Learning Equilibria in Games by Stochastic Distributed Algorithms.

Olivier Bournez

Ecole Polytechnique,
Laboratoire LIX
91128 Palaiseau Cedex,
FRANCE
Olivier.Bournez@lix.polytechnique.fr

Johanne Cohen

Centre National de la Recherche Scientifique,
Laboratoire PRiSM
Université de Versailles
45, avenue des Etats-Unis, 78000 Versailles,
FRANCE
Johanne.Cohen@prism.uvsq.fr

Submitted to SODA 2010
on July 6th 2009.

Abstract

We consider a class of fully stochastic and fully distributed algorithms, that we prove to learn equilibria in games. Indeed, we consider a family of stochastic distributed dynamics that we prove to converge weakly (in the sense of weak convergence for probabilistic processes) towards their mean-field limit, i.e an ordinary differential equation (ODE) in the general case. 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 convergence of the ODE holds, considered stochastic algorithms converge towards Nash equilibria. For games admitting Lyapunov functions, that we call Lyapunov games, the stochastic dynamics converge. We prove that any ordinal potential game, and hence any potential game is a Lyapunov game, with a multiaffine Lyapunov function. For Lyapunov games with a multiaffine Lyapunov function, we prove that this Lyapunov function is a super-martingale over the stochastic dynamics. This leads a way to provide bounds on their time of convergence by martingale arguments. This applies in particular for many classes of games that have been considered in literature, including several load balancing game scenarios 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 stochastic, and fully distributed, since we expect agents to adapt their strategies based on their local knowledge of the market, and since agents are often involved in games where a global, and hence local, deterministic description of the whole global market is not possible.

Several such dynamics of adjustments have been considered recently in the algorithmic game theory 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, and still mainly best-response based. We want to consider here more general dynamics, and discuss when one may expect convergence. This could lead to consider any dynamics which is monotone with respect to the utility of players, in relation with evolutionary game theory literature [**?**]. We propose to restrict here to dynamics that lead to dynamics related to (possibly perturbed) replicator dynamics.

Somehow, as algorithmic game theory can be seen as an algorithmic version of classical game theory, our long term aim is to better understand algorithmic evolutionary game theory. Somehow, we could also say, that as best-response dynamics can be seen as strategies that visit corners of the simplex of (mixed) strategies, we are interested in a long term objective in learning methods that could be seen as interior point methods to find equilibria.

Basic game theory framework. Let $[n] = \{1, \dots, n\}$ be the set of players. Every player i has a set $\mathcal{S}_i$ of *pure strategies*. Let m_i be the cardinal of $\mathcal{S}_i$. A *mixed strategy* $q_i = (q_{i,1}, q_{i,2}, \dots, q_{i,m_1})$ 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 (mixed or pure) strategies of all players: q_i corresponds to the mixed strategy played by player i. Following classical convention, we write often write abusively $Q = (q_i, Q_{-i})$, where Q_{-i} denotes the vector of the strategies played by all other players.

We allow games whose payoffs may be random: we only assume that whenever the strategy profile $Q \in K$ is known, each player i gets a random *cost* of expected value $c_i(Q)$. In particular, the expected cost for player i for playing pure strategy e_ℓ is denoted by $c_i(e_\ell, Q_{-i})$.

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

In *load balancing games* [**?**], resources are called machines, and players compete for elements (i.e. singleton subsets) of $[m]$. Hence, the pure strategy space $\mathcal{S}_i$ of player i having a weight w_i corresponds to $[m]$ or a subset of $[m]$, and a pure strategy $q_i \in \mathcal{S}_i$ for player i is some element $r \in [m]$. The cost for player (task) i under profile of pure strategies (assignment) $Q = (q_1, \dots, q_n)$

corresponds to $c_i(Q) = C_{q_i}(\lambda_{q_i}(Q))$, where $\lambda_r(Q)$ is the load of machine r: $\lambda_r(Q) = \sum_{j:q_j=r} w_j$, that is to say the sum of the weights of the tasks running on it.

In *congestion games* [?], resources are called edges, and players compete for subsets of $[m]$. Hence, the pure strategy space $\mathcal{S}_i$ of player i is a subset of $2^{[m]}$ and a pure strategy $q_i \in Q$ for player i is a subset of $[m]$. 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 $r \in q_j$. In *weighted congestion games*, weights $(w_i)_i$ are associated to players, and one takes instead $\lambda_r(Q) = \sum_{j:r \in q_j} w_j$.

In *task allocation games* [?], as in load balancing games, resources are called machines, and players compete for elements (i.e. singleton subsets) of $[m]$. Each resource (machine) r is assumed to have a function C_r that takes as input a set of tasks $\lambda \subset [n]$ assigned to it, and outputs a cost $C_{r,j}$ for each participating player j. The cost of player i under profile of pure strategies Q is then given by $c_i(Q) = C_{q_i,i}(\{j | q_j = q_i\})$. Functions C_r can be considered as speed and scheduling policies, and associated costs as corresponding completion time for player (task) i. For example, SPT and LPT are policies that schedule the jobs without preemption respectively in order of increasing or decreasing weights (processing times) [?].

Clearly, load balancing games are particular task allocation games, and load balancing games are particular weighted congestion games. A load balancing game whose weights are unitary is a particular congestion game.

Ordinal and potential games. All these classes of games can be related to ordinal and potential games introduced by [?]: 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

Generic Stochastic Learning Algorithm. We want basically to consider learning algorithms of the following form, over the most possible general games, where b is a parameter, intended to be positive but close to 0.

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

In a first step, consider functions $F_i^b(r_i(t), s_i(t), q_i(t)))$ as generic as possible, maintaining that the $q_i(t)$ always stay validity probability vectors: that is to say, $q_{i,\ell}(t) \in [0,1]$ and $\sum_\ell q_{i,\ell}(t) = 1$ is preserved. Functions $F_i^b(r_i(t), s_i(t), q_i(t))$ can be random (formally a random variable). We only assume that its expectation $\mathbb{E}[\ F_i^b(r_i(t), s_i(t), q_i(t))\ |Q(t)\ ]$ is always defined.

This corresponds indeed to fully distributed algorithms[1]. Decisions made by players are com-

[1] We of course understand that for some games (like congestion games), the size of the involved probability vectors might be non-polynomial. However, by restricting to function $F_i^b(r_i(t), s_i(t), q_i(t)))$, or close dynamics, which guarantee a support of polynomial size for $q_i(t)$, can solve the problem: restrict to function which are equal to

pletely decentralized: At each time step, player i only needs r_i and q_i, that is to say respectively her cost and her current mixed strategy, to update his own strategy q_i.

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. Assume that $G_i(Q) = \lim_{b \to 0} \mathbb{E}[\ F_i^b(r_i(t), s_i(t), q_i(t))\ |Q\ ]$ exists and is some continuous function G_i of Q.

Results. In the general case (Theorem 1), any stochastic algorithm in the considered class converges[2] weakly (in the sense of weak convergence for probabilistic processes) towards solutions of initial value problem (ordinary differential equation (ODE)) $\frac{dq_i}{dt} = p_i G_i(Q)$, given $Q(0)$, i.e. to its mean-field limit approximation.

A *replicator-like dynamics* F_i^b is a dynamic where

$$F_i^b(r_i(t), s_i(t), q_i(t)) = \gamma(r_i(t))(e_{s_i(t)} - q_i(t)) + \mathcal{O}(b),$$

or where this holds for its expectation, where $\gamma : \mathbb{R} \to [0,1]$ is some *decreasing*[3] function with value in $[0,1]$. Recall that $e_{s_i(t)}$ is the unit vector of dimension m_i with component number $s_i(t)$ unity.

Notice that we allow perturbed dynamics: $\mathcal{O}(b)$ denotes some perturbation that stay of order of parameter b.

We can also allow randomly perturbed dynamics: a *perturbed replicator-like dynamic* is of the form

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

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

We claim that such dynamics have a mean-field approximation which is isomorphic to a multi-population 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 of the game. Notice, that there is no reason that convergence of mean-field approximation holds for generic games, but if it holds, then its stable limit points will be Nash equilibria.

We claim (Theorem 3) that ordinal games (and hence (exact) potential games) are *Lyapunov* games: their mean-field limit approximation admits some Lyapunov function. Furthermore, this Lyapunov function, that can be taken as the expectation of the potential and is of a special type, that we call *multiaffine.*

We show that for Lyapunov games with multiaffine Lyapunov function (hence this includes ordinal and (exact) potential games such as load balancing, task allocation and congestion games), the Lyapunov function is a super-martingale over stochastic dynamics.

We deduce results on the convergence of stochastic algorithms for this class. We claim (Theorem 4) that for generic Lyapunov games with multiaffine Lyapunov function, the convergence towards Nash equilibria happens in expected time of order $\frac{F(Q(0))}{\epsilon}$, taking b of order ϵ.

$-bq_{i,\ell}$ for components ℓ outsides a polynomial (or fixed) sized support, for example. If this is too problematic to our reader, please consider that we restrict to games where the m_i stay polynomial, as for load balancing games and task allocation games.

[2] See discussion in Appendix A.1.1.

[3] If 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 for example $\gamma(x) = \frac{M-x}{M}$.

Related work. 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 (exact) potential games [?]. Actually, it is known that a game is an (exact) potential game iff its is isomorphic to a congestion game [?]. It has been proved in [?] that task allocation games are ordinal potential games, for SPT and LPT policies. In fact, it is proved in [?] that under SPT and LPT policies, one can build some function ϕ, which takes values of the form $(l_1, \cdots, l_n)$, that is lexicographically decreasing iff a player is doing a best response move. As the l_i (which corresponds to loads) are bounded by some constant K, function $\phi = \sum_i l_i K^{n-i}$ is decreasing iff a player is doing a best response move.

In other words, task allocation games under SPT and LPT policies are indeed ordinal potential games, under the terminology of [?].

An ordinal 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 [?]. This proof of existence of pure Nash equilibria can be turned into a dynamic: players play in turn, and move to resources with a lower cost.

For load-balancing games, following this idea, bounds on the convergence time of best-response dynamics have been investigated in [?]. Since players play in turns, this is often called the *Elementary Stepwise System*. Other results of convergence in this model, have been investigated in [?, ?, ?], but all 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 [?, ?]. It is proved to terminate in expected $O(\log \log n + m^4)$ rounds for uniform tasks, and uniform machines. This has been extended to weighted tasks and uniform machines in [?]. The expected time of convergence to an ϵ-Nash equilibrium is in $\mathcal{O}(nmW^3\epsilon^2)$ where W denotes the maximum weight of any task.

For congestion games, the problem of finding pure Nash equilibria in congestion games is PLS-complete [?]. Efficient convergence of particular best-response dynamics to approximate Nash equilibria in symmetric congestion games have been investigated in [?], in the particular case where each resource cost function satisfies a *bounded jump assumption*. In this context, the convergence to ϵ-Nash equilibria occurs within a number of steps that is polynomial in the number of players. This has been extended to different classes of asymmetric congestion games in [?].

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 [?] 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 [?, ?]. Evolutionary game theory has been applied to routing problems in the Wardrop traffic model in [?, ?]. Potential games have been generalized to continuous player sets in [?]. 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 [?]. No time convergence discussion is done in [?].

A replicator equation for the routing games considered has been considered in [?], where a Lyapunov function is established. The dynamics considered in [?] considers marginal costs. Moreover, in [?, ?], the replicator dynamics for particular allocation games are studied to converge to a pure Nash equilibrium by modified the game cost in order to obtain Lyapunov function.

3 Mean-Field Approximation For Generic Stochastic Algorithms

Recall that we are interested in discussing the evolution of $Q(t)$, where $Q(t) = (q_1(t), ..., q_n(t)) \in K$ denotes the state of the player team 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

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

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

Convergence of the stochastic algorithms towards ordinary differential equations defining their mean-field limit approximation can be formalized as follows: Consider the piecewise-linear 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[4] to a random variable Y, if $E[h(X_t)]$ converges to $E[h(Y)]$ for each bounded and continuous function h[5].

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

$$\frac{dq_i}{dt} = p_i G_i(Q),\ \ i = 1, \cdots, n, \tag{2}$$

with $Q(0) = Q^b(0)$.

4 General Games and Replicator-Like Dynamics

From now on, we restrict to (possibly perturbed) replicator-like dynamics, as defined in page 3.

For any such dynamic[6], Equation (2) leads to the following ordinary differential equation which turns out to be (a rescaling of) (multipopulation) classical replicator dynamic

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

whose limit points are related to Nash equilibria (through so-called Folk's theorems of evolutionary game theory [?])[7].

[4]See discussion in Appendix A.1.3.
[5]Proof of Theorem 1 can be found in Appendix A.2.
[6]Full details in Appendix A.4.
[7]Proof of Theorem 2 can be found in Appendix A.6.

Here, $u_i(Q)$ is taken as $u_i(Q) = \mathbb{E}[\ \gamma(r_i(Q))\ |Q\ ]$ for replicator-like dynamics, and $u_i(Q) = \mathbb{E}[\ \frac{1}{\alpha}\gamma(r_i(Q))\ |Q\ ]$ for perturbed replicator-like dynamics. The game whose costs are defined by u_i is clearly isomorphic to the original game. Notice that when γ is affine, this is just introducing a(n other) rescaling in (3).

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* (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[8], the following are true for solutions of (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 (3): 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 [**?**] some patches on the dynamics that would guarantee Non-complacency[9]. *Non-Complacency (NC)* is the following property: $G(Q) = 0$ implies that Q is a Nash equilibrium (3) (i.e. stationarity implies Nash).

This can be thought as the price to pay for purely deterministic models[10], and actually, when dealing with stochastic dynamics, all this can be avoided by taking profit of the unstability of non-Nash stationary points: this is the idea behind the randomized replicator dynamics already defined. This guaranteed unstable points to be left almost-surely by the associated stochastic algorithm: technically, this ensures ergodicity of the underlying Markov Chain. Notice that a purely deterministic replicator-like dynamics where $\mathcal{O}(b) = 0$ is not: an unstable stationary point, like a corner of K is invariant for ever, and the underlying Markov is hence not irreducible.

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.

Observe, that roughly speaking, for non-degenerated games, learning interior (hence mixed) Nash equilibria by such method is often problematic (and hence practically only pure Nash equilibrium may be learned) since the following is known:

Proposition 1 ([?, ?],[?, page 218]) *If a closed set $X \subset K$ belongs to the relative interior of some face of K, then X is not asymptotically stable by dynamics* (3).

[8] See discussion in Appendix A.5.
[9] See discussion in Appendix A.1.2.
[10] And perhaps somehow as artifacts of modeling.

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. As this practically often relies on some Lyapunov function argument, we propose the following terminology.

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* $\mathcal{C}^1$ *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 (and hence (exact) 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.*

Proof:
Consider $F(Q) = \mathbb{E}[\ \phi(Q) \mid \text{players play pure strategies according to probability distribution } Q\ ]$ where ϕ is the potential of the ordinal potential game. By linearity of expectation, $F(Q)$ is clearly multiaffine.

Now, by linearity of expectation, we have that $F(q_i, Q_{-i}) = \sum_\ell q_{i,\ell} F(e_\ell, Q_{-i})$, and hence $\frac{\partial F}{\partial q_{i,\ell}}(Q) = F(e_\ell, Q_{-i})$. Now, for dynamics (3), left-hand side of (4) rewrites to

$$\begin{array}{rcl}
\sum_{i,\ell} p_i \frac{\partial F}{\partial q_{i,\ell}}(Q) G_{i,\ell}(Q) & = & -\sum_i p_i \sum_\ell F(e_\ell, Q_{-i}) q_{i,\ell} (u_i(e_\ell, Q_{-i}) - u_i(q_i, Q_{-i})) \\
& = & -\sum_i p_i \sum_\ell \sum_{\ell'} q_{i,\ell} q_{i,\ell'} F(e_\ell, Q_{-i}) (u_i(e_\ell, Q_{-i}) - u_i(e_{\ell'}, Q_{-i})) \\
& = & -\frac{1}{2} \sum_i p_i \sum_{\ell < \ell'} q_{i,\ell} q_{i,\ell'} (F(e_\ell, Q_{-i}) - F(e_{\ell'}, Q_{-i})) (u_i(e_\ell, Q_{-i}) - u_i(e_{\ell'}, Q_{-i}))
\end{array}$$

Since the game is ordinal, $(F(e_\ell, Q_{-i}) - F(e_{\ell'}, Q_{-i}))(u_i(e_\ell, Q_{-i}) - u_i(e_{\ell'}, Q_{-i}))$ is always non-negative, by definition, and hence F is a Lyapunov function.

□

More precisely, if ϕ is the potential of the ordinal potential game, then one can take its expectation $F(Q) = \mathbb{E}[\ \phi(Q)\ ] = \mathbb{E}[\ \phi(Q) \mid \text{players play pure strategies according to } Q\ ]$as a Lyapunov function with respect to dynamics (3).

The following class of games have been introduced [**?**, **?**].

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

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

Proposition 2 *A continuous potential game is a Lyapunov game with respect to dynamics* (3)*. Furthermore, its has some multiaffine Lyapunov function.*

Proof:

By definition, F has a multiaffine Lyapunov function: this is clear as all its partial derivative are known, given by $\frac{\partial F}{\partial q_{i,\ell}}(Q) = c_i(e_\ell, Q)$.

Now, in this case, for dynamics (3), left-hand side of (4) rewrites to

$$\begin{array}{rcl} \sum_{i,\ell} p_i \frac{\partial F}{\partial q_{i,\ell}}(Q) G_{i,\ell}(Q) & = & -\sum_i p_i \sum_\ell u_i(e_\ell, Q_{-i}) q_{i,\ell}(u_i(e_\ell, Q_{-i}) - u_i(q_i, Q_{-i})) \\ & = & -\sum_i p_i \sum_\ell \sum_{\ell'} q_{i,\ell} q_{i,\ell'} u_i(e_\ell, Q_{-i})(u_i(e_\ell, Q_{-i}) - u_i(e_{\ell'}, Q_{-i})) \\ & = & -\frac{1}{2} \sum_i p_i \sum_{\ell<\ell'} q_{i,\ell} q_{i,\ell'} (u_i(e_\ell, Q_{-i}) - u_i(e_{\ell'}, Q_{-i}))^2 \end{array}$$

hence is positive on non-stationary points. □

Recall that exact potential games have been defined page 2, following [**?**], in terms of pure strategies. Notions turn out to be equivalent[11] when F is assumed at least $\mathcal{C}^2$.

Proposition 3 *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 $\mathcal{C}^2$ to pure strategies of a potential in the sense of above definition leads to an (exact) potential.*

Proof: In other words, a game is a continuous potential game if there exists some $\mathcal{C}^1$ function whose gradient ∇f equals the cost vector $H = (u_i(e_l, Q))_{i,l}$. Function F, which is unique up to an additive constant, is called the potential function of the game.

When F is $\mathcal{C}^2$, condition (5) is equivalent to *externality symmetry* [**?**, **?**]:

$$\frac{\partial u_i(e_\ell, Q)}{\partial q_{j,\ell'}} = \frac{\partial u_j(e'_\ell, Q)}{\partial q_{i,\ell}}, \tag{6}$$

for all i, j, ℓ, ℓ'. In that case, by a well-known result (characterization of exact forms), if we fix any $z \in K$, F is given by

$$F(Q) = \sum_{i=1}^{n} \sum_{\ell=1}^{m_i} \int_0^1 u_i(e_\ell, x(t)) x'_i(t) dt, \tag{7}$$

where $x : [0, 1] \to K$ is any piecewise continuous differentiable path in K that connects z to Q (i.e. $x(0) = z$, $x(1) = Q$).

In particular it must be independent of the used path. Considering paths from pure strategies to pure strategies, the second part of the proposition follows, from characterizations of (exact) potential games in [**?**]. The first part of the proposition is easy to establish, in the same vein as we established $\frac{\partial F}{\partial q_{i,\ell}}(Q) = F(e_\ell, Q_{-i})$ in the proof of Theorem 3 above.

□

A Lyapunov game can have some non-multiaffine potential function, hence not all Lyapunov games with respect to dynamics (3) are ordinal games. We believe Lyapunov game with respect to dynamics (3) with a multiaffine potential function to differ from ordinal games.

The interest of Lyapunov functions is that they provide convergence. Recall that the $\omega(Q_0)$ limit set of a point Q_0 is the set of accumulation points of the trajectories that start from Q_0: considering a trajectory starting from Q_0, this is the set of Q^* with $Q^* = \lim_{n\to\infty} Q(t_n)$, for some increasing sequence $(t_n)_{n\geq 0} \in \mathbb{R}$.

[11]Proof of Proposition 3 can be found in Appendix **??**.

Proposition 4 *In any Lyapunov game with respect to any dynamic* (2) *over* K*, the solutions of mean-field approximation* (2) *have their limit set* $\omega(Q)$ *non-empty, compact, connected, and consisting entirely of stationary points of the dynamic. On this limit sets,* F *is constant.*

Proof: This is made of well-known fact, and is for example present for example as Lemma A.1 of [**?**].

For self-contentedness, here is mainly a slight adaptation of the proof of Lyapunov Stability theorem [**?**, page 194].

$F(Q(t))$ must be monotone along trajectories, since Equation (4) guarantees $\frac{dF(Q(t))}{dt} < 0$. Let $Q(t)$ be some solution of ordinary differential equation (2) with $Q(0) = x$. Let $Q_0 \in \omega(x)$: that is to say $Q(t_n) \to Q_0$ for some sequence $t_n \to \infty$. We claim that Q_0 must be some stationary point of the dynamics, that is to say, $G(Q_0) = 0$. To see this, observe that $F(Q(t)) > F(Q_0)$ since $F(Q(t))$ decreases and $F(Q(t))$ converges to $F(Q_0)$ by continuity of F.

Suppose that $G(Q_0) \neq 0$. Let $Z(t)$ be the solution of the ordinary differential equation starting from Q_0. For any $s > 0$, we have $F(Z(s)) < F(Q_0)$. Hence, for any solution $Y(s)$ starting sufficiently near Z_0 we have $F(Y(s)) < F(Q_0)$. Setting $Y(0) = Q(t_n)$ for sufficiently large n yields the contradiction $F(Q(t_n + s)) < F(Q_0)$. Therefore, $G(Q_0) = 0$.

This proves that any limit set must be non-empty and consisting entirely of stationary point of the dynamics.

By continuity of F, $F(Q_0) = \lim_{n\to\infty} F(Q(t_n))$ for any limit point Q_0. Now this must be equal to $\inf_t(F(Q(t)))$ and hence independent of Q_0.

The subset $\omega(x)$ of limit points Q_0, being equal to $\cap_t Closure(F(s \geq t))$, hence a decreasing intersection of compact connected sets must be compact and connected.

□

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.*

If mean-field approximation (2) has the (NC) property, then this guarantees that limit points are Nash equilibria. Otherwise, unstable limit stationary may exist for the mean-field approximation.

6 Replicator-Like Dynamics for Multiaffine Lyapunov Games

Fortunately, this is possible to go further, observing that many of the previous classes (ordinal, (exact) potential, continuous potential, load balancing games, congestion games, task allocation games) turn out by previous discussion to have a multiaffine Lyapunov function.

When this holds, this is indeed possible to talk directly about the stochastic algorithms, avoiding passage through ordinary differential equation (2), and the double limit $b \to 0$, $t \to \infty$. 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,*

$$\mathbb{E}[\ \Delta F(Q(t+1))\ |Q(t)\ ] = \sum_{i=1}^{n}\sum_{\ell=1}^{m_i} \frac{\partial F}{\partial q_{i,\ell}}(Q(t))\mathbb{E}[\ \Delta q_{i,\ell}\ |Q(t)\ ], \tag{8}$$

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

Proof:

Let us denote $R(Q,\Delta) = F(Q+\Delta) - F(Q) - \sum_{i=1}^{n}\sum_{\ell=1}^{m_i} \frac{\partial F}{\partial q_{i,\ell}}(Q)\Delta_{i,\ell}$ when Δ is a vector, so that by definition taking $\Delta = \Delta Q(t)$, we have $\Delta F(t) = F(Q(t+1)) - F(Q(t)) = \sum_{i=1}^{n}\sum_{\ell=1}^{m_i} \frac{\partial F}{\partial q_{i,\ell}}(Q)\Delta q_{i,\ell} + R(Q, \Delta Q(t))$.

We then have

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

It only remains to prove that $\mathbb{E}[\ R(Q,\Delta Q(t))\ |Q(t)\ ] = 0$ when F is multiaffine.

A multiaffine function F is particular polynomial function, of degree 1 in each variable. By definition, $R(Q,\Delta Q)$ is hence also a polynomial function, of degree 1 in each variable $\Delta Q_{i,\ell}$. By construction, it has no-constant term, and no monomial of the form $\beta_{i,\ell}\Delta Q_{i,\ell}$. Hence, all its monomials are of the form $\beta_{i,\ell,j,\ell'}\Delta Q_{i,\ell}(t)\Delta Q_{j,\ell'}(t)$, with $(i,\ell) \neq (j,\ell')$.

By definition of multiaffine function used in this paper, there can not be terms $\Delta Q_{i,\ell}(t)\Delta Q_{j,\ell'}(t)$ with $i = j$ among these monomials.

Observe that $\Delta Q_{i,\ell}(t)\Delta Q_{j,\ell'}(t) = 0$ for $i \neq j$: indeed, at any time t, at most one player moves in the considered class of algorithms: in other words, we use the fact that considered algorithms are elementary stepwise.

□

When considering a Lyapunov game with respect to replicator-like dynamics, using Equation (1) and the fact that $G_i(Q) = \lim_{b\to 0} \tilde{F}_i^b(Q)$ the right hand side of Equation (8) is

$$b\sum_{i=1}^{n}\sum_{\ell=1}^{m_i} p_i \frac{\partial F}{\partial q_{i,\ell}}(Q)G_{i,\ell}(Q) + \mathcal{O}(b^2), \tag{9}$$

and hence expected to be negative by Equation (4) when $G(Q) \neq 0$ and b is sufficiently small.

In other words, when b is small, $(F(Q(t))_t$ will be a super-martingale until reaching a point where (9) is close to 0.

More precisely, for a replicator-like dynamics, Equation (9) rewrites to[12]

$$-b\frac{1}{4}\sum_i p_i \sum_{\ell\neq\ell'} q_{i,\ell}q_{i,\ell'}(u_i(e_\ell, Q_{-i}) - u_i(e_{\ell'}, Q_{-i}))^2 + \mathcal{O}(b^2).$$

As expected, on corners of K, this is expected to be close to 0, and hence not (neccesarily) a super-martingale.

[12] See Appendix 5.

For the perturbed replicator-like dynamics, taking the perturbation $\mathcal{O}(b)$ in page 3 to be 0, Equation (9) rewrites to

$$-b\alpha\frac{1}{4}\sum_i p_i \sum_{\ell\neq\ell'} q_{i,\ell}q_{i,\ell'}(u_i(e_\ell,Q_{-i})-u_i(e_{\ell'},Q_{-i}))^2 + b^2(1-\alpha)\sum_{i=1}^{n}\sum_{\ell=1}^{m_i}\frac{\partial F}{\partial q_{i,\ell}}(Q(t))(\frac{1}{m_i}-q_{i,\ell}).$$

which can be written

$$-b\alpha\frac{1}{4}\sum_i p_i \sum_{\ell\neq\ell'} q_{i,\ell}q_{i,\ell'}(u_i(e_\ell,Q_{-i})-u_i(e_{\ell'},Q_{-i}))^2 + \mathcal{O}(b^2).$$

When talking about stochastic perturbed dynamics, using this super-martingale argument, one gets the following stability result : we write $L(\mu)$ for the subset of states Q on which $F(Q) \leq \mu$.

Proposition 5 *Let $\lambda > 1$. Let $Q(t_0)$ be some state. Consider b enough small so that* (8) *is non-positive outside of $L(F(Q(0)))$. Then $Q(t)$ will be such that $Q(t) \in L(\lambda F(Q(t_0)))$ forever after time $t \geq t_0$ with a probability greater than $1-\frac{1}{\lambda}$.*

Proof: Consider sequence $Z_n = \max_{t\leq n} F(Q(t))$ and $\mathcal{F}_i$ the sigma-algebra generated by $(Q(j))_{j\leq i}$, and apply Proposition 8 for $\lambda' = \lambda\mathbb{E}[\ Z_0\ ]$:

$$P[\forall n, F(Q(n)) \leq \lambda F(Q(0)] = P[\sup_n Z_n \geq \lambda'] \leq \frac{\mathbb{E}[\ Z_0\ ]}{\lambda'} = \frac{1}{\lambda}.$$

□ If dynamic is perturbed, then the underlying Markov chain is ergodic. It follows that any neighborhood is visited with a positive probability: a dynamic will be said *perturbed* if for all $Q \in K$, for any neighborhood V with Q in its closure, the probability that $Q(t+1) \in V$ when $Q(t) = Q$ is positive.

Then if in some neighborhood of such a point we can apply previous proposition, one would get that almost surely, after some time, $Q(t)$ will be close to some Nash equilibria forever with high probability. The default of such an approach is clearly on the fact that it does not provide bounds on the time required to reach such a neighborhood.

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[13]: we write $L(\mu)$ for the subset of states Q on which $F(Q) \leq \mu$.

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

If one prefers, in an ϵ-Nash equilibrium, no player can improve its situation by more than ϵ times its current cost by changing unilaterally its strategy.

In a non ϵ-Nash equilibrium, we have some i and ℓ, with $u_i(e_\ell, Q_{-i}) < (1-\epsilon)u_i(q_i, Q_{-i})$. This means, $u_i(q_i - e_\ell, Q_{-i}) > \epsilon u_i(q_i, Q_{-i})$.

[13]Proof of Theorem 4 can be found in Appendix A.8.

For the perturbed replicator-like dynamics, taking the perturbation $\mathcal{O}(b)$ to be 0 in the definition of this dynamics, we have

$$\mathbb{E}[\ \Delta q_{i,\ell}\ |Q(t)\ ] = -\alpha b p_i q_{i,\ell}(u_i(e_\ell, Q_{-i}) - u_i(q_i, Q_{-i})) + b^2(1-\alpha)(\frac{1}{m_i} - q_{i,\ell}).$$

Assume without loss of generality that all costs are greater than 1. Let $\zeta = p_i u_i(q_i - e_\ell, Q_{-i})$ and $\beta = 1 - \alpha$. Previous equation is of the form $b(\alpha q_{i,\ell}\zeta + b\beta(\frac{1}{m_i} - q_{i,\ell}))$, hence some strictly increasing function of $q_{i,\ell}$ as soon as $b < \frac{\alpha}{\beta} p_i \epsilon u_i(q_i, Q_{-i})$ and $\zeta > p_i \epsilon u_i(q_i, Q_{-i})$. In that case, its minimal value, obtained for $q_{i,\ell} = 0$ is $\delta = \frac{b^2\beta}{m_i}$.

So, as soon as $\zeta > p_i \epsilon u_i(q_i, Q_{-i})$, that is to say $u_i(q_i - e_\ell, Q_{-i}) > \epsilon u_i(q_i, Q_{-i})$, we will have $\mathbb{E}[\ \Delta q_{i,\ell}\ |Q(t)\ ] \geq \delta$, that implies $\mathbb{E}[\ q_{i,\ell}(t+1)\ |Q(t)\ ] \geq \delta$.

This implies that the opposite of $\mathbb{E}[\ \Delta F(Q(t+1))\ |Q(t)\ ]$ will be greater than

$$V = b\alpha\frac{1}{4}p_i\delta\sum_{\ell\neq\ell'} q_{i,\ell'}(u_i(e_\ell, Q_{-i}) - u_i(e_{\ell'}, Q_{-i}))^2 + \mathcal{O}(b^2).$$

Taking $b < (1-\mu)\frac{\alpha}{\beta}p_i\epsilon u_i(q_i, Q_{-i})$ for any $\mu > 0$ guarantees that the factor in $q_{i,\ell}$ in previous discussed expression is greater than $\mu\zeta\alpha$, and hence that its iterations growth exponentially fast near 0. Reasoning by sequences of k steps, i.e. about the opposite of $\mathbb{E}[\ \Delta F(Q(t+k))\ |Q(t)\ ]$, will greater than a term of order

$$V = b\alpha\frac{1}{4}p_i\sum_{\ell\neq\ell'} q_{i,\ell'}(u_i(e_\ell, Q_{-i}) - u_i(e_{\ell'}, Q_{-i}))^2$$

in a non-ϵ-Nash equilibrium.

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 $b = \mathcal{O}(\epsilon)$, 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 $T(\epsilon)$ satisfies*

$$T(\epsilon) \leq \mathcal{O}(\frac{F(Q(0))}{\epsilon}).$$

We believe these bounds are tight for generic ordinal games. The point is that in arbitrary ordinal games, there is no necessarily relation between the gain in utility and the gain in potential: only sign of variation must be preserved.

Of course better bounds can be hoped for particular games, and in particular for congestion games. For generic congestion games, there is a strong relation between the potential and utilities of players. In congestion games, using notations from page 2, the potential is given by $F(Q) = \mathbb{E}[\ \sum_{r=1}^{m}\sum_{t=1}^{\lambda_r(t)} C_r(t)\ ]$. One has in particular $F(Q) \leq \mathbb{E}[\ \sum_{i=1}^{n} u_i(c_i, Q)\ ]$, since $c_i(Q) = \sum_{r\in q_i} C_r(\lambda_r(Q))$.

In particular, following [?], a congestion game is said to satisfy the *α-bounded jump condition* if its cost functions satisfy $C_r(t+1) \leq \alpha C_r(t)$ for all $t \geq 1$. This ensures the following property for $\delta = \frac{1}{\alpha n}$ (see [?]): whenever Q is not an ϵ-Nash equilibrium, then for at least a player i, the relative cost of adopting some pure strategy ℓ would induce a gain at least δ times the resulting gain in potential.

We believe perturbed replicator-like dynamics to converge very fast (hence in polynomially many steps) on such games.

A Proofs

A.1 Comments

A.1.1 Informal Analysis of the Dynamics of Stochastic Algorithms

Assume we replace $\mathbb{E}[\ \Delta q_i(t)\ |Q(t)\ ]$ by $\Delta q_i(t)$ in $\mathbb{E}[\ \Delta q_i(t)\ |Q(t)\ ] = bp_i\tilde{F}_i^b(Q(t))$, in the discussion that follows the description of the algorithm, where $\tilde{F}_i^b(Q(t)) = \mathbb{E}[\ F_i^b(r_i(t), s_i(t), q_i(t))\ |Q(t)\ ]$.

Through the change of variable $t \leftarrow tb$, this would become $q_i(t+b) - q_i(t) = bp_i\tilde{F}_i^b(Q)$. Approximating $q_i(t+b) - q_i(t)$ by $b\frac{dq_i}{dt}(t)$ for small b, we may expect the system to behave like ordinary differential equation (ODE)

$$\frac{dq_i}{dt} = p_i G_i(Q), \tag{10}$$

when b is close to 0.

A.1.2 Turning Replicator Dynamics Into a Non-Complacency Dynamics

Following the discussion after Proposition 6, to avoid 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 [?] some patches on the dynamics that would guarantee Non-complacency. *Non-Complacency (NC)* is the following property: $G(Q) = 0$ (stationary) implies that Q is a Nash equilibrium (3).

Indeed, one we may consider for example the following class of dynamics: Let $\delta(.) : K \to [0,1]$ be some continuous function. A *patched replicator* dynamics corresponds to a dynamics of form

$$F_i^b(r_i(t), s_i(t), q_i(t)) = \begin{cases} \frac{M - r_i(t)}{M}(e_{s_i(t)} - q_i(t)) & \text{with probability } 1 - \delta(q_i) \\ e_{s_j} - q_i(t) & \text{with probability } \delta(q_i), \\ & \quad \text{where } j \in \{1, \dots, m_i\} \text{ is chosen uniformly,} \end{cases}$$

Dynamics (2) becomes

$$\frac{dq_{i,\ell}}{dt} = -p_i \frac{1 - \delta(q_i)}{M} q_{i,\ell}(c_i(e_\ell, Q_{-i}) - u_i(q_i, Q_{-i})) + p_i\delta(q_i)(\frac{1}{m_i} - q_{i,\ell}) \tag{11}$$

whose stationary points are now exactly Nash equilibria, if $\delta(.)$ is well-chosen. Indeed, follow the idea of the construction in Appendix A.3 of [?]. Roughly, take δ to be 0 everywhere except on neighborhood of non-Nash stationary points of dynamic (3). On such a neighborhood, define it as positive so that to guarantee that right-hand side of Equation (11) stay positive for the i and ℓ for which $u_i(e_\ell, Q_{-i}) - u_i(q_i, Q_{-i})$ is not 0, that must exist in a non-Nash equilibrium.

When dealing with stochastic dynamics, all this can be avoided, by exploiting unstability of non-Nash stationary equilibrium points.

A.1.3 On Weak Convergence in Theorems 1 and 2

Using techniques from [?], this is possible to reinforce weak convergence in Theorems 1 and 2 into stronger notions of convergence, over a finite horizon, as in [?] where mean square (and thus in probability) convergence results are obtained, under wide hypotheses. Several results relating the stochastic dynamics and its ODE approximation, including results about their asymptotic behavior can also be established following constructions from [?].

Notice, that this would not help to get convergence of the underlying mean-field limit, nor help with the double limit $b \to 0$, $t \to \infty$.

A.2 A General Theorem about Approximation of Diffusions

We will use the following theorem from [**?**, theorem 11.2.3]. The following presentation is inspired by the presentation of it in [**?**, Theorem 5.8, page 96].

Suppose that for all integers $b > 0$, we have an homogeneous Markov chain $(Y_k^{(b)})$ in $\mathbb{R}^d$ with transition kernel $\pi^{(b)}(x, dy)$, meaning that the law of $Y_{k+1}^{(b)}$, conditioned on $Y_0^{(b)}, \cdots, Y_k^{(b)}$, depends only on $Y_k^{(b)}$ and is given, for all Borelian B, by $P(Y_{k+1}^{(b)} \in B | Y_k^{(b)}) = \pi^{(b)}(Y_k^{(b)}, B)$, almost surely.

Define for $x \in \mathbb{R}^d$,

$$\begin{aligned}
d^{(b)}(x) &= \frac{1}{b} \int (y - x)\pi^{(b)}(x, dy), \\
a^{(b)}(x) &= \frac{1}{b} \int (y - x)(y - x)^* \pi^{(b)}(x, dy), \\
K^{(b)}(x) &= \frac{1}{b} \int (y - x)^3 \pi^{(b)}(x, dy), \\
\Delta_\epsilon^{(b)}(x) &= \frac{1}{b} \pi^{(b)}(x, B(x, \epsilon)^c),
\end{aligned}$$

where $B(x, \epsilon)^c$ denotes the complement of the ball with radius ϵ, centered at x.

The coefficients $d^{(b)}$ and $a^{(b)}$ can be interpreted as the instantaneous drift and the variance (or matrix of covariance) of $X^{(b)}$.

Define

$$X^{(b)}(t) = Y_{\lfloor t/b \rfloor}^{(b)} + (t/b - \lfloor t/b \rfloor)(Y_{\lfloor t/b+1 \rfloor}^{(b)} - Y_{\lfloor t/b \rfloor}^{(b)}).$$

Theorem 5 ([?, theorem 11.2.3], [?, Theorem 5.8, page 96]) *Suppose that there exist some continuous functions d, b, such that for all $R < +\infty$,*

$$\begin{aligned}
\lim_{b \to 0} sup_{|x| \le R} |a^{(b)}(x) - a(x)| = 0 \\
\lim_{b \to 0} sup_{|x| \le R} |d^{(b)}(x) - d(x)| = 0 \\
\lim_{b \to 0} sup_{|x| \le R} \Delta_\epsilon^{(b)} = 0, \forall \epsilon > 0 \\
\sup_{|x| \le R} K^{(b)}(x) < \infty.
\end{aligned}$$

With σ a matrix such that $\sigma(x)\sigma^*(x) = a(x)$, $x \in \mathbb{R}^d$, we suppose that the stochastic differential equation

$$dX(t) = d(X(t))dt + \sigma(X(t))dB(t), \qquad X(0) = x, \tag{12}$$

has a unique weak solution for all x. This is in particular the case, if it admits a unique strong solution.

Then for all sequences of initial conditions $Y_0^{(b)} \to x$, the sequence of random processes $X^{(b)}$ weakly converges *to the diffusion given by Equation* (12). *In other words, for all functions F : $\mathcal{C}(\mathbb{R}^+, \mathbb{R}) \to \mathbb{R}$ bounded and continuous, one has*

$$\lim_{b \to 0} E[F(X^{(b)})] = E[F(X)].$$

A.3 Proof of Theorem 1

Theorem 1 follows from previous theorem. Consider $(Y_k^{(b)})$ to be

$$Y_k^{(b)} = (Q(k))$$

with the corresponding b, which is indeed an homogeneous Markov chain. Let $\pi^{(b)}(Q, dy)$ be its transition kernel.

We have

$$\begin{array}{rcl} d_i^{(b)}(Q) & = & \frac{1}{b}\int (y_i - q_i)\pi^{(b)}(x, dy) \\ & = & \frac{1}{b}\mathbb{E}[\ \Delta q_i\ |Q\ ] \\ & = & \frac{p_i}{b} b\tilde{F}_i^b(Q) \\ & \rightarrow & p_i G_i(Q) \qquad \text{when } b \rightarrow 0 \end{array}$$

and

$$\begin{array}{rcl} a_{i,j}^{(b)}(Q) & = & \frac{1}{b}\int (y_i - q_i)(y_j - q_j)^* \pi^{(n)}(x, dy) \\ & = & \frac{b^2}{b}\mathbb{E}[\ p_i p_j \Delta q_i \Delta q_j\ |Q\ ] \\ & = & \mathcal{O}(b) \\ & \rightarrow & 0 \text{ when } b \rightarrow 0 \end{array}$$

In the same vein, clearly $K^{(b)}(Q)$ stay bounded, being in $\mathcal{O}(b^2)$.

Now, from the fact that compact K must be kept invariant by the dynamics, $F_i^b(.)$ must have a compact support. This means that $\pi^{(b)}(Q, B(Q, \epsilon)^c)$ is 0 for b sufficiently small. Hence $\lim_{b\rightarrow 0} sup_{|x|\leq R}\Delta_\epsilon^{(b)} = 0$, $\forall \epsilon > 0$.

Hence, we have all the hypotheses of previous theorem with $a(Q) = 0$ and

$$d(Q) = (p_1 G_1(Q), \cdots, p_n G_n(Q)),$$

observing that the corresponding stochastic differential equation $dQ(t) = d(Q(t))dt + \sigma(Q(t))dB(t)$ turns out to be an ordinary differential equation, whose solution is unique by (classical) Cauchy Lipschtiz theorem.

A.4 Derivation of Dynamics (3) For Replicator-Like Dynamics

For replicator-like dynamics set $\alpha = 1$ in what follows.

For replicator-like dynamics and perturbed replicator-like dynamics, the one-step dynamics of the stochastic algorithm can be rewritten componentwise:

$$\Delta q_{i,\ell}(t) = q_{i,\ell}(t+1) - q_{i,\ell}(t) = \alpha \begin{cases} 0 & +\mathcal{O}(b) \quad \text{if} \quad i \neq i(t) \\ -b\gamma(r_i(t))q_{i,\ell}(t) & +\mathcal{O}(b) \quad \text{if} \quad i = i(t) \text{ and } s_i(t) \neq l \\ -b\gamma(r_i(t))q_{i,\ell}(t) + b(\gamma(r_i(t))) & +\mathcal{O}(b) \quad \text{if} \quad i = i(t) \text{ and } s_i(t) = l, \end{cases}$$

and we have

$$\begin{array}{rcl} G_i(Q) & = & \lim_{b\rightarrow 0} \frac{1}{bp_i}\mathbb{E}[\ \Delta q_{i,\ell}(t)\ |Q(t)\ ] \\ & = & \lim_{b\rightarrow 0} \frac{1}{b}\sum_j q_{i,j}(t)\mathbb{E}[\ \Delta q_{i,\ell}(t)\ |Q(t), s_i(t) = j, i(t) = i\ ] \\ & = & +\alpha\sum_j q_{i,j}(t) q_{i,\ell}(t)\mathbb{E}[\ \gamma(r_i(t))\ |Q(t), s_i(t) = \ell, i(t) = i\ ]) \\ & & -\alpha\sum_j q_{i,j}(t)(q_{i,\ell}(t)\mathbb{E}[\ \gamma(r_i(t))\ |Q(t), s_i(t) = j, i(t) = i\ ]) \\ & = & q_{i,\ell}(\mathbb{E}[\ \gamma(r_i(t))\ |Q(t), s_i(t) = \ell, i(t) = i\ ] - \mathbb{E}[\ \gamma(r_i(t))\ |Q(t), i(t) = i\ ]). \end{array}$$

that is to say, if we introduce $u_i(Q) = \mathbb{E}[\ -\frac{1}{\alpha}\gamma(r_i(Q))\ |Q\ ]$ for all Q, then Equation (2) leads to dynamics,

$$\frac{dq_{i,\ell}}{dt} = -p_i q_{i,\ell}(u_i(e_\ell, Q_{-i}) - u_i(q_i, Q_{-i})).$$

by Theorem 1. This is Equation exactly (3).

A.5 Formal Statement about Theorem 2:

Formally, we have:

Proposition 6 *The following are true for the solutions of Equation* (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.*

The following are well-known (and obtained by just playing with definitions).

Lemma 2 *A strategy profile Q is a Nash Equilibrium iff $u_i(q_i, Q_{-i}) \leq u_i(e_\ell, Q_{-i})$ for all $1 \leq i \leq n$, $1 \leq \ell \leq m_i$.*

Corollary 2 *In a Nash Equilibrium, we have $u_i(q_i, Q_{-i}) = u_i(e_\ell, Q_{-i})$ for all $1 \leq i \leq n$, $1 \leq \ell \leq m_i$ with $q_{i,\ell} > 0$.*

Proposition 6 is then an instance of the so-called folk-theorems of Evolutionary Game Theory [**?**]. For completeness, the proof goes as follows: From Corollary 2, clearly any Nash equilibria must also vanish the right-hand side of Equation (3).

A non-Nash equilibrium Q is not stable: Indeed, if Q is not a Nash equilibrium, this means that for some i, and some ℓ we have $u_i(q_i, Q_{-i}) > u_i(e_\ell, Q_{-i})$. By bilinearity and continuity of u_i, function $u_i(q_i - e_\ell, Q_{-i})$ must be strictly positive (say greater than ϵ) on some neighborhood of Q. On this neighborhood, $\frac{dq_{i,\ell}}{dt}$ is greater than $p_i q_{i,\ell}\epsilon$, and hence the point is left exponentially faster (faster than exponential $q_{i,\ell}(0)\exp(p_i \epsilon t)$).

In a corner of K, we have for all i, $q_i = e_\ell$ for some ℓ. Then clearly $q_{i,\ell'} = 0$ for index $\ell' \neq \ell$, and $u_i(e_\ell, Q_{-i}) - u_i(q_i, Q_{-i}) = 0$ for index $\ell' = \ell'$. Hence, the right-hand side of Equation (3) is always null, and hence any corner is a stationary point.

More generally any state Q in which all strategies in its support perform equally well, is clearly a stationary point from the definition of the dynamic.

A.6 Proof of Theorem 2

Theorem 2 follows immediately. Observe that a limit point must necessarily be a stationary point [**?**], and hence a Nash equilibrium if it is stable.

A.7 Results About Semi-Martingales

Let $\{Z_i, i \geq 0\}$ be a sequence of real non-negative random variables, such that Z_i is measurable in the increasing family of sigma-algebra $\mathcal{F}_i$.

Proposition 7 (proof similar to [?, Theorem 2.1.1, page 17]) *Assume that Z_0 is constant. Denote by τ the $\mathcal{F}_n$-stopping time representing the epoch of the first entry into $[0, C]$ or in some measurable subset K, for $C > 0$, i.e. $\tau(\omega) = \inf\{n \geq 1 | Z_n(\omega) \leq C \vee Z_n(\omega) \in K\}$. Introduce the stopped sequence*

$$\tilde{Z}_n = Z_{n \wedge \tau},$$

where

$$n \wedge \tau = \begin{cases} n, & \text{if } n \leq \tau \\ \tau, & \text{if } n > \tau \end{cases}$$

We use the classical notation for the indicator function $1_{\mathcal{A}}$:

$$1_{\mathcal{A}} = \begin{cases} 1, & \text{if } \mathcal{A} \text{ is true} \\ 0, & \text{otherwise} \end{cases}$$

Assume $Z_0 > C$, and for some $\epsilon > 0$ and all $n \geq 0$,

$$\mathbb{E}[\ \tilde{Z}_{n+1}\ |\mathcal{F}_n\ ] \leq \tilde{Z}_n - \epsilon 1_{\tau > n}, \text{ almost surely.}$$

Then τ is almost surely finite and

$$\mathbb{E}[\ \tau\ ] < \frac{Z_0}{\epsilon} < \infty.$$

Proposition 8 ([?, Theorem 3.2, Chapter 7]) *Assume that for all n, $\mathbb{E}[\ Z_{n+1} - Z_n\ |\mathcal{F}_n\ ] \leq 0$. Then for all $\lambda' > 0$,*

$$P[\sup_n Z_n \geq \lambda'] \leq \frac{\mathbb{E}[\ Z_0\ ]}{\lambda'}.$$

A.8 Proof of Theorem 4

Consider $V^* = \min_{i,q_i} \sum_{\ell \neq \ell'} q_{i,\ell'} (u_i(e_\ell, Q_{-i}) - u_i(e_{\ell'}, Q_{-i}))^2$. Let $I(\epsilon)$ denote the states where the righthand side of Equation (8) is greater than $-b\alpha \frac{1}{4} \min_i p_i V^* \epsilon$.

If the initial state is already ϵ-stable then there is nothing to prove.

Otherwise, this follows from the analysis before Theorem 4, and from proposition 7, with $Z_i = F(Q(i))$, $\mathcal{F}_i$ the sigma-algebra generated by $(Q(j))_{j \leq i}$, $C = \mu$, $K = I(\epsilon)$: indeed, whenever $Q(t) \notin I(\epsilon) \cup L(\mu)$, this implies $\tau > t$, and we have $\mathbb{E}[\ \Delta F(t)\ |Q(t)\ ] = \mathbb{E}[\ Z_{t+1} - Z_t\ |\mathcal{F}_t\ ] \leq -\epsilon \mathcal{O}(b)$. In all other cases, $\mathbb{E}[\ \tilde{Z}_{n+1}\ |\mathcal{F}_n\ ] = \tilde{Z}_n$ and hence all the hypotheses of Proposition 7 are satisfied.