האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

BEST-REPLY DYNAMICS IN LARGE ANONYMOUS GAMES

By

YAKOV BABICHENKO

Discussion Paper # 600 March 2012

מרכז לחקר הרציונליות

CENTER FOR THE STUDY OF RATIONALITY

Feldman Building, Givat-Ram, 91904 Jerusalem, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

Best-Reply Dynamics in Large Anonymous Games

Yakov Babichenko*†

March 18, 2012

Abstract

We consider small-influence anonymous games with a large number of players n where every player has two actions. For this class of games we present a best-reply dynamic with the following two properties. First, the dynamic reaches Nash approximate equilibria fast (in at most $cn\ logn$ steps for some constant $c > 0$). Second, Nash approximate equilibria are played by the dynamic with a limit frequency of at least $1 - e^{-c'n}$ for some constant $c' > 0$.

*Center for the Study of Rationality, Institute of Mathematics, The Hebrew University of Jerusalem. e-mail:yak@math.huji.ac.il.

†This paper is part of the Ph.D. research of the author at the Hebrew University of Jerusalem. The author wishes to thank his supervisor, Sergiu Hart, for his support and guidance, and Itai Arieli for useful discussions and comments. The author would like to acknowledge financial support by ERC 030-7950-411 and ISF 039-7679-411.

1 Introduction

The issue of convergence of dynamics to Nash equilibria has been widely studied in game theory. Two questions naturally arise in regard to this issue: Are there natural adaptive dynamics that lead to Nash equilibria in general classes of games? If so, how fast could these dynamics reach an equilibrium?

Regarding the first question, there are several impossibility results for convergence to equilibria [12] and [13]. There are also several possibility results, i.e., dynamics that converge to Nash equilibria in all n-player games, e.g., the regret-testing dynamic [7] and [9] or trial-and-error learning [18]. All these dynamics are based on the exhaustive search principle where players go over the action profile space using some stochastic (or deterministic) rule. Once the players reach an equilibrium action profile, they play it for a long time. An advantage of these exhaustive search dynamics is the fact that they converge to equilibria in *every* game. A disadvantage of exhaustive search dynamics lies in the answer to the second question, they take a long time. Specifically, for games with a large number of players n, the size of the action profile space is exponential in n (we write this as $\exp(n)$); therefore the expected time for reaching an equilibrium is $\exp(n)$. The fact that the established dynamics converge to equilibria in $\exp(n)$ periods of time is not coincidental. Hart and Mansour [11] prove that for every dynamic satisfying the natural requirement of uncoupledness there exists a game where it will take $\exp(n)$ periods in average to reach a Nash equilibrium. This impossibility result leads us to the conclusion that in order to find a dynamic that converges fast to Nash equilibria in games with a large number of players we must consider some subclass of games. Among the games whose rate of convergence has been studied to some extent are potential games [2], population acyclic games in evolutionary game theory [1], and some specific games, see for example [5] and [6].

In this paper we consider an important class of *anonymous games* introduced by Schmeidler [17]. In this class of games the payoff of each player depends only on his own action and on the proportions of the other players playing the various actions, rather than the exact action profile of the other players. Anonymous games have received much attention in various fields. In economics, for example, the price depends on the total amount of demand and production in classical markets. In computer science, congestion games are anonymous games because the cost of an edge depends on the total number of users there. In biology, single population games are anonymous games.

Assume that a large number of players n should decide whether to use technology A or technology B. The payoff function of every player depends on the proportion of A and B users, and is a λ-Lipschitz function of the proportion. Note that unlike population games, here the payoff functions of two different players could be different. At every step a single player is chosen uniformly at random[1]; then the chosen player updates his action according to his best reply to the *observable proportion.* We assume that the observable proportion is not the exact proportion but a rounding function of it (e.g., the integer percentage of players using the technology); this assumption could be explained, for example, by the difficulties (or costs) of computing the exact proportion. Our main results claim that in $O(n \log n)$ steps in expectation the played action profile will be a pure Nash approximate equilibrium; in addition, most of the time the played action profiles

[1]For example, if the players play in continuous time and the updating opportunities for players are governed by i.i.d. Poisson arrival processes then it leads exactly to the case where the chosen player will be picked uniformly.

will be pure Nash approximate equilibria.

2 Related Literature

A key negative result on convergence to Nash equilibria in large games was established by Hart and Mansour [11]. They considered *uncoupled* dynamics, i.e., dynamics where the action of every player i at every step depends on the payoff function of player i and the past realized action profiles, but does *not* depend on the payoff functions of players $j \neq i$. Uncoupledness is a very natural assumption that holds for most known dynamics, such as fictitious play, best-reply, and replicator dynamic. Using communication complexity tools Hart and Mansour show that for every uncoupled dynamic there exists an n-player game that has pure Nash equilibrium (PNE) where the expected time of reaching PNE is exponential in n. The same considerations hold also for the case of PNE_ε: for every uncoupled dynamic there exists an n-player game that has PNE_ε where the expected time to reach PNE_ε is exponential in[2] n. Therefore in large games a reasonable rate of convergence can be obtained only by subclasses of games.

Arieli and Young [1] focus on rapid convergence to equilibria in weakly acyclic population games. As in the present work, they assume that updating at each stage is carried out by a single player. In their dynamic the updating to a better action occurs not with probability 1, but with some positive probability that depends on the profit from deviation and on some additional random-walk variable. They prove that the dynamic converges to PNE in $O(n)$ steps in expectation in every weakly acyclic population game. Note that

[2]This result is a straightforward corollary from Theorem 3 in [11], and in particular from the fact that in every game $\{u_{l,i}\}$ the set of pure Nash equilibria coincides with the set of pure Nash ε-equilibria for $\varepsilon < 1/2$.

whereas in population games all the players have the same payoff function, our paper focuses on the general case where players may have different payoff functions.

Anonymous games have been the object of study for over three decades. Schmeidler [17] first considered non-atomic games with a continuum of players. Rashid [16] studied the discretization of Schmeidler's result for a finite number of players. Azriely and Shmaya [3] proved that every *small-influence* game where a deviation of player i has a small influence on the payoff of player j possesses a PNE_ε; in particular, λ-Lipschitz anonymous games belong to this class.

Daskalakis and Papadimitriou [4] prove that there exists an efficient algorithm for computing PNE_ε in anonymous games using the fact that PNE_ε is obtained from fixed points of the function $\phi : \Delta \to \Delta$, where ϕ maps every proportion δ onto the resulting proportion after all players update their function according to their best reply to δ. In the present paper the same function ϕ will play a central role. Application of the algorithm of Daskalakis and Papadimitriou to our case of two-action games generates an uncoupled communication protocol that requires $1/\varepsilon$ steps of communication in order to reach PNE_ε. Although the communication procedure reaches PNE_ε very fast, it is nevertheless far from reflecting a natural adaptive behavior. In other words, the algorithm of Daskalakis and Papadimitriou is computationally very efficient, but it does not solve the problem from a dynamical perspective since it is an "unnatural" exhaustive search. This paper uses some of the ideas from [4] to solve the dynamical case of the problem.

Fridman Halpern and Kash [8] consider a simultaneous best-reply dynamic (i.e., the best-reply updating is carried out simultaneously for all players) in anonymous games. They consider environments where players may not know the proportion played by other players, and may not know their own payoff function. They prove that this uncertainty might be overcome by stage learning; i.e., in every block, most of the time, a player plays some pure action, but sometime he or she tries different actions. Their result proves that in games where a simultaneous best-reply dynamic converges to PNE, so does the best-reply that uses stage learning. In addition, they prove that the convergence to PNE of the best-reply dynamic occurs at the same rate in both cases of fully informed players and in the case of players with bounded rationality. One problem of this approach is that it works only for games where a simultaneous best-reply dynamic converges to PNE, which does not include many interesting anonymous games, e.g., the minority game[3]. Another problem is that the rate of convergence of the simultaneous best-reply could be quite slow.

3 Formal Treatment

We consider n-player binary games. Let $A^i = \{0, 1\}$ be the action set of every player and $A = \{0, 1\}^n$ the action profile set. For $a = \{a^i\}_{i=1}^n \in A$ player i will be called a *b-player* if $a^i = b$, for $b = 0, 1$. For $a \in \{0, 1\}^m$ we denote by

$$p(a) = \frac{|\{i : a^i = 1\}|}{m}$$

the proportion of 1-players.

For $a \in A$ we denote by $a_{j \leftrightarrow k}$ the action profile where we exchange the j'-th coordinate with the k'-th coordinate. Let $u^i : A \to \mathbb{R}$ be the payoff

[3]For example, one may take the minority game where the payoff of a player that plays action $i = 0, 1$ is the proportion of players that plays the action $j = 1 - i$. In this game, if the players play according to the simultaneous best reply, then the played action profiles will be alternately $(0, 0, ..., 0)$ and $(1, 1, ..., 1)$, which are not PNE.

function of player i. A game will be called *anonymous* if $u^i(a) = u^i(a_{j\leftrightarrow k})$ for every player i and for every $j, k \neq i$, which means that the payoff of player i does not distinguish between any other players j and k. It is easy to show that in anonymous games the payoff of every player i depends only on i's action the proportion of play of the other players. Therefore, with a slight abuse of notation, we will write the payoff function of player i as $u^i(a^i, p(a^{-i}))$.

An action profile $a \in A$ is a pure Nash δ-equilibrium if for every player i, $u^i(a^i, p(a^{-i})) \geq u^i(1 - a^i, p(a^{-i})) - \delta$; i.e., no player can gain more than δ by deviating. Denote by PNE_δ the set of pure Nash δ-equilibria of the game.

We assume that the payoff functions u^i for all players i are λ-Lipschitz in the proportion; i.e., $|u^i(a^i, x) - u^i(a^i, y)| \leq \lambda|x - y|$, where $x, y \in \{0, 1/(n-1), 2/(n-1), ..., 1\}$.

Denote by $\overline{u}^i(0, \cdot), \overline{u}^i(1, \cdot) : [0, 1] \to \mathbb{R}$ the linear interpolations of $u^i(0, \cdot)$ and $u^i(1, \cdot)$ to the whole segment $[0, 1]$; which are also λ-Lipschitz.

Fix $\varepsilon > 0$. We will assume that player i does not observe the exact proportion $p(a^{-i})$ but rather an ε-floor rounding of it; i.e., $\lfloor p(a) \rfloor_\varepsilon$ where $\lfloor x \rfloor_\varepsilon := \max\{k\varepsilon : k \in \mathbb{N} \text{ and } k\varepsilon \leq x\}$. We emphasize that the ε-rounding is applied to $p(a)$, not to $p(a^{-i})$; thus the observed proportion is the same for all players. For a discussion of the case where players use different rounding functions see Remark 4 in Section 6. The convergence of the best-reply dynamic in case where players observe the exact proportion remain an open question, see Remark 5 in Section 6.

Under the assumption that player i observes the rounded proportion, we define $v^i_\varepsilon(a^i, x) \equiv v^i(a^i, x) = \overline{u}^i(a^i, \lfloor x \rfloor_\varepsilon)$ for $x \in [0, 1]$. Henceforth we will refer to v^i as the payoff function of player i, but we will have in mind that his actual payoff u^i may be different. The *best reply* of player i to the action profile $a \in A$ is $a^i = 1$ if $v^i(1, p(a)) \geq v^i(0, p(a))$, and $a^i = 0$ otherwise.[4] We will say that player i is *satisfied* if he is best replying with respect to v^i; otherwise we will say that player i is *unsatisfied*. We denote the set of unsatisfied players by $UNS(a)$.

The dynamic starts with an arbitrary pure action profile $a(0) \in A$.

At each step $t+1 \geq 1$ a single player i is chosen at random uniformly; this player i updates his action $a^i(t+1)$ to be his best reply (according to v^i_ε) to $a(t)$, whereas the other players play the same action ($a^{-i}(t+1) = a^{-i}(t)$). This dynamic will be denoted $BR^*(\varepsilon)$ (the dependence on ε appears in the best-replying rule that relies on the ε-rounding). We consider also another dynamic, denoted $BR(\varepsilon)$, where at time $t+1$ the updating player is chosen uniformly at random only among the unsatisfied players; i.e., $UNS(a(t))$. This dynamic will be called $BR(\varepsilon)$, If the chosen player is satisfied then there will be no change in the played action profile (i.e., $a(t+1) = a(t)$); therefore the only difference between the dynamics $BR^*(\varepsilon)$ and $BR(\varepsilon)$ is the speed of change. In the paper we will analyze the dynamic $BR(\varepsilon)$; the discussion on $BR^*(\varepsilon)$ is relegated to Remark 1 in Section 6.

The dynamic $BR(\varepsilon)$ induces a Markov chain over the state space A; with an abuse of notation we will refer to the Markov chain also as $BR(\varepsilon)$.

For a Markov chain M over a finite state space S with realization denoted by $(a(0), a(1), ...)$, we will denote by $\mathbb{P}_M$ the probability function of M. Thus $\mathbb{P}_M(a \to C)$ denotes the transition probability (in a single step) from $a \in S$ to the set of states $C \subset S$. For $B \subset S$ we will say that $\mathbb{P}_M(B \to C) \geq d$ if $\mathbb{P}_M(a \to C) \geq d$ for every $a \in B$. Similarly, $\mathbb{P}_M(B \to C) \leq d$ if $\mathbb{P}_M(a \to C) \leq d$ for every $a \in B$. Given an initial state $a(0) \in S$ and $C \subset S$, we define the stopping time $\tau = \tau(a(0), C) = \inf\{t : a(t) \in C\}$, and we denote

[4] Without loss of generality we assume that ties are broken in favor of the action 1.

by $T_M(a(0) \rightsquigarrow C) = E_M(\tau)$ the expected number of steps for reaching C from the initial state $a(0)$. Given $B \subset S$ we denote by

$$T_M(B \rightsquigarrow C) = \sup_{a(0) \in B} T_M(a(0) \rightsquigarrow C)$$

the maximal expected time for reaching C from B. In the case where $B = S$ we will write for short by $T_M(C) := T_M(S \rightsquigarrow C)$.

The first theorem asserts that for every anonymous game the dynamic $BR(\varepsilon)$ reaches a pure Nash $(2\lambda\varepsilon)$-equilibrium in $O(n \log n)$ steps in expectation.

Theorem 1. For every $\varepsilon > 0$ there exists $n_0(\varepsilon)$ such that for any λ-Lipschitz anonymous n-person game with $\lambda > 0$ and $n \geq n_0(\varepsilon)$, the expected number of steps that the $BR(\varepsilon)$ dynamic reaches the set of pure Nash $(2\lambda\varepsilon)$-equilibria is at most $9n \log n$, i.e.,[5] $T_{BR(\varepsilon)}(PNE_{2\lambda\varepsilon}) \leq 9n \log n$.

The second theorem asserts that for every anonymous game, these $(2\lambda\varepsilon)$-equilibria are played with a limit frequency that is exponentially close to 1. Since $BR(\varepsilon)$ is a Markov chain we can state this result using invariant distributions instead of limit frequency.

Theorem 2. For every $\varepsilon > 0$ there exists $n_0(\varepsilon)$ and a constant $c = c(\varepsilon) > 0$ (independent of n) such that for any λ-Lipschitz anonymous n-person game with $\lambda > 0$ and $n \geq n_0(\varepsilon)$, the limit frequency of steps where the played action profile of the $BR(\varepsilon)$ dynamic is a pure Nash $(2\lambda\varepsilon)$-equilibria is at least $1 - e^{-cn}$; i.e.,[6] $\mu(PNE_{2\lambda\varepsilon}) \geq 1 - e^{-cn}$ for every invariant distribution μ of the Markov chain $BR(\varepsilon)$.

[5] The proof of the theorem yields, for instance, that $n_0 = 210/\varepsilon^3$ is sufficient; one may decrease the convergence rate to $cn \log n$ where $2 < c < 9$ by taking higher n_0.

[6] The proof of the theorem yields, for instance, that $n_0 = 210/\varepsilon^3$ and $c = \varepsilon^2/8$ are sufficient.

4 Informal Sketch of the Proof

In this section we present the ideas of the proof of the main theorems, in parallel with the presentation of several notations and definitions that will be useful in the actual proof.

The proof is divided into two main steps. In the first step (Proposition 1) we prove that the proportion $p(a(t))$ reaches a *stable value* (see Definition 1) fast (in $O(n)$ steps). In the second part (Proposition 3) we prove that once the proportion has reached a stable value, the dynamic $BR(\varepsilon)$ reaches an pure Nash approximate equilibrium fast (in $O(n \log n)$ steps).

First part of the proof

A useful function that will help us to analyze the dynamic, and is similar to the function defined in [4], is $\phi : [0,1] \to [0,1]$, which is defined by

$$\phi(x) := \frac{|\{i : v^i(1,x) \geq v^i(0,x)\}|}{n}.$$

In other words, $\phi(x)$ is the resulting proportion of 1-players in the case where all the players observe the rounded proportion x and then all the players simultaneously update their actions according to their best reply.

Denote by $UNS_b(a)$, for $b = 0,1$, the set of unsatisfied b-players, and by $SAT_b(a)$ the set of satisfied b-players. Clearly $|UNS_1(a)|+|SAT_1(a)| = np(a)$ and $|UNS_0(a)| + |SAT_1(a)| = n\phi(p(a))$; by subtraction we get

$$|UNS_0(a)| - |UNS_1(a)| = n(\phi(p(a)) - p(a)). \tag{1}$$

Let $p(t)$ be the proportion of 1-players at time t. At step $t+1$ a single player among the unsatisfied players is chosen. Therefore, at step $t+1$ the proportion will be $p(t+1) = p(t) + 1/n$ if a player from $UNS_0(a(t))$ is chosen, and $p(t+1) = p(t) - 1/n$ if a player from $UNS_1(a(t))$ is chosen. From (1) we deduce that if $\phi(p(t)) > p(t)$, then the probability of moving to the right (i.e., $p(t+1) = p(t) + 1/n$) is above $1/2$, and vice versa, for a proportion that satisfies $\phi(p(t)) < p(t)$ the probability of moving to the left is above $1/2$. The basic idea is to apply random-walk arguments to prove that such a random walk reaches fast a *down-crossing fixed point* of ϕ; i.e., a q such that $\phi(q) = q$, $\phi(q-\delta) > q - \delta$ and $\phi(q+\delta) < q + \delta$ for δ small enough. The problem is that the function ϕ does not necessarily satisfy good properties (such as Lipschitz) and may not have a fixed point at all. This is where rounding plays a role. The functions v^i are piecewise constant; therefore ϕ is also piecewise constant. So the idea is to analyze the behavior of the proportion over segments where ϕ is constant (Lemmas 2, 3, 4, and 5) and then to deduce rapid convergence to a stable value by considering the Markov chain M_2 (see the formal definition in Section 5.4) that represents the movement from one constant segment to another.

For simplicity of notation we assume that $\varepsilon = m/n$ for some integer $m \in \mathbb{N}$. The functions $v^i(1,x)$ and $v^i(0,x)$ are constant in x over the segments $[k\varepsilon, (k+1)\varepsilon)$. The relevant domain of v^i and ϕ in the k'-th segment is the set of points $\{km/n, (km+1)/n, ..., ((k+1)m-1)/n\}$, so we denote by $\sigma_k := [k\varepsilon, (k+1)\varepsilon - 1/n]$ the relevant segment where the functions v^i and ϕ are constant. The center of the segment will be denoted by $c_k = (k+1/2)\varepsilon - 1/(2n)$.

In the case where $\phi(c_k) \geq c_k$ we will say that ϕ is *high* in the segment σ_k, while in the case where $\phi(c_k) \leq c_k$ we will say that ϕ is *low*. If we follow the intuition presented two paragraphs above, then for a segment where ϕ is high the proportion should move from left to right; indeed, Lemma 2 shows that the probability of crossing such a segment from left to right is constant whereas Lemma 4 shows that the probability of crossing such a segment from right to left is exponentially small. For a segment where ϕ is low the proportion should move from right to left; indeed, Lemmas 3 and 5 show that happening. Therefore, we say that the *correct direction on a segment* is right if $\phi(c_k) \geq c_k$ and left if $\phi(c_k) \leq c_k$.

Definition 1. A proportion p of the form $p = k\varepsilon$ will be called *stable* if $\phi(c_{k-1}) \geq c_{k-1}$ and $\phi(c_k) \leq c_k$.

Lemmas 4 and 5 show that if the dynamic is in a stable proportion then the probability of crossing a segment in both directions left and right is exponentially small; therefore the proportion will stay ε-close to the stable value for a long time.

Note that if the random walk of the proportion follows the correct direction in every segment then it will reach a stable value in $O(n)$ steps. In Proposition 1 we show that the $BR(\varepsilon)$ dynamic also reaches a stable value in $O(n)$ steps, using the fact that movement in the correct direction occurs with a constant probability and movement in the incorrect direction occurs with an exponentially small probability.

Second part of the proof

The relation between fixed points of a continuous version of ϕ and pure Nash approximate equilibria of an anonymous game was established in the literature; see, e.g., [4] or [3]. Every fixed point $\phi(p) = p$ induces an pure Nash approximate equilibrium where all the players are best replying according to p. The inverse is false: *not* every action profile a such that $\phi(p(a)) = p(a)$ is an pure Nash approximate equilibrium: an action profile a such that $UNS_0(a) = UNS_1(a) \neq \emptyset$ satisfies $\phi(p(a)) = p(a)$ but there could be players from $UNS(a)$ that are not ε-best replying. Therefore, reaching a fixed point of ϕ is not sufficient to reach PNE_ε; we need in addition that the unsatisfied players who gain more than ε from switching to their best reply indeed do so.

For a stable proportion $p = k\varepsilon$ we denote by I_k the set of players whose best reply is identical in the segments σ_{k-1} and σ_k. The following trivial lemma provides a sufficient condition for an action profile to constitute a pure Nash approximate equilibrium.

Lemma 1. Let $k\varepsilon$ be a stable proportion; then every action profile $a \in A$ such that $p(a) \in \sigma_{k-1} \cup \sigma_k$ and $UNS(a) \cap I_k = \emptyset$ is a pure $(2\lambda\varepsilon)$-equilibrium.

Proof. A player $i \notin I_k$ has an indifference point in a 2ε-environment of $p(a)$ (because his best reply is different in the segments σ_{k-1} and σ_k); therefore by playing his best reply the player cannot improve his payoff by more than $2\lambda\varepsilon$.

A player $i \in I_k$ is best replying according to the ε-rounding of the proportion, therefore he cannot improve his payoff by more than $\lambda(\varepsilon + 1/(n-1)) \leq 2\lambda\varepsilon$. $\square$

By Lemma 1, in order to reach a pure Nash approximate equilibrium it is sufficient to reach a situation where the proportion is in $\sigma_{k-1} \cup \sigma_k$ and all the players in I_k are best replying.

Since the proportion is stable, the probability of leaving the segments $\sigma_{k-1} \cup \sigma_k$ is exponentially small, and once a player from I_k has been chosen he will not update his action as long as the proportion stays in $\sigma_{k-1} \cup \sigma_k$. Therefore in order to reach an pure Nash approximate equilibrium it is sufficient to choose all the players from I_k before the proportion leaves $\sigma_{k-1} \cup \sigma_k$. Using arguments related to the Coupon Collector's problem, the expected time of choosing all the players from I_k is $O(n \log n)$, whereas the expected time of leaving the segments $\sigma_{k-1} \cup \sigma_k$ is exponentially large in n. In Proposition 3 we use these ideas in order to prove that the $BR(\varepsilon)$ dynamic reaches $PNE_{2\lambda\varepsilon}$ in $O(n^2 \log n)$ steps. More accurate considerations appear in the proof of Theorem 1 where we prove that the $BR(\varepsilon)$ dynamic reaches $PNE_{2\lambda\varepsilon}$ in $O(n \log n)$ steps.

Proof of Theorem 2

We consider the expected time of reaching Nash approximate equilibria versus the expected time of leaving it. By Theorem 1 the expected time of reaching it is $O(n \log n)$ whereas by Lemmas 4 and 5 (see inequality (5)) the expected time of leaving the approximate equilibria is exponentially small. By known results from Markov chain theory, this implies that $BR(\varepsilon)$ is most of the time in approximate equilibria.

5 Formal Proof of the Main Theorems

We will use the following standard notations. We will write $g(n) \leq O(f(n))$ if there exists n_0 and a constant $c > 0$ such that $g(n) \leq cf(n)$ for every $n > n_0$, and we will write $g(n) \geq \Omega(f(n))$ if there exists n_0 and constant $c > 0$ such that $g(n) \geq cf(n)$ for every $n > n_0$. In the statement of the propositions and lemmas we will specify those c and n_0, but in the proofs we will skip those simple computations. Note that we refer to ε as a fixed number; therefore c could depend on ε. We assume for simplicity that $\varepsilon \leq 1$.

5.1 The Markov chain M_1 induced by the stopping time t_m

We denote by $l_k = k\varepsilon$ the left corner of σ_k and by $r_k := (k+1)\varepsilon - 1/n$ the right corner of σ_k. The corresponding subsets of A will be denoted by $L_k = \{a \in A : p(a) = l_k\}$ and by $R_k = \{a \in A : p(a) = r_k\}$. $AS := \{a : UNS(a) = \emptyset\}$ denotes the set of action profiles where all the players are satisfied.

We would like to analyze the transition probabilities from proportion $p = r_k$ to proportions $p = l_k$ and $p = l_{k+1}$, and from l_k to r_{k-1} and r_k. Therefore we define the following sequence of stopping times:

$$\begin{aligned} t_1 &= \min\{t : a(t) \in L \cup R \cup AS\}. \\ \text{If } a(t_m) &\in AS \text{ then } t_{m+1} = t_m + 1. \\ \text{If } a(t_m) &\in R_k \text{ then } t_{m+1} = \inf\{t > t_m : a(t) = L_k \cup L_{k+1} \cup AS\}. \\ \text{If } a(t_m) &\in L_k \text{ then } t_{m+1} = \inf\{t > t_m : a(t) = R_{k-1} \cup R_k \cup AS\}. \end{aligned}$$

Lemmas 2 and 3 claim that with constant probability (independent on n) the proportion will cross a segment in the correct direction.

Lemma 2. If $\phi(c_k) \geq c_k$ then[7] $\mathbb{P}_{M_1}(L_k \to R_k \cup AS) \geq \Omega(1)$.

Lemma 3. If $\phi(c_k) \leq c_k$ then[8] $\mathbb{P}_{M_1}(R_k \to L_k \cup AS) \geq \Omega(1)$.

Lemmas 4 and 5 claim that with an exponentially (in n) small probability the proportion will cross a segment in the incorrect direction.

Lemma 4. There exists $n_0(\varepsilon)$ such that for every $n \geq n_0$ if $\phi(c_k) \geq c_k$ then[9] $\mathbb{P}_{M_1}(R_k \to L_k) \leq e^{-\Omega(n)}$.

Lemma 5. There exists $n_0(\varepsilon)$ such that for every $n \geq n_0$ if $\phi(c_k) \leq c_k$ then[10] $\mathbb{P}_{M_1}(L_k \to R_k) < e^{-\Omega(n)}$.

The proofs of Lemmas 3 and 4 are similar to the proofs of Lemmas 2 and 5; therefore we will present only the proofs of Lemmas 2 and 5 in Section 5.3.

[7] The proof of the lemma yields, for instance, that $\Omega(1) = \varepsilon/2$ is sufficient.

[8] The proof of the lemma yields, for instance, that $\Omega(1) = \varepsilon/2$ is sufficient.

[9] The proof of the lemma yields, for instance, that $n_0 = 2/\varepsilon$ and $\Omega(n) = \varepsilon^2 n/12$ are sufficient.

[10] The proof of the lemma yields, for instance, that $n_0 = 2/\varepsilon$ and $\Omega(n) = \varepsilon^2 n/12$ are sufficient.

5.2 Sampling without replacement and the probability space Γ

One of the central ideas in the proof is the following: once a player has updated his action according to his best reply he will not be chosen again as long as the proportion $p(a(t))$ stays in the same segment. Therefore the behavior of $BR(\varepsilon)$ in a segment depends on the order in which the unsatisfied players are chosen, when the choice is done *without* replacement. The probability space that we will consider to represent random choice without replacement is $\Gamma(s_0, s_1)$, where $s_0, s_1 \in \mathbb{N}$ and every element in $\Gamma(s_0, s_1)$ is chosen with equal probability. $\Gamma(s_0, s_1)$ is the set of functions: $\gamma : \{0, 1, 2, ..., s_0+s_1\} \to \mathbb{Z}$ such that $\gamma(0) = 0$, $\gamma(s_0 + s_1) = s_0 - s_1$ and $\gamma(t + 1) = \gamma(t) \pm 1$ for every $t = 0, 1, ..., s_0 + s_1 - 1$. Every $\gamma \in \Gamma(s_0, s_1)$ can be viewed as a path in $\mathbb{Z}^2$ from the point $(0, 0)$ to the point $(s_0 + s_1, s_0 - s_1)$.

The size of $\Gamma(s_0, s_1)$ is given by

$$|\Gamma(s_0, s_1)| = \binom{s_0 + s_1}{s_0}$$

because every choice of a function γ is equivalent to the choice of the s_0 places where the function γ increases.

For $x, y \in \mathbb{Z}$ and $\gamma \in \Gamma$ we will say that γ *reaches* x if there exists t such that $\gamma(t) = x$. We denote $x \prec_\gamma y$ if $\inf\{t|\gamma(t) = x\} < \inf\{t|\gamma(t) = y\}$, which means that the path γ reaches x before it reaches y (and indeed x is reached at some finite time). We denote $x \preccurlyeq_\gamma y$ if $\inf\{t|\gamma(t) = x\} \le \inf\{t|\gamma(t) = y\}$, which means that the path γ reaches x before it reaches y, or it never reaches any of them.

The following simple lemma is proved by implementation of Andre's reflection method (see, e.g., [14]), similar to the classical Bertrand's Ballot problem.

Lemma 6. Let $s_0, s_1, k \geq 0$ be such that $-1 < s_0 - s_1 < k$; then

(a) $|\{\gamma : \gamma \text{ reaches } -1\}| = \binom{s_0+s_1}{s_1-1}$;

(b) $|\{\gamma : \gamma \text{ reaches } k\}| = \binom{s_0+s_1}{s_1+k}$.

Proof. (a) We define a bijection function f between the sets $\{\gamma : \gamma \text{ reaches } -1\}$ and $\Gamma(s_1 - 1, s_0 + 1)$. $f(\overline{\gamma})$ is equal to $\overline{\gamma}$ up to the first time t_0 such that $\overline{\gamma}(t_0) = -1$, and $f(\overline{\gamma}(t))$ is the reflection of $\overline{\gamma}(t)$ with respect to the line $y = -1$ for $t > t_0$. Formally, let $\overline{\gamma}$ be a path that reaches -1 and let $t_0 = \arg\min\{t : \overline{\gamma}(t) = -1\}$. We define

$$f(\overline{\gamma})(t) = \begin{cases} \overline{\gamma}(t) & t \leq t_0 \\ -2 - \overline{\gamma}(t) & t > t_0 \end{cases}.$$

It is easy to show that indeed $f(\overline{\gamma}) \in \Gamma(s_1 - 1, s_0 + 1)$ and f is a bijection because composition of two reflections is the identity function.

(b) Similar to (a) we define a bijection function f between the sets $\{\gamma : \gamma \text{ reaches } k\}$ and $\Gamma(s_1 + k, s_0 - k)$, where $f(\overline{\gamma})$ is equal to $\overline{\gamma}$ up to the first time t_0 where $\overline{\gamma}(t_0) = k$, and $f(\overline{\gamma})(t)$ is the reflection of $\overline{\gamma}(t)$ with respect to the line $y = k$ for $t > t_0$. □

5.3 Proof of Lemmas 2 and 4

Proof of Lemma 2. Let $a(t_m) \in L_k$ be the initial state, and denote $s_b = |UNS_b(a(t_m))|$ for $b = 0, 1$. $\phi(k\varepsilon) = \phi(c_k) \geq (k + 1/2)\varepsilon - 1/2n$; therefore by equality (1) we know that $s_0 - s_1 \geq n\varepsilon/2 - 1/2$.

The order in which the $s_0 + s_1$ players will be chosen is equivalent to a choice of a path $\gamma \in \Gamma(s_0, s_1)$. The relation between the two probability spaces $\mathbb{P}_{M_1}$ and $\mathbb{P}_{\Gamma(s_0,s_1)}$ is as follows:

(-) $n\varepsilon - 1 \prec_\gamma -1$ is equivalent to $a(t_{m+1}) \in R_k$.

(-) γ reaches neither -1 nor $n\varepsilon - 1$ is equivalent to $a(t_{m+1}) \in AS$, because if all the $s_0 + s_1$ unsatisfied players change their actions and the proportion stays in the segment σ_k then the dynamic reaches an action profile where all the players are satisfied.

Now we can express the probability $\mathbb{P}_{M_1}(L_k \to R_k \cup AS)$ in terms of the space Γ:

$$\mathbb{P}_{M_1}(L_k \to R_k \cup AS) = \mathbb{P}_{\Gamma(s_0,s_1)}(n\varepsilon - 1 \preccurlyeq_\gamma -1) = \pi_0.$$

Since $\mathbb{P}_\Gamma$ is the uniform distribution, π_0 can be expressed as:

$$\pi_0 = \frac{|\{\gamma \in \Gamma(s_0,s_1) : n\varepsilon - 1 \preccurlyeq_\gamma -1\}|}{|\Gamma(s_0,s_1)|}.$$

Clearly the set of paths that reach -1 not before they reach $n\varepsilon - 1$ includes in the set of paths that *never* reach -1; therefore,

$$\pi_0 \geq 1 - \frac{|\{\gamma \in \Gamma(s_0,s_1) : \gamma \text{ reaches } -1\}|}{|\Gamma(s_0,s_1)|} = \pi_1.$$

By Lemma 6 (a) we get

$$\pi_1 = 1 - \frac{\binom{s_0+s_1}{s_1-1}}{\binom{s_0+s_1}{s_1}} = \frac{s_0 - s_1 + 1}{s_0 + 1} \geq \frac{n\frac{\varepsilon}{2} + \frac{1}{2}}{n} \geq \frac{\varepsilon}{2}.$$

□

Proof of Lemma 5. Let $a \in L_k$ by the initial state, and denote $s_b = |UNS_b(a)|$ for $b = 0, 1$. Since $\phi(k\varepsilon) = \phi(c_k) \leq (k + 1/2)\varepsilon - 1/2n$ we know by equality (1) that $s_0 - s_1 \leq n\varepsilon/2 - 1/2$. So

$$\begin{aligned}\mathbb{P}_{M_1}(R_k \to L_k) = \mathbb{P}_{\Gamma(s_0,s_1)}(n\varepsilon - 1 \prec_\gamma -1) \leq \\ \leq \frac{|\{\gamma \in \Gamma(s_0,s_1) : \gamma \text{ reaches } n\varepsilon - 1\}|}{|\Gamma(s_0,s_1)|} = \pi_3.\end{aligned}$$

By Lemma 6 (b) we have

$$\pi_3 = \frac{\binom{s_0+s_1}{n\varepsilon+s_1-1}}{\binom{s_0+s_1}{s_0}} = \frac{s_0!s_1!}{(s_1 + n\varepsilon - 1)!(s_0 - n\varepsilon + 1)!}.$$

The only remaining thing is to bound the expression π_3:

$$\begin{aligned}
\pi_3 &= \frac{s_0}{s_1+n\varepsilon-1}\cdot\frac{s_0-1}{s_1+n\varepsilon-2}\cdot\ldots\cdot\frac{s_0-(n\varepsilon-2)}{s_1+n\varepsilon-(n\varepsilon-2)} < \\
&< \left(\frac{s_0}{s_1+n\varepsilon-1}\right)^{n\varepsilon-1} \leq \left(\frac{s_1+n\frac{\varepsilon}{2}-\frac{1}{2}}{s_1+n\varepsilon-1}\right)^{n\varepsilon-1} < \left(\frac{n+n\frac{2\varepsilon}{3}}{n+n\varepsilon}\right)^{n\varepsilon-1} = \\
&= \left(1-\frac{\varepsilon}{3(1+\varepsilon)}\right)^{n\varepsilon-1} < e^{-\frac{\varepsilon}{3(1+\varepsilon)}(n\varepsilon-1)} = e^{-\Omega(n)}.
\end{aligned}$$

□

5.4 The Markov chain M_2 induced by the stopping time t_{2m}

Let us consider the subsequence $\{a(t_{2m})\}_{m=1}^{\infty}$, which is also a Markov chain, denoted by M_2. Every single step of $BR(\varepsilon)$ (i.e., $\{a(t)\}_{t=1}^{\infty}$) will be called a *step*, and every single step of M_2 (i.e., $\{a(t_{2m})\}_{m=1}^{\infty}$) will be called a *period*.

Remark 1. Note that $t_{m+1}-t_m \leq n$, because a player cannot be chosen twice in the same segment; therefore, after n steps, either $BR(\varepsilon)$ switches to a new segment or $BR(\varepsilon)$ reaches AS. Therefore $t_{2m+2}-t_{2m} \leq 2n$, which means that every period is at most $2n$ steps.

Since the sequence $a(t_m)$ switches every time between $R\cup AS$ and $L\cup AS$, it follows that $a(t_{2m}) \in R\cup AS$ for every m or $a(t_{2m}) \in L\cup AS$ for every m. Without loss of generality we assume that $a(t_{2m}) \in L\cup AS$.

To deal with the extreme segments σ_0 and $\sigma_{\lfloor 1/\varepsilon\rfloor}$ in the same way as we do with the other segments, we put $\phi(c_{-1}) \geq c_{-1}$ and $\phi(c_{\lfloor 1/\varepsilon\rfloor+1}) \leq c_{\lfloor 1/\varepsilon\rfloor+1}$.

Using Lemmas 2, 3, 4 and 5 we can deduce the following bounds on the transition probabilities of M_2:

If $\phi(c_{k-1}) \geq c_{k-1}$ and $\phi(c_k) \geq c_k$, then

$$\begin{aligned} \mathbb{P}_{M_2}(L_k \to L_{k-1}) &< e^{-\Omega(n)}; \\ \mathbb{P}_{M_2}(L_k \to L_{k+1} \cup AS) &\geq \Omega(1)(1 - e^{-\Omega(n)}) \geq \Omega(1) \end{aligned} \tag{2}$$

If $\phi(c_{k-1}) \geq c_{k-1}$ and $\phi(c_k) \geq c_k$, then

$$\begin{aligned} \mathbb{P}_{M_2}(L_k \to L_{k-1} \cup AS) &\geq \Omega(1); \\ \mathbb{P}_{M_2}(L_k \to L_{k+1}) &< e^{-\Omega(n)}. \end{aligned} \tag{3}$$

If $\phi(c_{k-1}) \leq c_{k-1}$ and $\phi(c_k) \geq c_k$, then

$$\begin{aligned} \mathbb{P}_{M_2}(L_k \to L_{k-1} \cup AS) &\geq \Omega(1); \\ \mathbb{P}_{M_2}(L_k \to L_{k-1} \cup AS) &\geq \Omega(1). \end{aligned} \tag{4}$$

If $\phi(c_{k-1}) \geq c_{k-1}$ and $\phi(c_k) \leq c_k$ (i.e., l_k is a stable proportion), then

$$\mathbb{P}_{M_2}(L_k \to L_{k-1}) < e^{-\Omega(n)};$$
$$\mathbb{P}_{M_2}(L_k \to L_{k+1}) < e^{-\Omega(n)};$$

therefore

$$\mathbb{P}_{M_2}(L_k \to L_k) \geq 1 - e^{-\Omega(n)}. \tag{5}$$

It is easy to show that the inequalities hold also for the extreme segments σ_0 and $\sigma_{\lfloor 1/\varepsilon \rfloor}$.

An example of a function ϕ and the derived inequalities is presented in Figure 1.

5.5 The Markov chain M_2 reaches a stable proportion fast

Let $l^* = \{l_k : \phi(c_{k-1}) \geq c_{k-1} \text{ and } \phi(c_k) \leq c_k\}$ be the set of stable proportions (see Definition 1), and $L^* = \{L_k : l_k \in l^*\}$ be the set of action profiles with stable proportion.

The following proposition claims that the expected number of periods required to reach L^* is constant (independent in n):

Proposition 1. There exists $n_0(\varepsilon)$ such that $T_{M_2}(L^* \cup AS) \leq O(1)$ for every[11] $n \geq n_0(\varepsilon)$.

The proposition implies the following corollary using Remark 1.

Corollary 1. There exists $n_0(\varepsilon)$ such that $T_{BR(\varepsilon)}(L^* \cup AS) \leq 2n \cdot O(1) = O(n)$ for $n \geq n_0(\varepsilon)$.

In the proof of Proposition 1 we use Proposition A.3 of Gorodeisky [10] which exactly suits to our situation. The proposition claims the following:

Proposition 2 (Gorodeisky [10])**.** Given a Markov chain M over S, let $S_0, S_1, ..., S_m$ be pairwise disjoint sets such that $\overset{m}{\underset{i=0}{\cup}} S_i = S$. If

$$\mathbb{P}_M\left(S_j \to \overset{j-1}{\underset{i=0}{\cup}} S_i\right) \geq f_j \text{ and } \mathbb{P}_M\left(S_j \to \overset{m}{\underset{i=j+1}{\cup}} S_i\right) \leq g_j$$

for every $j = 1, 2, ..., m$, then

$$T_M(S_0) \leq \left(\sum_{i=0}^{m} \frac{1}{f_i}\right)\left(\prod_{i=0}^{m-1}(1 + \frac{g_i}{f_i})\right).$$

Proof of Proposition 1. For every $k = 0, 1, ..., \lfloor 1/\varepsilon \rfloor$ we define the distance $d(k)$ of l_k from l^* as follows:

-If $\phi(c_{k-1}) \leq c_{k-1}$ and $\phi(c_k) \leq c_k$ then $d(k) = k - \max\{k' < k : l_{k'} \in l^*\} = d_l(k)$;[12] i.e., the distance from the closest stable proportion from the left.

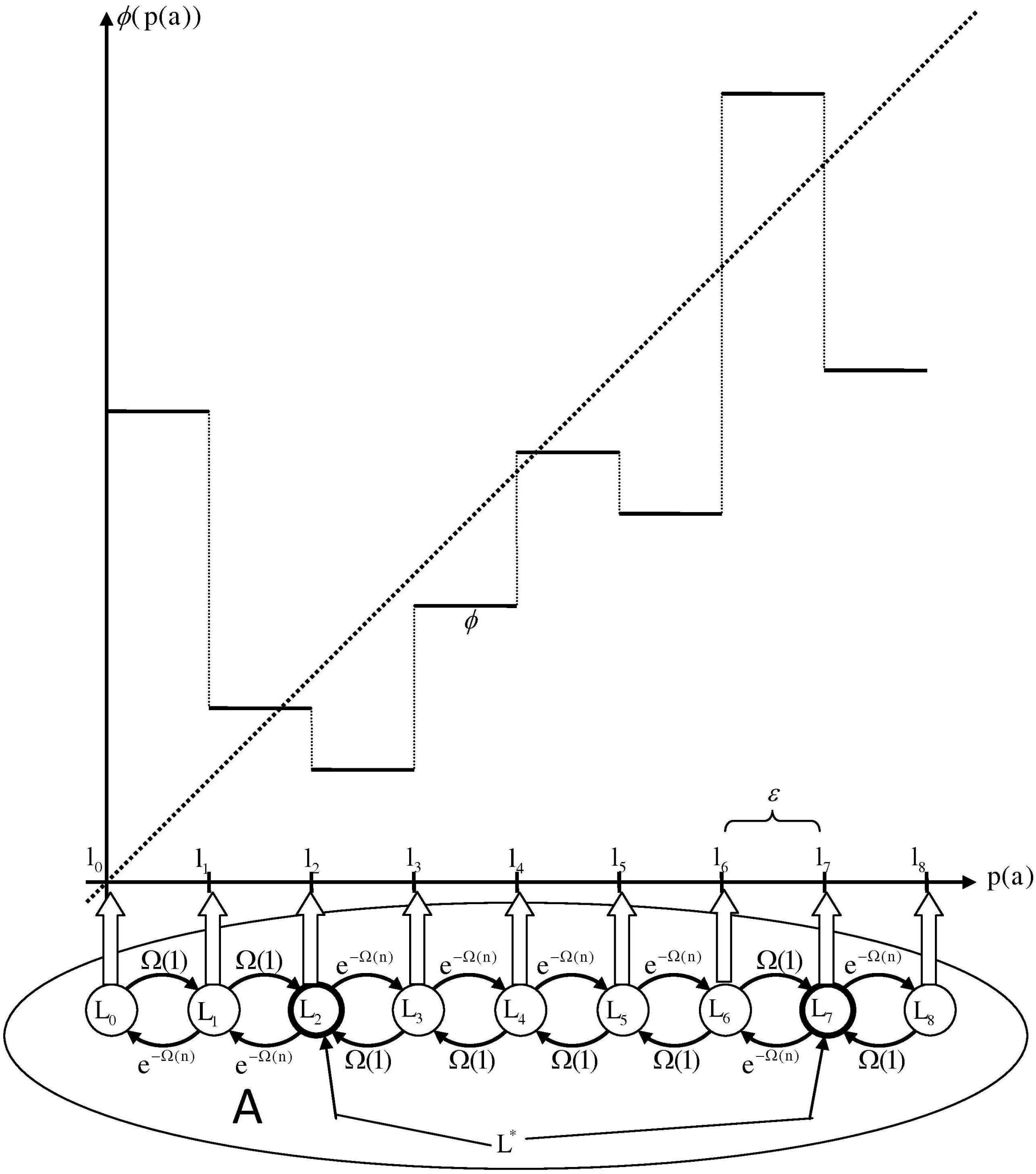
-If $\phi(c_{k-1}) \geq c_{k-1}$ and $\phi(c_k) \geq c_k$ then $d(k) = \min\{k' > k : l_{k'} \in l^*\} - k = d_r(k)$; i.e., the distance from the closest stable proportion from the right.

-If $\phi(c_{k-1}) \leq c_{k-1}$ and $\phi(c_k) \geq c_k$ then $d(k) = \max(d_l(k), d_r(k))$; i.e., the maximal distance between the left distance and the right distance.

[11] The proof of the proposition yields, for instance, that $n_0 = \frac{12}{\varepsilon^2}$ and $O(1) = \frac{24}{\varepsilon^2}$ are sufficient.

[12] Note that $-1 \in \{k' \leq k : l_{k'} \in l^*\} \neq \emptyset$; therefore the distance is well defined.

Figure 1:



φ(p(a))
φ
ε
l0
l1
l2
l3
l4
l5
l6
l7
l8
p(a)
Ω(1)
e−Ω(n)
L0
L1
L2
L3
L4
L5
L6
L7
L8
A
L*

In the example presented in Figure 1 the distances are $d(0) = 2, d(1) = 1, d(2) = 0, d(3) = 1, d(4) = 2, d(5) = 3, d(6) = 4, d(7) = 0, d(8) = 1$.

Now we define sets of states $S_j = \{L_k : d(k) = j\}$ for $j = 1, 2, ..., \lfloor 1/\varepsilon \rfloor$ and $S_0 = L^* \cup AS$. By inequalities (2), (3), and (4) we get that $\mathbb{P}_{M_2}(S_j \to S_{j-1} \cup AS) \geq \Omega(1)$, and $\mathbb{P}_{M_2}(S_j \to \bigcup_{i=j+1}^{m} S_i) \leq e^{-\Omega(n)}$; therefore by Proposition 2 we have

$$T(L^* \cup AS) \leq \left(\frac{1}{\varepsilon} \cdot \frac{1}{\Omega(1)}\right)\left(1 + \frac{e^{-\Omega(n)}}{\Omega(1)}\right)^{\frac{1}{\varepsilon}} \leq O(1).$$

□

5.6 The Markov chain M_2 reaches $PNE_{2\lambda\varepsilon}$ fast

Up to now we have proved that $BR(\varepsilon)$ reaches $L^* \cup AS$ fast; now we would like to prove that once $BR(\varepsilon)$ reaches $L^* \cup AS$ it reaches pure $(2\lambda\varepsilon)$-equilibria fast.

Let l_k be a stable proportion; we split the action profiles in L_k into $n+1$ sets. For $j = 0, 1, ..., n$ we denote $L_k(j) = \{a \in L_k : |UNS(a) \cap I_k| = j\}$; recall that I_k is the set of players with identical best replies in the segments σ_{k-1} and σ_k (see Section 4). We denote

$$L^*(j) := \bigcup_{k: l_k \in l^*} L_k(j).$$

Let us bound the transition probabilities over $L_k(j)$.

Lemma 7. Let l_k be a stable proportion; then there exists $n_0(\varepsilon)$ such that

$$\mathbb{P}_{M_2}\left(L_k(j) \to \bigcup_{i=0}^{j-1} L_k(i) \cup AS\right) \geq \frac{j}{3n}$$

for every[13] $n \geq n_0(\varepsilon)$.

[13]The proof of the Lemma yields, for instance, that $n_0 = \frac{4}{\varepsilon^2}$ is sufficient.

Proof. Let $a \in L_k(j)$ be the initial state. Denote by $m_0 = |UNS_0(a) \cap I_k|$ the number of unsatisfied 0-players with identical best replies of 1; similarly $m_1 = |UNS_1(a) \cap I_k|$. Note that $m_0 + m_1 = j$. Let B be the event that the Markov chain M_2 stays in $L_k \cup AS$. By inequality (5) we know that $\mathbb{P}_{M_2}(B) \geq 1 - e^{-\Omega(n)}$. Given the event B, no player from $I_k \backslash UNS(a)$ switches his action because the proportion stays in the segments σ_{k-1},σ_k. In addition, at least one player from UNS_0 and one player from UNS_1 have been chosen. If in at least one of these choices a player from I_k has been chosen then M_2 moves to $\overset{j-1}{\underset{i=0}{\cup}} L_k(i) \cup AS$. Therefore

$$\mathbb{P}_{M_2}(L_k(j) \to \overset{j-1}{\underset{i=0}{\cup}} L_k(i) \cup AS \mid B) \geq$$
$$\geq 1 - \left(1 - \frac{m_0}{|UNS_0(a)|}\right)\left(1 - \frac{m_1}{|UNS_1(a)|}\right) \geq$$
$$\geq 1 - \left(1 - \frac{m_0}{2|UNS_0(a)|} - \frac{m_1}{2|UNS_1(a)|}\right)^2 \geq$$
$$\geq 1 - \left(1 - \frac{j}{2n}\right)^2 = \frac{j}{n}\left(1 - \frac{j}{2n}\right) \geq \frac{j}{2n}.$$

Therefore

$$\mathbb{P}_{M_2}\left(L_k(j) \to \overset{j-1}{\underset{i=0}{\cup}} L_k(i) \cup AS\right) \geq \frac{j}{2n}(1 - e^{-\Omega(n)}) \geq \frac{j}{3n}.$$

□

Now we are able to prove the second part of the proof.

Proposition 3. There exists $n_0(\varepsilon)$ such that $T_M(L^*(0) \cup AS) \leq O(n \log n)$ for every[14] $n \geq n_0(\varepsilon)$.

Note that by Lemma 1 $L^*(0) \subset PNE_{2\lambda\varepsilon}$. Using Remark 1 an immediate corollary from Proposition 3 is

[14]The proof of the proposition yields, for instance, that $n_0 = \frac{141}{\varepsilon^3}$ and $O(n \log n) = 36 n \ln n$ are sufficient.

Corollary 2. There exists $n_0(\varepsilon)$ such that $T_{BR(\varepsilon)}(PNE_{2\lambda\varepsilon}) \leq 2nO(n \log n) = O(n^2 \log n)$ for $n \geq n_0(\varepsilon)$.

Proof of Proposition 3. The proof uses the same ideas as the proof of Proposition 1. We will again use Proposition 2 but in this case we need to define additional state sets: for $j = 0, 1, ..., n$ we define $S'_j = L^*(j)$. For $j = n+1, n+2, ..., n + \lfloor 1/\varepsilon \rfloor$ we take $S'_j = S_{j-n}$ where S_i are defined in the proof of Proposition 1. Using the notations of Proposition 2 on the state sets S'_j: for $j = 1, ..., n$ we have $f_j = j/(3n)$ (by Lemma 7) and $g_j = e^{-\Omega(n)}$. For $j = n+1, n+2, ..., n + \lfloor 1/\varepsilon \rfloor$ (similar to Proposition 1), we have $f_j = \Omega(1)$ and $g_j = e^{-\Omega(n)}$. Therefore by Proposition 2

$$T_{M_2}(L^*(0) \cup AS) \leq \left(\frac{1}{\varepsilon} \cdot \frac{1}{\Omega(1)} + \sum_{i=1}^{n} \frac{2n}{i} \right) \cdot$$
$$\cdot \left(1 + \frac{e^{-\Omega(n)}}{\Omega(1)} \right)^{\frac{1}{\varepsilon}} \prod_{i=1}^{n} \left(1 + \frac{2n \cdot e^{-\Omega(n)}}{i} \right) \leq O(n \log n).$$

□

5.7 Proof of the main theorems

Corollary 2 provides a bound of $O(n^2 \log n)$ on the rate of convergence, while Theorem 1 claims a bound of $O(n \log n)$. Indeed in the proof of Proposition 3 it was assumed that in every period of size $2j$ there is *one* left movement and *one* right movement, while in fact there are j left movements and j right movements.

Proof of Theorem 1. Let us bound $T_{BR(\varepsilon)}(L^* \rightsquigarrow L^*(0) \cup AS)$ using Corollary 2 and simple probability considerations.

We define an event B in the probability space induced by the Markov chain $BR(\varepsilon)$ with initial state $a(0) \in L_k \subset L^*$. B is the event where M_2

stays in L_k for n^3 consecutive periods,[15] and during these n^3 periods all the players from I_k are chosen. It should be emphasized that the event B requires n^3 *periods* of staying in L_k and not n^3 steps.

Given event B, we split the steps of the first n^3 periods into two categories. For $b = 0, 1$ let w_b be the set of steps where an unsatisfied b-player has been chosen. Note that $|w_0| = |w_1| \geq n^3$ because in every period the proportion leaves $k\varepsilon$ and then returns to $k\varepsilon$; therefore there is at least one choice of a 0-player and one choice of a 1-player. For $b = 0, 1$ we denote by $I_k(b)$ the set of players with identical best replies of b (regardless of the played action profile); note that $I_k = I_k(0) \cup I_k(1)$. At each step $t \in w_b$ the probability of choosing a player from $I_k(b)$ is at least $|I_k(b) \cap UNS(a(t))|/n$. Without the restriction that all the players from I_k are chosen during those n^3 periods, the expected number of steps required to choose all the players in $I_k(b)$ is $O(n \log n)$ for $b = 0, 1$ by standard arguments of the Coupon Collector's problem. Therefore, given the event B, the expected number of steps required to choose all the players in $I_k(b)$ is at most $O(n \log n)$, because event B simply bounds this number of steps by at most n^3 periods. Therefore $T_{BR(\varepsilon)}(L^* \cup AS \rightsquigarrow L * (0) \cup AS \mid B) \leq O(n \log n)$.

Let us bound $\mathbb{P}_{BR(\varepsilon)}(B)$. By the Markov inequality, the probability of not choosing all the players in $I_k(b)$ is at most $O(n \log n)/n^3$ for $b = 0, 1$. Therefore, the probability of choosing all the players in I_k is at least $(1 - O(\log n)/n^2)^2$. In addition, the probability of staying at L_k for n^3 consecutive periods is at least $(1 - e^{-\Omega(n)})^{n^3}$. Therefore

$$\mathbb{P}_{BR(\varepsilon)}(B) \geq (1 - e^{-\Omega(n)})^{n^3} \left(1 - \frac{O(\log n)}{n^2}\right)^2 \geq 1 - \frac{O(\log n)}{n^2}.$$

In the case where B does not occur we can bound $T_{BR(\varepsilon)}(L^*(0) \cup AS)$ by

[15] Any number of periods m that satisfies $n^2 \log n \leq m << exp(n)$ is good enough for the proof, and not necessarily $m = n^3$.

$O(n^2 \log n)$ using Corollary 2. Therefore we get

$$T_{BR(\varepsilon)}(L^* \cup AS \rightsquigarrow L^*(0) \cup AS) \leq O(n \log n)\mathbb{P}_{BR(\varepsilon)}(B) +$$
$$+O(n^2 \log n)(1 - \mathbb{P}_{BR(\varepsilon)}(B)) \leq O(n \log n)\left(1 - \frac{O(\log n)}{n^2}\right) +$$
$$+O(n^2 \log n)\frac{O(\log n)}{n^2} \leq O(n \log n)$$

Therefore, using Corollary 1, we get

$$T_{BR(\varepsilon)}(PNE_{2\lambda\varepsilon}) \leq T_{BR(\varepsilon)}(L^*(0) \cup AS) \leq T_{BR(\varepsilon)}(L^* \cup AS) +$$
$$+T_{BR(\varepsilon)}(L^* \cup AS \rightsquigarrow L^*(0) \cup AS) \leq O(n) + O(n \log n) \leq O(n \log n).$$

□

The proof of Theorem 2 is based only on an expected arrival time analysis. We will use the following simple lemma on Markov chains, which follows directly from Gorodeisky [10], Proposition A.5:

Lemma 8 (Gorodeisky [10])**.** Let M be a Markov chain over S, and let $S_1 \subset S_2 \subset S$ be two subsets such that $T_M(S_1) \leq g_1$ and $T_M(S_1 \rightsquigarrow S \backslash S_2) \geq g_2$. Then $\mu(S_2) \geq \frac{g_2}{g_1+g_2}$ for every invariant distribution μ of M.

Proof of Theorem 2. Let $S_1 = L^*(0) \cup AS$ and let

$$S_2 = \bigcup_{k:l_k \in l^*} \{a : p(a) \in \sigma_{k-1} \cup \sigma_k \text{ and } UNS(a) \cap I = \emptyset\} \cup AS.$$

which is an extension of $L_k(0)$ to action profiles a with proportion $p(a)$ that is not necessarily exactly $k\varepsilon$ but some proportion in the segments σ_{k-1}, σ_k. By Lemma 1, $S_2 \subset PNE_{2\lambda\varepsilon}$.

Once $BR(\varepsilon)$ reaches $L^*(0)$ we know by inequality (5) that the expected time of leaving the segments σ_{k-1}, σ_k is at least $e^{\Omega(n)}$, and as long as the proportion stays in $\sigma_{k-1} \cup \sigma_k$ only action profiles from S_2 will be played.

Therefore we get $T_{BR(\varepsilon)}(S_1 \rightsquigarrow S \backslash S_2) \geq e^{\Omega(n)}$, while $T_{BR(\varepsilon)}(S_1) \leq O(n \log n)$ (by Theorem 1). Therefore

$$\mu(PNE_{2\lambda\varepsilon}) \geq \mu(S_2) \geq \frac{e^{\Omega(n)}}{O(n \log n) + e^{\Omega(n)}} \geq 1 - e^{-\Omega(n)}.$$

□

6 Remarks

1. Best-reply dynamic where the updating player is chosen uniformly from all players. We have assumed that the updating player is chosen only from the *unsatisfied* players. The dynamic where the updating player is chosen from *all* players was denoted by BR^*. A trivial bound on the expected time for reaching PNE is $T_{BR^*}(PNE_{2\lambda\varepsilon}) \leq nO(n \log n) = O(n^2 \log n)$ because at each step the probability of choosing an unsatisfied player is at least $1/n$; therefore in expectation it will take at most n steps to choose such a player. More accurate considerations can prove $T_{BR^*}(A \rightsquigarrow PNE_{2\lambda\varepsilon}) \leq O(n^2)$; this is because by observing the second part of the proof, we can see that we do not use there the fact that the chosen player is unsatisfied, while in the first part we do use it. Therefore

$$T_{BR^*}(PNE_{2\lambda\varepsilon}) \leq T_{BR^*}(L^* \cup AS) + T_{BR^*}(L^* \cup AS \rightsquigarrow L^*(0) \cup AS) \leq$$
$$\leq O(n^2) + O(n \log n) \leq O(n^2).$$

2. Deterministic choice of updating player. Unlike in potential games or congestion games where every choice of updating players leads eventually to PNE, in anonymous games this is not the case. We show a counterexample where the updating players are chosen deterministically in a round robin order; i.e., at time t the updating player is $t(\text{mod } n)$. Similar counterexamples can be constructed for other deterministic roles.

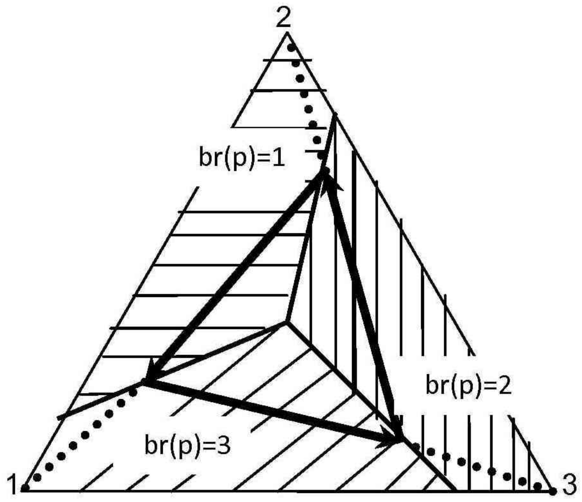
Let the number of players be $n = 2k$. Let us consider the game where players $\{0, 1, ..., k-1\}$ have a minority payoff function and players $\{k, k+1, ..., n-1\}$ have a majority payoff function; in addition we want that for exact proportion $p = 1/2$, the majority players will prefer to play 1 and the minority players will prefer to play 0. Formally, we define the payoff of players $i = 0, 1, ..., k-1$ to be $u^i(0, p) = 0$ and $u^i(1, p) = 1/2 - 1/(2n) - p$ and the payoff of players $i = k, k+1, ..., n-1$ to be $u^i(0, p) = 0$ and $u^i(1, p) = p + 1/(2n) - 1/2$.

Assume that the dynamic starts at the proportion $p = 0$ where all players play 0. First players $\{0, 1, ..., k-1\}$ will update their action to 1, because at the updating moment 1 is the minority action. Afterwards, players $\{k, k+1, ..., n-1\}$ will update their action to 1, because at the updating moment 1 is the majority action. Then, players $\{0, 1, ..., k-1\}$ will update their actions to action 0 because 0 is the minority, then players $\{k, k+1, ..., n-1\}$ would update their actions to action 0, and so forth.

In this example, at steps where $p(a) \in (1/2 - \varepsilon, 1/2 + \varepsilon)$ the played action profile is a pure Nash ε-equilibrium. If we change only the payoffs of player 0 and $n-1$ to be $u^0(0, p) = 0$,$u^0(1, p) = 1/3 - p$, $u^{n-1}(0, p) = 0$,$u^{n-1}n(1, p) = 2/3 - p$, then the dynamic remains in the same cycle, but the played action profiles including those action profiles where $p(a) \in (1/2 - \varepsilon, 1/2 + \varepsilon)$ are no longer pure Nash ε-equilibria for $\varepsilon < 1/6$.

3. Games with more than two actions. In the case where players have more than two actions, best-reply dynamics fail to converge. Consider for example a single-population game where the payoff function of the players

Figure 2:



is the following version of the rock-paper-scissors game:

	1	2	3
1	4	5	0
2	0	4	5
3	5	0	4

Clearly, the population game presented in the above figure is a 5-Lipschitz anonymous game. In this game, all the approximate PNE are obtained for action profiles with proportion $p(a)$ in a neighborhood of $(1/3, 1/3, 1/3)$. In this classical game evolutionary dynamics such as best-reply, fictitious-play, or replicator dynamic converge to the cycle presented in Figure 2; see [15]. Our version of a best-reply dynamic is also most of the time near this cycle.

4. Use of different rounding functions for the observable proportion. In the proof of the main theorems we only used the fact that ϕ is piecewise constant over segments of size $\Omega(1)$. In the case where every player observes the proportion according to some different rounding, the function ϕ is still pairwise constant; therefore the same proofs still hold in this case.

5. The case where players observe exact proportions. Whether there is convergence, and if so, the rate of convergence when every player i observes the *exact* proportion $p(a)$ or the *exact* proportion $p(a^{-i})$ (rather than ε-rounding), remains an open question. I conjecture that the resulting best-reply dynamic does indeed converge to[16] PNE_{ε}.

References

[1] Arieli, I. and Young, P.H. (2011) "Fast Convergence in Population Games", *Department of Economics, University of Oxford,* DP 15262.

[2] Awerbuch, B., Azar, Y., Epstein, A., Mirrokni, V.S. and Skopalik, A (2008) "Fast Convergence to Nearly Optimal Solutions in Potential Games", *EC '08 Proceedings of the 9th ACM Conference on Electronic Commerce* 2, 264–273.

[3] Azrieli, Y., Shmaya, E. (2011) "Lipschitz Games", *Ohio State University* Mimeo.

[4] Daskalakis, C and Papadimitriou, C.H. (2007) "Computing Equilibria in Anonymous Games", *The 48th Annual IEEE Symposium on Foundations of Computer Science* 83–93.

[16] The intuition behind this conjecture is that, in most cases, it appears that the expected time for a proportion to move by more than ε (for a constant ε) is significantly longer than the expected time for choosing all the unsatisfied players, which, by Coupon Collector arguments, is $O(n \log n)$.

[5] Fischer, S., Räcke, H. and Vöcking, B. (2010) "Fast Convergence to Wardrop Equilibria by Adaptive Sampling Methods", *SIAM Journal on Computing* 39, 3700–3735.

[6] Fischer, S. and Vöcking, B. (2004) "On the Evolution of Selfish Routing", *Algorithms- ESA, Lecture Notes in Computer Scince* 3221, 323–334.

[7] Foster, D. and Young, P.H. (2006) "Regret Testing: Learning to Play Nash Equilibrium without Knowing you have an Opponent", *Theoretical Economics* 1, 341–367.

[8] Friedman, E.J., Kash, I.A. and Halpern, J.Y. (2009) "Multiagent Learning in Large Anonymous Games", *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems* 2, 765–772.

[9] Germano, F. and Lugosi, G. (2007) "Global Nash Convergence of Foster and Young's Regret Testing", *Games and Economic Behavior* 60, 135–154.

[10] Gorodeisky, Z. (2006) "Evolutionary Stability for Large Populations", *Mathematics of Operations Research* 31, 369–380.

[11] Hart, S. and Mansour, Y. (2010) "How Long to Equilibrium? The Communication Complexity of Uncoupled Equilibrium Procedures", *Games and Economic Behavior* 69, 107–126.

[12] Hart, S. and Mas-Colell, A. (2003) "Uncoupled Dynamics Do Not Lead to Nash Equilibrium", *American Economic Review* 93, 1830–1836.

[13] Hart, S. and Mas-Colell, A. (2006) "Stochastic Uncoupled Dynamics and Nash Equilibrium", *Games and Economic Behavior* 2, 286–303.

[14] Hilton, P. and Pedersen, J. (1991) "Catalan Numbers, Their Generalization, and Their Uses", *The Mathematical Intelligencer* 13, 64–75.

[15] Hofbauer, J. and Sigmund, K. (1998) *Evolutionary Games and Population Dynamics*, Cambridge University Press.

[16] Rashid, S. (1983) "Equilibrium Points of Nonatomic Games: Asymptotic Results", *Economics Letters* 12, 7–10.

[17] Schmeidler, D. (1973) "Equilibrium Points of Nonatomic Games", *Journal of Statistical Physics* 7, 295–300.

[18] Young, P.H. (2009) "Learning by Trial and Error", *Games and Economic Behavior* 65, 626–643.