > 0$ let E^ϵ be the set of $x \in X$ such that

$$\inf_{e \in E} \|x - e\| < \epsilon.$$

Define the *indicator function* $\mathbf{1}_{E^\epsilon}(\cdot)$ as follows,

$$\mathbf{1}_{E^\epsilon}(x) = 1 \text{ if } x \in E^\epsilon$$
$$\mathbf{1}_{E^\epsilon}(x) = 0 \text{ if } x \notin E^\epsilon.$$

Let

$$D_N^\epsilon(t) = \frac{1}{t} \int_0^t \mathbf{1}_{E^\epsilon}(X^N(s))ds.$$

This is the proportion of time that the process spends within ϵ of one or more Nash equilibria up to time t.

Definition 15. Given $\epsilon > 0$ the process $X^N(t)$ *converges with precision* ϵ, if for all initial conditions x,

$$\mathbb{P}(\liminf_{t \to \infty} D_N^\epsilon(t) > 1 - \epsilon) > 1 - \epsilon.$$

Note that the precision ϵ refers simultaneously to how close the process is to equilibrium, the proportion of times that the process is close to equilibrium, and also how likely it is that the process is *not* close to equilibrium.

Given $\epsilon > 0$ and $L > 0$ let

$$D_{L,N}^\epsilon(t) = \frac{1}{L} \int_t^{t+L} \mathbf{1}_{E^\epsilon}(X^N(s))ds.$$

Definition 16. Given $\epsilon > 0$ the process $X^N(t)$ exhibits *fast convergence with precision* ϵ if there exists a window of length $L > 0$, depending on ϵ, such that for all sufficiently large N, all initial conditions $x(0)$, and all $t \geq 0$,

$$\mathbb{P}(D_{L,N}^\epsilon(t) > 1 - \epsilon) > 1 - \epsilon.$$

In the case of finite normal form games, weak acyclicity guarantees fast convergence in the following sense. Suppose that players are drawn at random to revise their strategies. Every time a player gets to revise he chooses a better reply given the others' strategies, and he chooses among his better replies uniformly. By assumption there exists a better reply path to a Nash equilibrium. Since the game is finite this path will be followed with positive probability. It follows that given $\epsilon > 0$ there is a time T_ϵ such that, after T_ϵ has elapsed, the probability is at least $1 - \epsilon$ that the process has reached a pure strategy Nash equilibrium.

A similar argument applies to population games that are derived from weakly acyclic normal form games, provided that the population size N is held fixed. However, it need not be the case that the waiting time is bounded as N grows to infinity. We now demonstrate this with an example.

Consider the following family of three player games $\Gamma_{\gamma,\delta}$ where $\gamma, \delta \geq 0$.

Example 3.

$\mathcal{L}$

	L	M	R
T	$-1,-1,1$	$\gamma,0,1$	$0,\gamma,1$
M	$0,\gamma,0$	$-1,-1,1$	$\gamma,0,1$
B	$\gamma,0,1$	$0,\gamma,1$	$-1,-1,1$

$\mathcal{R}$

	L	M	R
T	$0,0,0$	$0,0,0$	$0,0,0$
M	$2,2,\delta$	$0,0,0$	$0,0,0$
B	$0,0,0$	$0,0,0$	$0,0,0$

Population 1 plays row, population 2 plays column, and population 3 chooses between the two matrices $\mathcal{L}$, $\mathcal{R}$.

Proposition 4.

(i) For every $\gamma > 0$ and $\delta > 0$, the population game $\Gamma_{\gamma,\delta}$ is weakly acyclic.

(ii) For every $\epsilon > 0$, every $N > 0$, and every starting point x,

$$\mathbb{P}(\liminf_{t\to\infty} D_N^{\epsilon}(t) = 1) = 1.$$

(iii) Given a sufficiently small $\epsilon > 0$ there exists $\gamma, \delta > 0$ and starting point y, such that for every finite t,

$$\mathbb{P}(D_N^{\epsilon}(t) = 0) > 1 - \epsilon \text{ for all sufficiently large } N. \tag{9}$$

Note that (9) implies a failure of fast convergence in a particularly strong sense. Namely, the process stays bounded away from all the Nash equilibria for arbitrarily long periods of time when N is arbitrarily large.

Proof. It can be verified that the underlying normal form game is weakly acyclic. Indeed from any pure strategy-tuple there exists a better-reply path to the unique pure Nash equilibrium $(M, L, \mathcal{R})$. Furthermore, even though the game is not generic, it can be checked that the proof of Proposition 1 remains valid for this case. This establishes the first claim.

To establish the second claim, note that for every fixed N and initial state $y \in \chi^N$, there exists a time T such that with positive probability the process will arrive at the unique pure Nash equilibrium by time T. Since the number of states χ^N is finite and the equilibrium is absorbing, claim (ii) holds.

To establish claim (iii), we begin by showing that we can choose γ and δ, and a starting point $y \in X$, such that the orbit of y has the following properties.

$$(x^1, x^2, x^3) \in o(y) \Rightarrow x^3 = (1, 0) \tag{10}$$

$$(x^1, x^2, x^3) \in o(y) \Rightarrow x^1 \cdot x^2 \leq 1/100. \tag{11}$$

These statements imply in particular that the deterministic dynamic (starting from y) stays more that .20 away from any Nash equilibrium. Indeed the unique Nash equilibrium when population 3 plays $\mathcal{L}$ is:

$$e = ((1/3, 1/3, 1/3), (1/3, 1/3, 1/3), \mathcal{L}).$$

However, $x^1 \cdot x^2 \leq 1/100$ implies that for at least one p and j, $x_j^p \leq 1/10$, hence $\|x - e\| \geq 1/3 - 1/10 > .20$. To establish the existence of a state y with these properties, let us consider the sub-game for players 1 and 2 when player 3 is held fixed at $\mathcal{L}$.[8]

	L	M	R
T	$-1, -1$	$\gamma, 0$	$0, \gamma$
M	$0, \gamma$	$-1, -1$	$\gamma, 0$
B	$\gamma, 0$	$0, \gamma$	$-1, -1$

For each $0 \leq \gamma < 1$ let $\phi^\gamma : [0, \infty) \times X \to X$ be the semi-flow for the two-player game derived from the revision protocol ρ.

A compact set $\mathcal{A} \subseteq X$ is an *attractor* for the semi-flow ϕ if it is *invariant* that is, $\forall\, t \geq 0$, $\Phi_t(\mathcal{A}) = \mathcal{A}$, and there is a neighbourhood $\mathcal{U}$ of $\mathcal{A}$ such that,

$$\lim_{t \to \infty} \sup_{x \in \mathcal{U}} d(\phi_t(x), \mathcal{A}) = 0.$$

Let $\mathcal{A} \subset X$ be the set of states where the diagonal strategy combinations have mass zero that is,

$$\mathcal{A} = \{(x^1, x^2) : x_j^1 \cdot x_j^2 = 0, \text{ for } j = 1, 2, 3\}.$$

$\mathcal{A}$ is an attractor for the semi-flow ϕ^0.

By theorem 9.B.5 in Sandholm [20, 2010] for all sufficiently small $\gamma > 0$, there exists an attractor $\mathcal{A}^\gamma$ of ϕ^γ such that $\mathcal{A}^0 = \mathcal{A}$ and the map $\gamma \to \mathcal{A}^\gamma$

[8]One may verify that for small enough γ and δ there are only two Nash equilibria; the mixed equilibrium e, and the pure equilibrium, $(M, L, \mathcal{L})$.

is upper-hemicontinuous. It follows that for all sufficiently small $\gamma > 0$, all elements $(x^1, x^2) \in \mathcal{A}^\gamma$ have the property that $x^1 \cdot x^2 \leq 1/100$.

Fix such a value $\gamma > 0$. From the properties of the revision protocol ρ it follows that among all states in $\mathcal{A}^\gamma$ the mass on the strategy pair (M, L) is at most $1 - 2\tau$ for some $\tau > 0$.

Choose $\delta \equiv \tau$ in the definition of the matrix $\mathcal{R}$. (We have not relied on the value of δ up to this point.) Then for *every* state $(x^1, x^2) \in \mathcal{A}^\gamma$ population 3 strictly prefers matrix $\mathcal{L}$ to matrix $\mathcal{R}$.

Now let $y = (y^1, y^2, y^3)$, where $y^3 = (1, 0)$ and puts all the mass on $\mathcal{L}$. By construction the full three-population game is such that the flow from y stays in the set $A^\gamma \times (1, 0)$ hence both (10) and (11) holds.

The proof of claim (iii) is now completed as follows. Fix a finite time T and let $\epsilon < 1/10$. By Lemma 1 in Benaïm and Weibull [3, 2003], the stochastic process $X^N(t)$ stays within ϵ of the path of the deterministic dynamic with probability at least $1 - \epsilon$ over the period $[0, T]$ provided that N is sufficiently large (given ϵ and T). Therefore, starting from the state y identified above, the probability is at least $1 - \epsilon$ that $X^N(t)$ is at least $.2 - \epsilon > \epsilon$ away from any Nash equilibrium in $\Gamma_{\gamma,\delta}$ over the *entire* period $[0, T]$. This establishes (9), and completes the proof of the proposition. $\square$

7 Environmental Interference

We now introduce an additional feature of the evolutionary process that leads to fast convergence with arbitrarily high precision. The idea is that players do not learn about the payoff of other strategies with *certainty*, but rather with some probability that is determined by the environment. In particular we shall assume that the switching rate between two strategies is determined not only by their payoffs, but also by the probability that a player *knows* these payoffs.

To be specific, suppose that the i-players currently have payoff $\pi_i^p(t)$ and the j-players have payoff $\pi_j^p(t)$. If an i player has a revision opportunity, he switches to j with probability proportional to $\alpha_{ji}^p(t)\rho_{ji}(\pi^p(t))$, where $\rho_{ji}^p(\pi^p(t))$ is the usual switching rate (determined by the current payoffs and the protocol), and $\alpha_{ji}^p(t) \in [0,1]$ to the probability that i "hears about" the payoffs of the j-players.[9] This latter probability is a feature of the environment, which evolves according to a separate stochastic process to be discussed below.

To take a concrete example, Smith [23, 1984] originally proposed his dynamic as a model of route choices by drivers. The idea is that a driver switches from route i to route j at time t with a probability proportional to the current payoff difference $(\pi_j^p(t)-\pi_i^p(t))$. In our version of the model , the driver switches from i to j with a probability proportional to $\alpha_{ji}(t)(\pi_j^p(t) - \pi_i^p(t))$ where $\alpha_{ji}(t)$ is the probability he hears about $\pi_j(t)$ at time t, say through a radio broadcast.

We can formalize these ideas as follows. The *environment* at time t is described by a vector of matrices $\vec{A} = (A^1(t), \dots, A^m(t))$ where $A^p(t) = (\alpha_{ij}^p(t))_{ij\in S^p}$ represents the *accessibility* of information about the j-players in population p by the i-players (we assume that $\alpha_{ii} \equiv 1$). The environment evolves over time according to a Poisson random walk with a constant step size. To be specific, each $\alpha_{ij}^p(t)$ is a Poisson random variable with arrival rate a. At each arrival α_{ij}^p increases by $\frac{1}{\sqrt{a}}$ with probability one-half (provided $\alpha_{ij}^p+\frac{1}{\sqrt{a}} \leq 1$) and decreases with probability one-half (provided $\alpha_{ij}^p-\frac{1}{\sqrt{a}} \geq 0$). The state of the process can be written in the compact form as $(\vec{A}, x)$. Given

[9] Recall that, by assumption, ρ_{ji}^p actually depends only on π_j^p and π_i^p.

a state $(\vec{A}, x)$ the new switching rates are determined as follows:

$$\bar{\rho}^p_{ij}(\pi^p, \alpha^p) = \alpha^p_{ij} \cdot \rho^p_{ij}(\pi).$$

Let λ be as in equation (1). Revision opportunities by some member of the population arrive at the rate $\lambda^N = N\lambda + ca$. Assume that the current state of the population is $(\vec{A}, x)$. Whenever a revision opportunity arrives, the process moves to a new state $(\vec{B}, y)$ according to the following probabilities:

$$\begin{cases} \frac{N \cdot x^p_i \rho^p_{ij}(F^p(x))\alpha^p_{ij}}{\lambda^N} & \text{where } \beta = \alpha, \ y = x + \frac{1}{N}(e^p_j - e^p_i) \\ \frac{a}{2\lambda^N} & \text{where } y = x, \ \beta - \alpha = e^p_{ij} \cdot \min\{\frac{1}{\sqrt{a}}, [1 - (\alpha^p_{ij} + \frac{1}{\sqrt{a}})]_+\} \\ \frac{a}{2\lambda^N} & \text{where } y = x, \ \beta - \alpha = -e^p_{ij} \cdot \max\{\frac{1}{\sqrt{a}}, [\alpha^p_{ij} - \frac{1}{\sqrt{a}}]_+\}. \end{cases} \tag{12}$$

Equation 12 has the following simple interpretation. With a probability of $\frac{N \cdot x^p_i \rho^p_{ij}(F^p(x))\alpha^p_{ij}}{\lambda^N}$ a member of population p playing strategy i switches to strategy j. With probability $\frac{a}{2\lambda^N}$ the variable α^p_{ij} increases by $\frac{1}{\sqrt{a}}$, and with $\frac{a}{2\lambda^N}$ it decreases by $\frac{1}{\sqrt{a}}$ provided that the constraints $0 \leq \alpha^p_{ij} \leq 1$ are satisfied. With the complementary probability the process stays in the state $(\vec{A}, x)$.

Denote this process by $(\vec{A}(t), X^N_a(t))$. Our main theorem states that if the environment evolves in sufficiently small steps, we obtain fast convergence.

Theorem 1. For every weakly acyclic population game Γ, revision protocol ρ, and $\epsilon > 0$, for all sufficiently large $a > 0$, $X^N_a(t)$ exhibits fast convergence with precision ϵ.

Remark 3. We have assumed that the environment evolves according to a Poisson random walk, but in fact the theorem holds for many other stochastic models of the environment. For example, one could assume that each α^p_{ij} changes according to a Poisson process, and that every time a change occurs α^p_{ij} is drawn uniformly from the interval $[0, 1]$.

Remark 4. Let us recall Example 3, where we showed that convergence is slow when the process follows a revision protocol with no environmental interference. When there is environmental interference the process is kicked out of the cycle. This occurs because there is now a positive probability of being in a state where the strategy-tuple $(M, L, \mathcal{L})$ has mass larger than $1 - \delta$. When this happens population 3 is better-off switching to the right-hand matrix, and the process converges rapidly to the Nash equilibrium $(M, L, \mathcal{R})$.

Remark 5. Fast convergence means that the waiting time to come close to equilibrium is bounded above for all population sizes N. It leaves open the question of how long the waiting time actually is; this may depend on the details of the revision protocol. Our result also says nothing about the time it takes to reach a particular equilibrium, such as a potential-maximizing equilibrium in a potential game. This issue examined by Shah and Shin [22, 2010].

8 Informal Sketch of the Proof

Although the proof of this theorem is rather technical, the main ideas are quite intuitive. Here we shall sketch the general gist of the argument.

Given $\epsilon > 0$, let us say that a state is "good" if it lies within ϵ of a Nash equilibrium, and "very good" if it lies in the interior of the ϵ-basin of some Nash equilibrium. Otherwise the state is "bad." Starting from any state y, we know that there exists a better-reply path to a very good state. Lemmas 1-4 show that with positive probability there is a realization of the environment such that the expected motion of the process is very close to this path. Given such a realization the process follows a nearby path with high probability when N is sufficiently large (Lemma 4).

It then follows that there is a time T^y such that, if the process $X_a^N(t)$ starts close enough to y and N is large enough, there is a positive probability that by time $T^y + t$ the process will be in a very good state. Using compactness arguments, one can show that there is a T such that this statement holds for all y. Moreover, once the process is in the neighbourhood of a very good state, it is very improbable that it will leave for a long period of time (much longer than T) provided N is large enough.

The formal proof of Theorem 1 relies on a series of technical lemmas. The first of these states that any better-reply path can be approximated by a better-reply path with a continuous Lipschitz selection of θ (recall equation 4).

Lemma 1. Let $z : [0,T] \to X$ be a better-reply path and let ρ be a revision protocol. For every $\epsilon > 0$ there exists a Lipschitz continuous function $K : [0,T] \to [0,M]^d$ and a better-reply path $w : [0,T] \to X$ such that, for all $t \in [0,T]$, p, and $i \in S^p$:

$$\dot{w}_i^p = \sum_{j \in S^p} K_{ji}^p \rho_{ji}^p(F^p(y))w_j^p - K_{ij}^p \rho_{ij}^p(F(w))w_i^p, \tag{13}$$

and

$$\sup_{0 \leq t \leq T} \|z(t) - w(t)\| < \epsilon. \tag{14}$$

Let $h(t,x) : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ be a piecewise continuous function in t that is Lipschitz with respect to x with Lipschitz constant ν. Let $g(t,x) : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ be a piecewise continuous function that is Lipschitz with respect to x. By Picard's Theorem, along the interval $[0,T]$, for any $z_0, y_0 \in \mathbb{R}^n$ there are a unique solutions to the differential equations

$$\dot{z}(t) = h(t, z(t)),\ z(0) = z_0,$$
$$\dot{y}(t) = g(t, y(t)),\ y(0) = y_0.$$

The following Lemma provides a standard approximation to the distance between these two solutions, as a function of the initial conditions, and the distance between the functions h and g.

Lemma 2. If $\|h(t,x) - g(t,x)\|_\infty < \beta$ and $\|z_0 - y_0\| < \delta$ then

$$\sup_{t\in[0,T]}\|z(t) - y(t)\| < (\delta + \beta T)\exp(\nu T).$$

The following Lemma shows that if the step size $\frac{1}{\sqrt{a}}$ defining the evolution of the environment is sufficiently small, then the environment follows any pre-specified continuous path very closely with positive probability.

Lemma 3. Let $K : [0,T] \to [0,1]^d$ be a continuous path, and let $\alpha(t)$ evolve according to a Poisson random walk with arrival rate a. For every $\gamma > 0$ there exists $\delta > 0$ and a_0 such that for every $a > a_0$:

$$\mathbb{P}(\sup_{\gamma\leq s\leq T+\gamma}|\alpha(s+\gamma) - K(s)| < \gamma : \alpha(0)) > \delta.$$

Given a population of size N, let $\alpha : [0,T] \to \mathbb{R}^d$ be a piecewise constant, right-continuous realization of the Poisson random walk (with arrival rate a). Let $X_a^N(t)$ be the stochastic process defined by ρ and α beginning in state $x_0 \in \chi^N$. Let $0 = \tau_0 < \tau_1 < \ldots < \tau_k$ be the sequence of times at which one or more environmental variables α_{ij}^p change for some (p,i,j). Between these distinct times the α's remain fixed. Let $z(t)$ be the unique solution of the following differential equation:

$$\dot{z}_i^p(t) = \sum_{j\in S^p} z_j^p \rho_{ij}^p(F^p(z))\alpha_{ij}^p - z_i^p \rho_{ji}^p(F^p(z))\alpha_{ij}^p. \tag{15}$$

Let $Z(t)$ be the stochastic process defined by equation 15. The following Lemma is a variant of Lemma 1 in Benaïm and Weibull [3] (see also Kurtz [15]) and shows that, given a realization of the environment, the stochastic process $X_a^N(t)$ can be approximated by $Z(t)$ provided that N sufficiently large.

Lemma 4. Let $(\alpha(t))_{0\leq t\leq T}$ be any realization of the environment with $c(T) \leq k$ jumps along $[0,T]$. For every $\epsilon > 0$, there exists $N_{T,k,\epsilon}$ such that,

$$\forall N > N_{T,k,\epsilon},\ \mathbb{P}_x(\sup_{t\in[0,T]}\|X_a^N(t) - Z(t)\| > \epsilon : (\alpha(t))_{0\leq t\leq T}) < 1-\epsilon, \quad (16)$$

Furthermore,

$$\forall N > N_{T,a,\epsilon},\ \mathbb{P}_x(\sup_{t\in[0,T]}\|X_a^N(t) - Z(t)\| > \epsilon) < 1-\epsilon, \quad (17)$$

9 Proof of The Main Theorem

Fix $\epsilon > 0$ and recall that E^ϵ is the set of points within ϵ of some Nash equilibrium. Let $y \in (E^\epsilon)^c$. By definition of weakly acyclicity there exists a better-reply path $z^y : [0,T^y] \to X$ and an equilibrium x such that $z(T^y) \in \mathrm{int}(\mathcal{BA}_\epsilon(x))$. By Lemma 1 we may assume that there exists a Lipschitz continuous function $K^y : [0,T] \to [0,1]^d$ such that

$$\forall\ p,\ \dot{z}_i^{y,p} = \sum_{j\in S^p} z_j^{y,p}\rho_{ji}^p(F^p(z^y))K_{ji}^{y,p} - z_i^{y,p}\rho_{ij}^p(F(z^y))K_{ij}^{y,p}.$$

It follows that there exists $\delta^y > 0$ such that $\mathrm{B}_{\delta^y}(z^y(T^y)) \subseteq \mathcal{BA}_\epsilon(x)$. (In general $\mathrm{B}_r(x)$ denotes the ball of radius r centered at x.)

Let $\alpha : [0,T^y] \to [0,1]^d$ be any realization of the environment, which is piecewise constant, hence piecewise Lipschitz continuous. Let w be the solution for the following differential equation starting at y':

$$\forall\ p,\ \dot{w}_i^p = \sum_{j\in S^p} w_j^p\rho_{ji}^p(F^p(y))\alpha_{ji}^p - w_i^p\rho_{ij}^p(F(w))\alpha_{ij}^p.$$

By Lemma 2, if y' is close enough to y and the realization of the environment, α, is close enough to K^y, then the distance between $w(T^y)$ and $z^y(T^y)$ is smaller than δ^y. That is, there exists a small $\theta^y > 0$ such that,

$$\begin{aligned} &\sup_{0\le t\le T^y}\|\alpha(t) - K^y(t)\| < 2\theta^y \text{ and } \|y - y'\| < 2\theta^y, \\ &\Rightarrow \|w(T^y) - z(T^y)\| < \delta^y. \end{aligned} \tag{18}$$

Assume that the current population state at time t_0 lies within θ^y of y. It follows from Lemma 3 that for every $\gamma > 0$ there exists a^y (which represents the arrival rate of the environment) such that for all $a > a^y$, $\alpha(t)$ is close to $K^y(t-(t_0+\gamma))$ with positive probability, along the interval $[t_0+\gamma, t_0+T^y+\gamma]$. By Lemma 4, one can choose γ^y such that for every for every $N > N^y(a, \gamma^y)$, $X_a^N(t_0 + T^y + \gamma^y)$ lies within δ^y of $z(T^y)$, with probability greater than some $\epsilon^y > 0$. This may be summarized as follows.

Lemma 5. There exists small enough $\gamma^y > 0$, large enough a^y, and $\epsilon^y > 0$ such that for every $a > a_y$ there exists $N^y(a, \gamma^y)$ such that

$$\forall\, N > N^y,\ \mathbb{P}(X_a^N(t_0+\gamma^y+T^y) \in B_{\delta^y}(z^y(T^y)) \mid \|X_a^N(t_0)-y\| < \theta^y) > \epsilon^y \tag{19}$$

Proof. See Appendix. □

We show next that there exists a large T and a small $\epsilon' > 0$, such that within time T, the process reaches an interior point of an ϵ-basin of some equilibrium.

Note that $\{B_{\theta^y}(y)\}_{y\in(E^\epsilon)^c}$ is a cover of the compact set $(E^\epsilon)^c$. By compactness there exists a finite set $\{y_1, \ldots y_l\}$ such that,

$$(E^\epsilon)^c \subseteq \bigcup_{m=1}^{l} B_{\theta^{y_m}}(y_m).$$

Let $\mathcal{M} = \bigcup_{m=1}^{l} \mathrm{B}_{\delta^{y_m}}(z^{y_m}(T^{y_m}))$ and let $\mathcal{L} = \cup_{x\in E}\mathrm{int}(\mathcal{B}_\epsilon(x))$ be the set of interior points of all ϵ-basins of some equilibrium. By construction, $\mathcal{M} \subset \mathcal{L}$.

In particular, $\mathcal{M} \subset E^\epsilon$. Let

$$T = \max\{T^{y_1}, \dots, T^{y_l}\},$$

$$\epsilon' = \min\{\epsilon^{y_1}, \dots, \epsilon^{y_l}\},$$

$$a > \max\{a^{y_1}, \dots, a^{y_l}\}.$$

Let

$$N > \max\{N^{y_1}(a, \gamma^{y_1}), \dots, N^{y_l}(a, \gamma^{y_l})\}.$$

Given $y \notin E^\epsilon$ it follows from Lemma 5 that

$$\mathbb{P}(\exists t \le T \text{ s.t. } X_a^N(t_0 + t) \in \mathcal{M} \mid X_a^N(t_0) = y) > \epsilon'. \tag{20}$$

Let $Z(t)$ be the stochastic process defined by (15) starting at $x \in \mathcal{M}$. By construction, $Z(t) \in \mathcal{L}$ for all $t > 0$. By (17),

$$\mathbb{P}(\sup_{0 \le t \le T} \|X_a^N(t_0 + t) - Z(t)\| > \epsilon \mid X_a^N(t_0) = x \in \mathcal{M}) \tag{21}$$

goes to 0 uniformly in x when N grows. Theorem 1 now follows from equations (20) and (21).

Appendix

We shall use the following form of Grönwall's inequality:

Lemma 6 (Grönwall's Inequality)**.** Let $I = [0, T]$ denote an interval of the real line. Let α, β and u be real-valued functions defined on I. Assume that β and u are continuous and that α is non-decreasing and is integrable on every closed subinterval of I. If β is non-negative and if u satisfies the integral inequality

$$u(t) \leq \alpha(t) + \int_0^t \beta(s)u(s)\mathrm{d}s \quad \forall\ t \in I \tag{22}$$

then,

$$u(t) \leq \alpha(t)\exp(\int_0^t \beta(s)u(s)\mathrm{d}s) \quad \forall\ t \in I. \tag{23}$$

Proof of Lemma 1. Since $\dot{z}(t)$ integrable, there exists a measurable selection $\theta : [0, T] \to [0, M]^d$ such that,

$$\dot{z}_i^p = \sum_{j \in S^p} z_j^p[F_i^p(z) - F_j^p(z)]_+\theta_{ji}^p - \sum_{j \in S^p} z_i^p[F_j^p(z) - F_j^p(z)]\theta_{ij}^p.$$

Let $C : [0, T] \to [0, M]^d$ be a Lipschitz ϵ-approximation of θ in L^1, that is,

$$\forall p,\ i, j \in S^p \quad \int_0^T |\theta_{ij}^p(t) - C_{ij}^p(t)|\mathrm{d}t < \epsilon.$$

Let $h : [0, M] \times X \to TX$ be the following function:

$$h_i^p(t, y) = \sum_j y_j^p[F_i^p(y) - F_j^p(y)]_+C_{ji}^p - \sum_j y_i^p[F_j^p(y) - F_i^p(y)]_+C_{ij}^p, \tag{24}$$

and let $y : [0, T] \to X$ be the solution for $h(t, y)$ starting at $z(0)$. We shall show that $y(t)$ is an approximation of $z(t)$ in the interval $[0, T]$. Letting

$u(t) = \|z(t) - y(t)\|$, we have:

$$\begin{aligned} u(t) &= \|z(t) - y(t)\| \\ &\leq \int_0^t \|\dot{z}(s) - h(s,y)\|\mathrm{d}s \\ &\leq \int_0^t \|\dot{z}(s) - h(s,z)\|\mathrm{d}s + \int_0^t \|h(s,z) - h(s,y)\|\mathrm{d}s. \end{aligned} \quad (25)$$

Let us evaluate the right-hand side of (25) first:

$$\begin{aligned} & \int_0^t |h_i^p(s,z) - h_i^p(s,y)|\mathrm{d}s \\ \leq & \int_0^t |\sum_j (z_j^p[F_i^p(z) - F_j^p(z)]_+ - y_j^p[F_i^p(y) - F_j^p(y)]_+)C_{ji}^p|\mathrm{d}s \quad (26) \\ + & \int_0^t |\sum_j z_i^p[F_j^p(z) - F_i^p(z)]_+ - y_i^p[F_j^p(y) - F_i^p(y)]_+ C_{ij}^p|\mathrm{d}s. \quad (27) \end{aligned}$$

As for (26) we have:

$$\begin{aligned} \leq & M \cdot \int_0^t \sum_j |z_j^p[F_i^p(z) - F_j^p(z)]_+ - y_j^p[F_i^p(y) - F_j^p(y)]_+|\mathrm{d}s \\ \leq & M \cdot \int_0^t \sum_j |z_j^p[F_i^p(z) - F_j^p(z)]_+ - z_j^p[F_i^p(y) - F_j^p(y)]_+|\mathrm{d}s \\ + & M \cdot \int_0^t \sum_j |z_j^p[F_i^p(y) - F_j^p(y)]_+ - y_j^p[F_i^p(y) - F_j^p(y)]_+|\mathrm{d}s \\ \leq & M \cdot \int_0^t D\|z(s) - y(s)\| + 2\|F\|\|z(s) - y(s)\|\mathrm{d}s \\ = & \int_0^t M(D + \|F\|)\|z(s) - y(s)\|\mathrm{d}s, \end{aligned}$$

where D is the maximal Lipschitz constant for the functions $z \to [F_i^p(z) - F_j^p(z)]_+$. A similar approximation may be provided for (27). Therefore,

$$\int_0^t \|h(s,z) - h(s,y)\|\mathrm{d}s \leq 2\int_0^t M(D + \|F\|)\|z(s) - y(s)\|\mathrm{d}s.$$

Turning to the left-hand side of (25) one can see that

$$\int_0^t \|\dot{z}(s) - h(s,z)\|\mathrm{d}s \leq 2\|F\|nT\epsilon.$$

It follows from Grönwall's inequality (23) that:

$$u(t) \leq 2\|F\|nT\epsilon \exp(\int_0^t M(D+\|F\|)u(s)\mathrm{d}s).$$

Since the integral is bounded, $y(t)$ approximates $z(t)$ for small ϵ.

In order to complete the proof of Lemma 1 we are going to show that every better-reply path y with Lipschitz selection C^p_{ji} can be approximated by a function w which has the desired properties (13)-(14). Define

$$h_i^{p,\delta}(t,y) = \sum_j y_j^p[[F_i^p(y)-F_j^p(y)]_+-\delta]_+C^p_{ji}-\sum_j y_i^p[[F_j^p(y)-F_i^p(y)]_+-\delta]_+C^p_{ji}. \tag{28}$$

Note that $h^\delta(t,y)$ converges uniformly to the function $h(t,y)$ defined in (24). Hence for any $\epsilon > 0$ there exists a δ_0 such that for every $0 < \delta < \delta_0$,

$$\sup_{t\leq T}\|y(t)-w(t)\| < \epsilon,$$

where w is a solution of (28) starting at $y(0)$. Given any $\delta < \delta_0$ we have

$$\dot{w}_i^p = \sum_j w_j^p[[F_i^p(w)-F_j^p(w)]_+-\delta]_+C^p_{ji}-\sum_j w_i^p[[F_j^p(w)-F_i^p(w)]_+-\delta]_+C^p_{ji}.$$

Define[10]

$$K^p_{ji}(t) = \frac{[[F_i^p(w(t))-F_j^p(w(t))]_+-\delta]_+C^p_{ji}(t)}{\rho^p_{ji}(F^p(w(t)))}.$$

It follows that,

$$\dot{w}_i^p = \sum_{j\in S^p} K^p_{ji}\rho^p_{ji}(F^p(w))w_j^p - K^p_{ij}\rho^p_{ij}(F(w))w_i^p.$$

Since $\rho^p_{ji}(F^p)$ is continuous and positive wherever $F^p_i > F^p_j$ it follows that $\rho^p_{ij}(F^p)$ is bounded away from zero whenever $F^p_i \geq F^p_j - \delta$. Note that the numerator is zero whenever $F^p_i < F^p_j - \delta$. This completes the proof of Lemma 1. $\square$

[10] We use the convention that $\frac{0}{0} = 0$.

Proof of Lemma 2. Assume that $\|h(t,x) - g(t,x)\|_\infty < \delta$ and $\|z_0 - y_0\| < \delta$. Let $u(t) = \|z(t) - y(t)\|$ for every $t \in [0,T]$. Then

$$\begin{aligned}
u(t) &= \|x_0 - y_0 + \int_0^t h(s,z(s)) - g(s,y(s))\mathrm{d}s\| \\
&= \|x_0 - y_0\| + \int_0^t \|h(s,z(s)) - g(s,y(s))\mathrm{d}s\| \\
&\leq \delta + \int_0^t \|h(s,z(s)) - f(s,y(s))\| + \|h(s,y(s)) - g(s,y(s))\|\mathrm{d}s \\
&\leq \delta + \int_0^t (\nu\|z(s) - y(s)\| + \beta)\mathrm{d}s \qquad (29) \\
&\leq \delta + \beta t + \int_0^t \nu u(t)
\end{aligned}$$

Hence, by Grönwall's inequality $u(t) \leq (\delta + \beta t)\exp(\nu t)$. □

Proof of Lemma 3. For simplicity, choose $d = 1$. (Similar arguments hold for any integer d.) Let τ be a Poisson process with an arrival rate a, and let $\{\xi_n\}_{n=1}^\infty$ be a sequence of i.i.d random variables with mean zero and variance 1. Let $\chi^a(t) = \frac{1}{\sqrt{a}}\sum_{n=1}^{\tau(t)} \xi_n$. By Theorem 7.1 in Durrett [5, 1996] (with $\mathbf{a} = 1$ and $\mathbf{b} = 0$) we have

$$\chi^a(t) \Rightarrow_{a\to\infty} B(t) \text{ in probability,} \qquad (30)$$

where $B(t)$ is a one dimensional Brownian motion that starts at the origin.

Assume first that $K(0) = 0$. Theorem 2.10 in Durrett [5, 1996] implies that for every $\epsilon > 0$ there is a positive probability that the maximal difference $B(t) - K(t)$ on the interval $[0,T]$ is smaller than ϵ:

$$\exists\ \delta > 0 \ \ \mathbb{P}(\sup_{0\leq t\leq T}|K(s) - B(s)| < \epsilon) > \delta. \qquad (31)$$

We claim that, given any $\epsilon > 0$, any $f : [0,1] \to [0,1]$, and any $0 \leq x \leq 1$, there exists $\delta' > 0$ such that

$$\mathbb{P}(\sup_{t_0+\epsilon\leq t\leq T+t_0+\epsilon}|K(t-t_0-\epsilon) - B(t)| < \epsilon \text{ and } B(t) \in (0,1) : B(t_0) = x) > \delta'.$$

To see this, consider the function that is linear between x and $K(0)$ on the interval $[t_0, t_0+\epsilon]$, and that agrees with $K(t-(t_0+\epsilon))$ on $[t_0+\epsilon, T+t_0+\epsilon]$.

Let τ be the first point such that $\chi^a(t) \in \{0,1\}$. The result now follows from (30), (31) and the fact that $\alpha(t)$ has the same distribution as $\chi^a(t)$ on the interval $[0,\tau]$. □

Proof of Lemma 4. Lemma 1 in Benaïm and Weibull implies the following:

Lemma 7. Let ρ be a revision protocol for the game F and, let $X^N(t)$ be the pure jump Markov process corresponding to ρ and F. Let $\xi(t,x)$ be the flow of the differential equation defined by (3), and let

$$D^N(T,x) = \max_{0\le t\le T} \|X^N(t) - \xi(t,x)\|_\infty.$$

There exists a scalar $c(T)$ such that for any $\epsilon > 0$, $T > 0$ and $N > \frac{\exp(\nu T)\nu T}{\epsilon}$:

$$\mathbb{P}_x(D^N(T,x) \ge \epsilon) \le 2n\exp(-\epsilon^2 c(T) N), \tag{32}$$

and,

$$c(T) = \frac{\exp(-2BT)}{8TA}.$$

(Here A and B are constants that depend on $\|\rho\|_2^2$ the Lipschitz constant of ρ, and the Lipschitz constant of the payoff function F.)

Let $\alpha : [0,T] \to [0,1]^d$ be any realization of the environment and let $(\tau_1, \ldots, \tau_k)$ be the sequence of distinct times at which the environment changed. Note first that on $[\tau_l, \tau_{l+1})$ the process $X_a^N(t)$ is distributed according to the stochastic process generated by the revision protocol $\bar{\rho}$ defined by,

$$\forall\ p,\ \forall\ i,j \in S^p,\ \bar{\rho}^p_{ij} = \rho^p_{ij}\alpha^p_{ij}(\tau_l),$$

with initial condition $X_a^N(\tau_l)$. Note also that for any realization α of the environment, the protocol $\bar{\rho}$ has a Lipschitz constant that is no higher than the Lipschitz constant for ρ.

Let $s(t)$ be the piecewise continuous process that is defined as follows:

$$\dot{s}_i^p(t) = \sum_{j \in S^p} s_j^p(t)\rho_{ji}^p(F^p(s))\alpha_{ji}^p - \sum_{j \in S^p} s_i^p(t)\rho_{ij}^p(F^p(s))\alpha_{ij}^p,$$

and $\forall\ 1 \leq l \leq k$ let $s(\tau_l) = X_a^N(\tau_l)$. Let $c(t) = \max\{k : \tau_k < t\}$, and note that $c(T) = k$. Let $\tau_{k+1} = T$. We then have,

$$\begin{aligned} & \sup\nolimits_{0 \leq t \leq T} \|X_a^N(t) - z(t)\| && (33) \\ \leq\ & \sup\nolimits_{0 \leq t \leq T} [\|X_a^N(t) - s(t)\| + \|s(t) - z(t)\|] \\ =\ & \sum_{l=1}^{k+1} \sup\nolimits_{\tau_{l-1} \leq t \leq \tau_l} [\|X_a^N(t) - s(t)\| + \|s(t) - z(t)\|]. \end{aligned}$$

Hence

$$\begin{aligned} & \mathbb{P}(\sup\nolimits_{0 \leq t \leq T} \|X_a^N(t) - z(t)\| > \epsilon : (\alpha(t))_{0 \leq t \leq T}) && (34) \\ \leq\ & \sum_{l=1}^{k+1} \mathbb{P}(\sup\nolimits_{\tau_{l-1} \leq t \leq \tau_l} \|X_a^N(t) - s(t)\| > \epsilon : (\alpha(t))_{0 \leq t \leq T}) \\ +\ & \mathbb{P}(\sup\nolimits_{\tau_{l-1} \leq t \leq \tau_l} \|s(t) - z(t)\| > \epsilon : (\alpha(t))_{0 \leq t \leq T}). && (35) \end{aligned}$$

$\sum_{l=1}^{k+1} \mathbb{P}(\sup_{\tau_{l-1} \leq t \leq \tau_l} \|X_a^N(t) - s(t)\| > \epsilon : (\alpha(t))_{0 \leq t \leq T})$ can be evaluated using Lemma 7 and goes to zero uniformly in N for any realization α with k jumps or less.

We claim that $\sum_{l=1}^{k+1} \mathbb{P}(\sup_{\tau_{l-1} \leq t \leq \tau_l} \|s(t) - z(t)\| > \epsilon : (\alpha(t))_{0 \leq t \leq T})$ goes to zero in N note that by Lemma 2,

$$\begin{aligned} & \sup\nolimits_{\tau_{l-1} \leq t \leq \tau_l} \|s(t) - z(t)\| && (36) \\ \leq\ & \exp(\nu(\tau_l - \tau_{l-1}))\|s(\tau_{l-1}) - z(\tau_{l-1})\|. && (37) \end{aligned}$$

Thus inductively one has,

$$\begin{aligned} \sup\nolimits_{0 \leq t \leq T} \|s(t) - z(t)\| &\leq \sum_{l=2}^{k+1} \exp(\nu\tau_l)\|s(\tau_1) - z(\tau_1)\| \\ \leq\ & k\exp(\nu T)\|X_a^N(\tau_1) - z(\tau_1)\|, \end{aligned}$$

which goes to zero uniformly in N. This establishes (16) of Lemma 4. Equation (17) follows from (16) and the fact that $E(c(T)) = aT < \infty$. $\square$

Proof of Lemma 5. By Lemma 3 for every $\gamma > 0$ there exists $a^y(\gamma)$ and $\delta' > 0$ such that, for every $a > a_y(\gamma)$:

$$\mathbb{P}(\sup_{t_0+\gamma \le t \le T_y+\gamma} \|\alpha(t) - K(t)\| < 2\theta_y\ ) > \delta'. \tag{38}$$

Choose $\gamma^y > 0$ such that,

$$\inf_N \mathbb{P}(\|X_a^N(t_0 + \gamma^y) - y\| < 2\theta^y \mid \|X_a^N(t_0) - y\| < \theta^y) > \frac{1}{2}. \tag{39}$$

Note that, (39) goes to 1, uniformly in a, when γ^y goes to zero. Let A be the following event,

$$A = \{X_a^N(t_0 + \gamma^y) \in B_{2\theta^y}(y),\ \sup_{0 \le t \le T_y} \|\alpha(t + t_0 + \gamma^y) - K^y(t)\| < 2\theta_y\}.$$

By Equations (38) and (39) for every $a > a^y(\gamma^y)$,

$$\mathbb{P}(A) > 0. \tag{40}$$

By Lemma 4 and Equation (18) there exists N_0 such that

$$\forall N > N_0,\ \mathbb{P}(X_a^N(t_0 + \gamma^y + T^y) \in \mathrm{B}_{\delta^y}(z^y(T^y)) \mid A) > 0. \tag{41}$$

Lemma 5 follows from (40) and (41). $\square$

References

[1] Babichenko, Yakov, 2010a, "Completely uncoupled dynamics and Nash equilibria," Working Paper 529, Center for the Study of Rationality, Hebrew University.

[2] Babichenko, Yakov, 2010a, "How long to Pareto efficiency?" Working Paper 562, Center for Rationality, Hebrew University.

[3] Benaïm, Michel, and Jorgen Weibull, 2003, "Deterministic approximation of stochastic evolution in games," *Econometrica, 71, 873-903.*

[4] Bjornerstedt, Jonas, and Jorgen Weibull, 1996 "Nash Equilibrium and Evolution by Imitatiom," *The Rational Foundations of Economic Behavior 155-171,* edited by Kenneth Arrow, London: Macmillan.

[5] Durrett, Richard 1996 "Stochastic Calculus: A Practical Introduction," Boca Raton: CRC Press.

[6] Foster, Dean P., and H. Peyton Young, 2003, "Learning, hypothesis testing, and Nash equilibrium," *Games and Economic Behavior, 45, 73-96.*

[7] 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.*

[8] Germano, Fabrizio, and Gabor Lugosi, 2007, "Global convergence of Foster and Youngs regret testing," *Games and Economic Behavior, 60, 135-154.*

[9] Hart, Sergiu, and Yishay Mansour, 2010, "How Long to Equilibrium? The Communication Complexity of Uncoupled Equilibrium Procedures," *Games and Economic Behavior, 69, 107-126.*

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

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

[12] Hofbauer, Josef, and William Sandholm, 2008, "Evolution in Games with Randomly Disturbed Payoffs," *Journal of Economic Theory 132, 47-69.*

[13] Hofbauer, Josef, and William Sandholm, 2008, "Survival of dominated strategies under evolutionary dynamics," *Theoretical Economics,* Forthcomming.

[14] Hofbauer, Josef, and Jeroen Swinkels, 1996, "A universal Shapley example," Discussion Paper, University of Vienna and Northwestern University.

[15] Kurtz, T. G., 1970, "Solutions of ordinary differential equations as limits of pure jump Markov processes," *Journal of Applied Probability, 7, 49-58.*

[16] Marden, Jason R., H. Peyton Young, Gurdal Arslan, and Jeff S. Shamma, 2009, "Payoff-based dynamics for multiplayer weakly acyclic games," *SIAM Journal on Control and Optimization, 48, No. 1, 373-396.*

[17] Papadimitriou, Christos H., 2005, "Computing correlated equilibria in multi-player games," In H. N. Gabow, R. Fagin, eds, *Proceedings*

of the 37th Annual ACM Symposium on Theory of Computing (STOC 2005), pp. 49-56.

[18] Sandholm, William H., 2009, "Large Population Potential Games," *Journal of Economic Theory 144, 1710-1725.*

[19] Sandholm, William H., 2010, "Pairwise comparison dynamics and evolutionary foundations of Nash equilibrium," *Games, 1, 3-17.*

[20] Sandholm, William H., 2010, "Population Games and Evolutionary Dynamics," MIT Press.

[21] Sandholm, William H., 2010, "Pairwise comparison dynamics and evolutionary foundations of Nash equilibrium," *Games, 1, 3-17.*

[22] Shah, Devavrat and Shin, Jinwoo, 2010, "Dynamics in Congestion Games," *ACM SIGMETRICS*

[23] Smith, Michael J., 1984,"The stability of a dynamic model of traffic assignment an application of a method of Lyapunov," *Transportation Science, 18, 245-252.*

[24] Young, H. Peyton, 1993, "The evolution of conventions," *Econometrica, 61, 57-84.*

[25] Young, H. Peyton, 1998, "Individual Strategy and Social Structure: An Evolutionary Theory of Institutions," Princeton NJ: Princeton University Press.

[26] Young, H. Peyton, 2009, "Learning by trial and error", *Games and Economic Behavior, 65, 626-643.*