Learning Equilibria in Games by Stochastic Distributed Algorithms.

Olivier Bournez and Johanne Cohen*

Abstract We consider a family of stochastic distributed dynamics to learn equilibria in games, that we prove to correspond to an Ordinary Differential Equation (ODE). We focus then on a class of stochastic dynamics where this ODE turns out to be related to multipopulation replicator dynamics. Using facts known about convergence of this ODE, we discuss the convergence of the initial stochastic dynamics. For general games, there might be non-convergence, but when the convergence of the ODE holds, considered stochastic algorithms converge towards Nash equilibria. For games admitting a multiaffine Lyapunov function, we prove that this Lyapunov function is a super-martingale over the stochastic dynamics and that the stochastic dynamics converge. This leads a way to provide bounds on their time of convergence by martingale arguments. This applies in particular for many classes of games considered in literature, including several load balancing games and congestion games.

1 Introduction

Consider a scenario where agents learn from their experiments, by small adjustments. This might be for example about choosing their telephone companies, or about their portfolio investments. We are interested in understanding when the whole market can converge towards rational situations, i.e. Nash equilibria in the sense of game theory. This is natural to expect dynamics of adjustments to be

---

Olivier Bournez
LIX-CNRS, 91128 Palaiseau Cedex,FRANCE.
e-mail: bournez@lix.polytechnique.fr

Johanne Cohen
PRiSM-CNRS, Univ. Versailles, 45 av. des Etats-Unis, 78000 Versailles, FRANCE
e-mail: Johanne.Cohen@prism.uvsq.fr

* The authors are supported in part by the ANR SHAMAN project.

stochastic, and fully distributed, since we expect agents to adapt their strategies based on their local knowledge of the market.

Several such dynamics of adjustments have been considered recently in the literature. Up to our knowledge, this has been done mainly for deterministic dynamics or best-response based dynamics: computing a best response requires a global description of the market. Stochastic variations, avoiding a global description, have been considered. However, considered dynamics are somehow rather ad-hoc, in order to get efficient convergence time bounds. We want to consider here more general dynamics related to (possibly perturbed) replicator dynamics, and discuss when one may expect convergence.

Basic game theory framework. Let $[n] = \{1,\ldots,n\}$ be the set of players. Every player i has a set $\mathscr{S}_i$ of *pure strategies*. Let m_i be the cardinal of $\mathscr{S}_i$. A *mixed strategy* $q_i = (q_{i,1}, q_{i,2}, \ldots, q_{i,m_1})$ corresponds to a probability distribution over pure strategies: pure strategy ℓ is chosen with probability $q_{i,\ell} \in [0,1]$, with $\sum_{\ell=1}^{m_i} q_{i,\ell} = 1$. Let K_i be the simplex of mixed strategies for player i. Any pure strategy ℓ can be considered as mixed strategy e_ℓ, where vector e_ℓ denotes the unit probability vector with ℓ^{th} component unity, hence as a corner of K_i.

Let $K = \prod_{i=1}^{n} K_i$ be the space of all mixed strategies. A *strategy profile* $Q = (q_1, ..., q_n) \in K$ specifies the strategies of all players: q_i corresponds to the mixed strategy of player i. In game theory, we often write $Q = (q_i, Q_{-i})$, where Q_{-i} denotes the vector of the strategies played by all other players. We admit games whose payoffs may be random: we assume that each player i gets a random *cost* of expected value $c_i(Q)$. In particular, the expected cost for player i for playing the pure strategy e_ℓ is denoted by $c_i(e_\ell, Q_{-i})$.

Some classes of games. Several games where players' costs are based on the shared usage of a common set of resources $[m] = \{1,2,\ldots,m\}$ where each resource $1 \leq r \leq m$ has an associated nondecreasing cost function denoted by $C_r : [n] \rightarrow \mathbb{R}$, have been considered in algorithmic game theory literature.

In *load balancing games* [9], the machines are the resources, and the players (task) choose a machine to execute : each player i has a weight w_i. The cost for player i under profile of pure strategies (assignment) $Q = (q_1, \ldots, q_n)$ corresponds to $c_i(Q) = C_{q_i}(\lambda_{q_i}(Q))$, where $\lambda_r(Q)$ is the load of machine r: $\lambda_r(Q) = \sum_{j: q_j = r} w_j$. In *congestion games* [13], the players compete for subsets of $[m]$. Hence, the pure strategy space $\mathscr{S}_i$ of player i is a subset of $2^{[m]}$ and a pure strategy $q_i \in Q$ for player i is a subset of $[m]$ resources. The cost of player i under profile of pure strategies Q corresponds to $c_i(Q) = \sum_{r \in q_i} C_r(\lambda_r(Q))$ where $\lambda_r(Q)$ is the number of q_j with $r \in q_j$.

Ordinal and potential games. All these classes of games can be related to potential games introduced by [11]: A game is an *ordinal potential game* if there exists some function ϕ from *pure* strategies to $\mathbb{R}$ such that for all pure strategies Q_{-i}, q_i, and q'_i, one has $c_i(q_i, Q_{-i}) - c_i(q'_i, Q_{-i}) > 0$ iff $\phi(q_i, Q_{-i}) - \phi(q'_i, Q_{-i}) > 0$. It is an *an exact potential game* if for all pure strategies Q_{-i}, q_i, and q'_i, one has $c_i(q_i, Q_{-i}) - c_i(q'_i, Q_{-i}) = \phi(q_i, Q_{-i}) - \phi(q'_i, Q_{-i})$.

2 Stochastic Learning Algorithms

We consider fully distributed algorithms of the following form where b is a positive real parameter close to 0.

Let $Q(t) = (q_1(t),...,q_n(t)) \in K$ denote the state of all players at instant t. Our interest is in the asymptotic behavior of $Q(t)$, and its possible convergence to Nash equilibria. Functions $F_i^b(c_i(t), s_i(t), q_i(t)))$ is defined as generic as possible, maintaining that the $q_i(t)$ always stays validity probability vectors. We only assume that $G_i(Q) = \lim_{b\to 0} \mathbb{E}[\, F_i^b(c_i(t), s_i(t), q_i(t)) \,|Q(t)\,]$ is always defined and that G_i is continuous.

• Initially, $q_i(0) \in K_i$ can be any vector of probability, for all i.
• At each round t,
  • Any player i selects strategy $\ell \in \mathscr{S}_i$ with probability $q_{i,\ell}(t)$. This leads to a cost $c_i(t)$ for player i.
  • Select some player $i(t)$: player $i(t)$ is selected with probability p_i, with $\sum_{i=1}^n p_i = 1$.
    This player $i(t)$ updates $q_i(t)$ as follows: $q_i(t+1) = q_i(t) + bF_i^b(c_i(t), s_i(t), q_i(t))$;
    Any other player keeps $q_i(t)$ unchanged: $q_i(t+1) = q_i(t)$.

Results. In the general case (Theorem 1), any stochastic algorithm in the considered class converges (see in [2]) weakly towards solutions of initial value problem (ordinary differential equation (ODE)) $\frac{dq_i}{dt} = p_i G_i(Q)$, given $Q(0)$.

A *replicator-like dynamics* F_i^b is a dynamic where $F_i^b(c_i(t), s_i(t), q_i(t)) = \gamma(c_i(t))$ $(q_i(t) - e_{s_i(t)}) + \mathscr{O}(b)$, where $\gamma : \mathbb{R} \to [0,1]$ is some *decreasing* function with value in $[0,1]$. We assume all costs to be positive, by linearity of expectation then all costs must be bounded by some constant M, and we can take $\gamma(x) = \frac{M-x}{M}$.

We can admit randomly perturbed dynamics: $\mathscr{O}(b)$ denotes some perturbation that stay of order of b. A *perturbed replicator-like dynamic* is of the form

$$F_i^b(r_i(t), s_i(t), q_i(t)) = \mathscr{O}(b) + \begin{cases} \gamma(r_i(t))(q_i(t) - e_{s_i(t)}) & \text{with probability } \alpha \\ b(q_i(t) - e_{s_j}) & \text{with probability } 1-\alpha, \\ & \text{where } j \in \{1,\dots,m_i\} \text{ is chosen uniformly,} \end{cases}$$

where $0 < \alpha < 1$ is some constant.

We prove that such dynamics have a mean-field approximation which is isomorphic to a multipopulation replicator dynamics. We claim (Theorem 2), that for general games, *if* there is convergence of the mean-field approximation, then stable limit points will correspond to Nash equilibria. Notice, that there is no reason that the convergence of mean-field approximation holds for generic games. We note (Theorem 3) that the ordinal games are *Lyapunov* games: their mean-field limit approximation admits some Lyapunov function. Furthermore, we show that for Lyapunov games with multiaffine Lyapunov function, the Lyapunov function is a supermartingale over stochastic dynamics. Finally, we deduce results on the convergence of stochastic algorithms for this class.

For lack of space, we refer to [2] for missing proofs.

Related work. A potential game always have a pure Nash equilibrium: since ordinal potential function, that can take only a finite number of values, is strictly decreasing in any sequence of pure strategies strict best response moves, such a sequence must be finite and must lead to a Nash equilibrium [13].This is clear that an (exact) potential game is an ordinal potential game. Congestion games, and hence load balancing games are known to be particular potential games [13].

For load-balancing games, the bounds on the convergence time of best-response dynamics have been investigated in [5]. Since players play in turns, this is often called the *Elementary Stepwise System*. Other results of convergence in this model, have been investigated in [7, 10], but they require some global knowledge of the system in order to determine what next move to choose. A Stochastic version of best-response dynamics has been investigated in [1].For congestion games, the problem of finding pure Nash equilibria is PLS-complete [8]. Efficient convergence of best-response dynamics to approximate Nash equilibria in particular symmetric congestion games have been investigated in [3] in the case where each resource cost function satisfies a *bounded jump assumption*.

All previous discussions are about best-response dynamics. A stochastic dynamic, not elementary stepwise like ours, but close to those considered in this paper, has been partially investigated in [12] for general games and for potential games: It is proved to be weakly convergent to solutions of a multipopulation replicator equation. Some of our arguments follow theirs, but notice that their convergence result (Theorem 3.1) is incorrect: convergence may happen towards non-Nash (unstable) stationary points. Furthermore, this is not clear that any super-martingale argument holds for such dynamics, as our proof relies on the fact that the dynamics is elementary stepwise.

Replicator equations have been deeply studied in evolutionary game theory [15]. Evolutionary game theory has been applied to routing problems in the Wardrop traffic model in [6]. Potential games have been generalized to continuous player sets in [14]. They have be shown to lead to multipopulation replicator equations, and since our dynamics are not about continuous player sets, but lead to similar dynamics, we borrow several constructions from [14]. No time convergence discussion is done in [14]. Moreover, in [4], a replicator equation for the routing games and for particular allocation games are studied to converge to a pure Nash equilibrium.

3 Mean-Field Approximation For Generic Algorithms

We focus on the evolution of $Q(t)$, where $Q(t) = (q_1(t), ..., q_n(t))$ denotes the strategy profile at instant t in the stochastic algorithm. Clearly, $Q(t)$ is an homogeneous Markov chain. Define $\Delta Q(t)$ as $\Delta Q(t) = Q(t+1) - Q(t)$, and $\Delta q_i(t)$ as $q_i(t+1) - q_i(t)$. We can write

$$\mathbb{E}[\ \Delta q_i(t)\ |Q(t)\ ] = bp_i \mathbb{E}[\ F_i^b(c_i(t), s_i(t), q_i(t))\ |Q(t)\ ], \tag{1}$$

with $G_i(Q) = \lim_{b \to 0} \mathbb{E}[\ F_i^b(c_i(t), s_i(t), q_i(t))\ |Q(t)\ ]$ assumed to be continuous under our hypotheses.

Convergence of the stochastic algorithms towards ODEs defining their mean-field limit approximation can be formalized as follows: Consider the piecewiselinear interpolation $Q^b(.)$ of $Q(t)$ defined by $Q^b(t) = Q(\lfloor t/b \rfloor) + (t/b - \lfloor t/b \rfloor)$ $(Q(\lfloor t/b+1 \rfloor) - Q(\lfloor t/b \rfloor))$. Function $Q^b(.)$ belongs to the space of all functions from $\mathbb{R}$ into K which are right continuous and have left hand limits (*cad-lag functions*). Now consider the sequence $\{Q^b(.) : b > 0\}$. We are interested in the limit $Q(.)$ of this sequence when $b \to 0$. Recall that a family of random variable $(Y_t)_{t\in\mathbb{R}}$ weakly converges (see in [2]) to a random variable Y, if $E[h(X_t)]$ converges to $E[h(Y)]$ for each bounded and continuous function h.

Theorem 1. *The sequence of interpolated processes {$Q^b(.)$} converges weakly, when $b \to 0$, to $Q(.)$, which is the solution of initial value problem*

$$\frac{dq_i}{dt} = p_i G_i(Q),\ i = 1, \cdots, n, \quad \textit{with } Q(0) = Q^b(0). \tag{2}$$

4 General Games and Replicator-Like Dynamics

Now, we restrict to (possibly perturbed) replicator-like dynamics, as defined in page 3. For any such dynamic (full details in [2]), Equation (2) leads to the following ordinary differential equation which turns out to be (a rescaling of) (multipopulation) classical replicator dynamic

$$\frac{dq_{i,\ell}}{dt} = p_i q_{i,\ell}(c_i(q_i, Q_{-i}) - c_i(e_\ell, Q_{-i})), \tag{3}$$

whose limit points are related to Nash equilibria (see in [2]). Using properties of dynamics (3), we get:

Theorem 2. *For general games, for any replicator-like or perturbed replicator-like dynamic, the sequence of interpolated processes {$Q^b(.)$} converges weakly, as $b \to 0$, to the unique deterministic solution of dynamic*(3) *with* $Q(0) = Q^b(0)$. If *the mean-field approximation dynamic* (3) *converges, its stable limit points correspond to Nash equilibria of the game.*

More precisely (see in [2]), the following are true for solutions of dynamic (3): (i) All Nash equilibria are stationary points. (ii) All stable stationary points are Nash equilibria. (iii) However, (unstable) stationary points can include some non-Nash equilibria.

Actually, all corners of simplex K are stationary points, as well as, from the form of (3), more generally any state Q in which all strategies in its support perform equally well. Such a state Q is not a Nash equilibrium as soon as there is an not used strategy (i.e. outside of the support) that performs better.

Unstable limit stationary points may exist for the mean-field approximation. Consider for example a dynamics that leave on some face of K where some well-performing strategy is never used. To avoid "bad" (non-Nash equilibrium, hence unstable) stationary points, following the idea of penalty functions for interior point methods, one can use as in Appendix A.3 of [14] some patches on the dynamics that would guarantee Non-complacency (see in [2]). *Non-Complacency (NC)* is the following property: $G(Q) = 0$ implies that Q is a Nash equilibrium (3) (i.e. stationarity implies Nash).

For general games, we get that the limit for $b \to 0$ is some ordinary differential equation whose stable limit points, when $t \to \infty$, *IF* there exist, can only be Nash equilibria. Hence, *IF* there is convergence of the ordinary differential equation, then one expects the previous stochastic algorithms to learn equilibria.

5 Lyapunov Games, Ordinal and Potential Games

Since general games have no reason to converge, we propose now to restrict to games for which replicator equation dynamic or more generally general dynamics (2) is provably convergent.

Definition 1 (Lyapunov Game). We say that a game has a *Lyapunov function* (with respect to a particular dynamic (2) over K), or that the game is *Lyapunov*, if there exists some non-negative $\mathscr{C}^1$ function $F : K \to \mathbb{R}$ such that for all i, ℓ and Q, whenever $G(Q) \neq 0$,

$$\sum_{i,\ell} p_i \frac{\partial F}{\partial q_{i,\ell}}(Q) G_{i,\ell}(Q) < 0. \tag{4}$$

Lyapunov games include ordinal potential games: we will say that a Lyapunov function $F : K \to \mathbb{R}$ is multiaffine, if it is defined as as polynomial in all its variables, it is of degree 1 in each variable, and none of its monomials are of the form $q_{i,\ell} q_{i,\ell'}$.

Theorem 3. *An ordinal potential game is a Lyapunov game with respect to dynamics* (3)*. Furthermore, its has some multiaffine Lyapunov function.*

If ϕ is the potential of the ordinal potential game, then one can take its expectation $F(Q) = \mathbb{E}[\ \phi(Q) \mid \text{players play pure strategies according to } Q\ ]$as a Lyapunov function with respect to dynamics (3). The following class of games have been introduced [14, 12].

Definition 2 (Potential Game [14]). A game is called a *continuous potential game* if there exists a $\mathscr{C}^1$ function $F : K \to \mathbb{R}$ such that for all i, ℓ and Q,

$$\frac{\partial F}{\partial q_{i,\ell}}(Q) = c_i(e_\ell, Q). \tag{5}$$

Proposition 1. *A continuous potential game is a Lyapunov game with respect to dynamics* (3)*. Its has some multiaffine Lyapunov function.*

Proposition 2. *An (exact) potential game of potential ϕ leads to a continuous potential game with $F(Q) = \mathbb{E}[\ \phi(Q)\ ]$, and conversely, the restriction of F of class $\mathscr{C}^2$ to pure strategies of a potential in the sense of above definition leads to an (exact) potential.*

A Lyapunov game can have some non-multiaffine potential function, hence not all Lyapunov games with respect to dynamics (3) are ordinal games. The interest of Lyapunov functions is that they provide convergence. Observing that all previous classes are Lyapunov games with respect to dynamics (3), this gives the full interest of this corollary.

Corollary 1. *In a Lyapunov game with respect to general dynamics* (3)*, whatever the initial condition is, the solutions of mean-field approximation* (2) *will converge. The stable limit points are Nash equilibria.*

6 Replicator Dynamics for Multiaffine Lyapunov Games

Fortunately, this is possible to go further, observing that many of the previous classes turn out to have a multiaffine Lyapunov function. The key observation is the following (the proof mainly relies on the fact that second order terms are null for multiaffine functions).

Lemma 1. *When F is a multiaffine Lyapunov function,*

$$\mathbb{E}[\,\Delta F(Q(t+1))\,|Q(t)\,] = \sum_{i=1}^{n}\sum_{\ell=1}^{m_i} \frac{\partial F}{\partial q_{i,\ell}}(Q(t))\mathbb{E}[\,\Delta q_{i,\ell}\,|Q(t)\,], \tag{6}$$

where $\Delta F(t) = F(Q(t+1)) - F(Q(t))$.

Notice that for Lyapunov game with a multiaffine Lyapunov function F, with respect to Dynamic (3) (this include ordinal, and hence potential games from above discussion), the points Q^* realizing the minimum value F^* of F over compact K must correspond to Nash equilibria.

Fortunately, this is possible to get bounds on the expected time of convergence (see in [2]): we write $L(\mu)$ for the subset of states Q on which $F(Q) \leq \mu$.

Definition 3 (ε-Nash equilibrium). Let $\varepsilon \geq 0$. A state Q is some ε-Nash equilibrium iff for all $1 \leq i \leq n, 1 \leq \ell \leq m_i$, we have $c_i(e_\ell, Q_{-i}) \geq (1-\varepsilon)c_i(q_i, Q_{-i})$.

Theorem 4. *Consider a Lyapunov game with a multiaffine Lyapunov function F, with respect to* (3)*. This includes ordinal, and hence potential games from above discussion. Taking $b = \mathscr{O}(\varepsilon)$, whatever the initial state of the stochastic algorithm is, it will almost surely reach some ε-Nash equilibrium. Furthermore, it will do it in a random time whose expectation $T(\varepsilon)$ satisfies $T(\varepsilon) \leq \mathscr{O}(\frac{F(Q(0))}{\varepsilon})$.*

References

1. P. Berenbrink, T. Friedetzky, L. A. Goldberg, P. Goldberg, Z. Hu, and R. Martin. Distributed selfish load balancing. In *SODA*, pages 354–363, ACM, 2006.

2. O. Bournez and J. Cohen. Learning equilibria in games by stochastic distributed algorithms. *Université de Versailles,CoRR*, abs/0907.1916, 2009.
3. S. Chien and A. Sinclair. Convergence to approximate nash equilibria in congestion games. In *SODA*, pages 169–178, 2007.
4. P. Coucheney, C. Touati, and B. Gaujal. Fair and efficient user-network association algorithm for multi-technology wireless networks. In *INFOCOM*, 2009.
5. E. Even-Dar, A. Kesselman, and Y. Mansour. Convergence time to Nash equilibrium in load balancing. *ACM Transactions on Algorithms*, 3(3), 2007.
6. S. Fischer, H. Räcke, and B. Vöcking. Fast convergence to Wardrop equilibria by adaptive sampling methods. *Proceedings of STOC*, pages 653–662, 2006.
7. Paul W. Goldberg. Bounds for the convergence rate of randomized local search in a multi-player load-balancing game. In *PODC '04*, pages 131–140, 2004.
8. D.S. Johnson, C.H. Papadimitriou, and M. Yannakakis. How easy is local search? *Journal of Computer System Science*, 37(1):79–100, 1988.
9. E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *Symposium on Theoretical Computer Science (STACS'99)*, pages 404–413, 1999.
10. L. Libman and A. Orda. Atomic Resource Sharing in Noncooperative Networks. *Telecommunication Systems*, 17(4):385–409, 2001.
11. D. Monderer and L.S. Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996.
12. M.A.L. Thathachar P.S. Sastry, V.V. Phansalkar. Decentralized learning of Nash equilibria in multi-person stochastic games with incomplete information. *IEEE transactions on system, man, and cybernetics*, 24(5), 1994.
13. R.W. Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2(1):65–67, 1973.
14. W.H. Sandholm. Potential games with continuous player sets. *Journal of Economic Theory*, 97(1):81–108, 2001.
15. J. W. Weibull. *Evolutionary Game Theory*. The MIT Press, 1995.