Selection of Efficient Pure Strategies in Allocation Games

Pierre Coucheney, Corinne Touati, and Bruno Gaujal
INRIA Rhône-Alpes, Grenoble, France
{pierre.coucheney, corinne.touati, bruno.gaujal}@inria.fr

Abstract— **In this work we consider allocation games and we investigate the following question: under what conditions does the replicator dynamics select a pure strategy?**

By definition, an allocation game is a game such that the payoff of a player when she takes an action only depends on the set of players who also take the same action. Such a game can be seen as a set of users who share a set of resources, a choice being an allocation to a resource.

A companion game (with modified utilities) is introduced. From the payoffs of an allocation game, we define the repercussion utilities: for each player, her repercussion utility is her payoff minus the decrease in marginal payoff that her presence causes to all other players. The corresponding allocation game with repercussion utilities is the game whose payoffs are the repercussion utilities. A simple characterization of those games is given.

In such games, if the players select their strategy according to a stochastic approximation of the replicator dynamics, we show that it converges to a Nash equilibrium of the game that is a locally optimal for the initial game. The proof is based on the construction of a potential function for the game. Furthermore, a spectral study of the dynamics shows that no mixed equilibrium is stable, so that the strategies of all players converge to a set of Nash equilibria. Then, martingale argument prove the convergence of the stochastic approximation to a pure point. A discussion of the global/local optimality of the limit points is also included.

Keywords: **allocation games; replicator dynamics; stochastic approximation; pure Nash equilibrium.**

I. Introduction

Deterministic evolutionary dynamics for games first appeared in the mathematical biology literature. Soon after, it has found many applications in non-biological fields like economics or learning theory in what is now called evolutionary game theory, and classical texts emerged [1], [2]. In biology, the replicator dynamics was first introduced by Taylor and Jonker as an explicit dynamics that models the evolution of species. In economics, the replicator dynamics is a way to describe the behavior of a large population of agents who are randomly matched to play normal form games. Nowadays, the evolutionary game theory has been successfully applied to allocation and routing problems in telecommunication systems [3], [4].

There is a connection between the Nash equilibria and replicator dynamics of a game. Indeed, according to the folk theorem of evolutionary games, Nash equilibria are stationary points of the dynamics and all stable stationary points are Nash equilibria. A central question for normal form game with multiple Nash equilibria is to know which of them will be selected. In some cases, replicator dynamics converges to a Nash equilibrium, but, unfortunately, it is known that for the general "rock-scissors-paper" game, the replicator dynamics may converge to a cycle and therefore does not select any strategy. Hence, in general setting, the replicator dynamics does not help one to select an equilibrium.

In this paper, we study mechanisms that select efficient Nash Equilibria for *allocation games*. Allocation games are a generalization of congestion games [5]. The mechanism we study here can be viewed as a stochastic approximation of the replicator dynamics.

Some works have shown that asymmetric games with two players admit no *interior* equilibrium stable for the replicator dynamics [6], [7]. Purity of the equilibrium point is an interesting property for many applications. Indeed, mixed strategies may be costly for players as they amount to a permanent change of choices [8]. Here, we provide a quite general framework under which allocation games with an arbitrary number of players and an arbitrary number of choices converge to pure strategies. Furthermore, the limit equilibrium point is also proved locally optimal.

Things work as follow: starting with an allocation game (introduced in Section II-A), we define a companion game by introducing the repercussion utilities (defined in Section II-B). For each player, the repercussion utility is defined to be her payoff in the original game minus the decrease in marginal payoff that her presence causes to all other players. We then give a useful characterization of those games.

Second, we study the replicator dynamics for the game with repercussion utilities (Section III-A) and we show that it admits a potential. Therefore, allocation games with repercussion utilities, for which we have a simple characterization, are a subclass of potential games [9], [10]. A classical result claims that the potential is a Lyapunov function for the replicator dynamics. We also show, by considering the divergence of the replicator dynamics, that it does not admit any stable equilibria in the interior of its domain (the cube in dimension N). Combining these two ingredients, we show that the stable points of the replicator dynamics for the game with repercussion utilities are faces of the domain. This induces the existence of a pure Nash equilibrium in allocation games with repercussion utilities. Finally we propose in Section III-B a distributed algorithm based on a stochastic approximation of the replicator dynamics that converges to a stable equilibrium of the replicator dynamics, and, furthermore, to a pure Nash equilibrium of the allocation

This work was performed within the common INRIA - Alcatel-Lucent Laboratory.

game with repercussion utilities (see Theorem 2). As the algorithm is stochastic, at each iteration and for each player, only *one* possible action is tested and evaluated. This limits the number of changes of actions, which can be attractive in certain types of applications [8] and contrasts with classical gradient-descent like algorithms [4].

The last part (Section IV) introduces a discussion on the shape of the basin of attraction of the stable equilibria in the case where the potential function admits several local maxima. We show that with two players and two choices, the center of the cube always belongs to the basin of attraction of the globally maximal point. This insures that the stochastic approximation algorithm converges to the optimal equilibrium with a large probability when starting from the center of the cube. However, with more players and/or more choices, we exhibit several examples showing that starting in the center may not lead to the globally optimal point.

II. Allocation Games Related to Potential Games

A. Allocation Games

We consider a normal-form game $(\mathcal{N},\mathcal{I},\mathcal{U})$ consisting of a set $\mathcal{N}$ of players ($|\mathcal{N}|=N$), player n taking actions in a set $\mathcal{I}_n \subset \mathcal{S}$ ($|\mathcal{I}_n| = I_n$), where $\mathcal{S}$ is the set of all actions. Let us denote by $s_n \in \mathcal{I}_n$ the action taken by player n, and $s=(s_n)_{n\in\mathcal{N}} \in \mathcal{I} = \bigotimes_{n=1}^{N} \mathcal{I}_n$. Then, $\mathcal{U}=(U_n)_{n\in\mathcal{N}}$ refers to the *utility* or *payoff* for each player: the payoff for player n is $U_n(s_1,\ldots,s_n,\ldots,s_N)$.

By definition, an *allocation game* is a game such that the payoff of a player when she takes action i only depends on the set of players who also take action i. One can interpret such a game as a set of users who share a common set of resources $\mathcal{S}$, and an action vector corresponds to an allocation of resources to users (hence the name of these games).

We define the *load* on action (or resource) i by $\ell^i(s) \in \{0;1\}^N$ as a vector such that $\ell^i_n(s)=1$ if and only if player n take action i. When there is no ambiguity, we will simplify the notation and use $\ell = \ell^i(s)$. We denote by $\ell^{s_n}(s)$ the load on the action taken by player n, and we denote the payoff for player n by $u_n(\ell^{s_n}(s)) \stackrel{\text{def}}{=} U_n(s_1,\ldots,s_n,\ldots,s_N)$.

Hence, allocations games are a wider class of games than *congestion games* where the payoff of each player depends on the *number* of players adopting the same strategy [5]. They represent systems where different users accessing a given resource may have a different impact.

B. Repercussion utilities

We build a companion game of the allocation game, denoted $(\mathcal{N},\mathcal{I},\mathcal{R})$. The new player utilities, called *repercussion utilities* are built from the payoffs of the original game, according to the following definition:

Definition 1 (allocation game with repercussion utilities): Let us consider the repercussion utility for player n to be:

$$r_n(\ell^{s_n}(s)) \stackrel{\text{def}}{=} u_n(\ell^{s_n}(s)) - \sum_{m\neq n: s_m=s_n} \left(u_m(\ell^{s_m}(s)-e_n) - u_m(\ell^{s_m}(s))\right).$$

where e_n denotes the vector whose entries are all 0 but the nth one, which equals 1.
An *allocation game with repercussion utilities* is a game whose payoffs are repercussion utilities.

The utilities defined in this manner have a natural interpretation: it corresponds to the player's payoff $(u_n(\ell^{s_n}(s)))$ minus the total increase in payoff for all users impacted by the presence of a given user on a given commodity ($\sum_{\substack{m\neq n:\\ s_m=s_n}} [u_m(\ell^{s_m}(s)-e_n) - u_m(\ell^{s_m}(s))]$). This is more obvious in the following equivalent formulation.

Remark 1: An equivalent formulation of the repercussion utilities is:

$$r_n(\ell^{s_n}(s)) = \sum_{m:\ell^{s_n}_m=1} u_m(\ell^{s_n}(s)) - \sum_{\substack{m\neq n:\\ \ell^{s_n}_m(s)=1}} u_m(\ell^{s_n}(s)-e_n).$$

C. Characterization of Allocation Games with Repercussion Utilities

We now give a characterization of a payoff which is a repercussion utility.

Proposition 1: An allocation game $(\mathcal{N},\mathcal{I},\mathcal{R})$ is an allocation game with repercussion utilities if and only if $\forall \ell, \forall n,m \in \mathcal{N}$ s.t. $s_m=s_n$,

$$r_n(\ell) - r_n(\ell-e_m) = r_m(\ell) - r_m(\ell-e_n). \qquad (1)$$

Proof: Suppose that r is a repercussion utility, then there exists a payoff u such that:

$$r_n(\ell) = \sum_{\ell_k=1} u_k(\ell) - \sum_{k\neq n:\ell_k=1} (u_k(\ell-e_n)).$$

Then, denote

$$A = \left(\sum_{\ell_k=1} u_k(\ell-e_m) - \sum_{k\neq n:\ell_k=1} u_k(\ell-e_n-e_m)\right).$$

Then,

$$\begin{aligned} r_n(\ell) - r_n(\ell-e_m) &= \sum_{\ell_k=1} u_k(\ell) - \sum_{k\neq n:\ell_k=1} u_k(\ell-e_n) - A \\ &= \sum_{\ell_k=1} u_k(\ell) - \sum_{k\neq m:\ell_k=1} u_k(\ell-e_m) - A \\ &= r_m(\ell) - r_m(\ell-e_n). \end{aligned}$$

Conversely, consider an allocation game $(\mathcal{N},\mathcal{I},\mathcal{R})$ such that Eq. 1 is satisfied. Consider an action i and ℓ the vector of load on action i. Then, $\ell \in \{0;1\}^N$ and $K \stackrel{\text{def}}{=} \sum_{k:\ell_k=1} 1$ is the number of players taking action i. Further, let $(a(k)), 1 \leq k \leq K$ be the subscripts of all players taking action i. If there are K such players, then $\ell = \sum_{k=1}^{K} e_{a(k)}$. Then, we claim that, for any permutation σ of $\{1, K\}$:

$$\sum_{k=0}^{K-1} r_{a(k+1)}(\ell - \sum_{j=1}^{k} e_{a(j)}) = \sum_{k=0}^{K-1} r_{a(\sigma(k+1))}(\ell - \sum_{j=1}^{k} e_{a(\sigma(j))}). \quad (2)$$

Indeed, note that, from Eq. 1:

$$r_{a(k+1)}\left(\ell - \sum_{j=1}^{k-1} e_{a(j)}\right) - r_{a(k+1)}\left(\ell - \sum_{j=1}^{k} e_{a(j)}\right) = r_{a(k)}\left(\ell - \sum_{j=1}^{k-1} e_{a(j)}\right) - r_{a(k)}\left(\ell - \sum_{j=1}^{k-1} e_{a(j)} - e_{a(k+1)}\right).$$

Therefore:

$$r_{a(k)}\left(\ell - \sum_{j=1}^{k-1} e_{a(j)}\right) + r_{a(k+1)}\left(\ell - \sum_{j=1}^{k} e_{a(j)}\right) = r_{a(k+1)}\left(\ell - \sum_{j=1}^{k-1} e_{a(j)}\right) + r_{a(k)}\left(\ell - \sum_{j=1}^{k-1} e_{a(j)} - e_{a(k+1)}\right).$$

Hence, for any k, the sum $\sum r_{a(k+1)}(\ell - \sum_{j=1}^{k} e_{a(j)})$ remains unchanged if one swaps $a(k)$ and $a(k+1)$ (elementary transposition). Then, Eq. 2 results from the fact that any permutation σ can be decomposed in a finite number of elementary transpositions.

We now construct a payoff u as follow: for any n such that $\ell_n = 1$, let us define:

$$u_n(\ell) \stackrel{\text{def}}{=} \frac{1}{K} \sum_{k=0}^{K-1} r_{a(k+1)}(\ell - \sum_{j=1}^{k} e_{a(j)}).$$

Then,

$$\sum_{\ell_m=1} u_m(\ell) - \sum_{m \neq n : \ell_m=1} (u_m(\ell - e_n)) = \sum_{k=0}^{K-1} r_{a(k+1)}(\ell - \sum_{j=1}^{k} e_{a(j)}) - \sum_{k=0}^{K-2} r_{b(k+1)}(\ell - e_n - \sum_{j=1}^{k} e_{b(j)}).$$

Note that the sequence a is identical to sequence b with the additional element n. From Eq. 2, we can choose a permutation σ such that $a(\sigma(1)) = n$. Then:

$$\sum_{\ell_m=1} u_m(\ell) - \sum_{m \neq n : \ell_m=1} (u_m(\ell - e_n))$$
$$= \sum_{k=0}^{K-1} r_{a(\sigma(k+1))}(\ell - \sum_{j=1}^{k} e_{a(\sigma(j))}) - \sum_{k=1}^{K-1} r_{a(\sigma((k+1))}(\ell - e_n - \sum_{j=2}^{k} e_{a(\sigma(j))})$$
$$= \sum_{k=1}^{K-1} r_{a(\sigma(k+1))}(\ell - e_{a(\sigma(1))} - \sum_{j=2}^{k} e_{a(\sigma(j))}) + r_{a(\sigma(1))}(\ell) - \sum_{k=1}^{K-1} r_{a(\sigma((k+1))}(\ell - e_n - \sum_{j=2}^{k} e_{a(\sigma(j))})$$
$$= r_n(\ell).$$

Hence $(\mathcal{N}, \mathcal{I}, \mathcal{R})$ is the allocation game with repercussion utilities associated to the $(\mathcal{N}, \mathcal{I}, \mathcal{U})$ allocation game. ∎

From Prop. 1, we conclude that allocation games with repercussion utilities are a special subset of allocation games. The results presented in the following are hence valid for any allocation game such that Eq. 1 is satisfied.

Example 1: Let M be the payoff matrix of a two-player game. This amounts to saying that the first (resp. second) player chooses the line and the second chooses the column. The payoff for the first player is given by the first (resp. second) component.

$$M = \begin{pmatrix} (a, A) & (b, B) \\ (c, C) & (d, D) \end{pmatrix}.$$

It follows from Proposition 1 that this is a game with repercussion utilities if and only if $a = A + b - C$ and $d = D + c - B$. Then, one can check the interesting property that there necessarily exists a pure Nash equilibrium (for instance (a, A) is a Nash equilibrium if $a \geq c$ and $A \geq B$).

D. Allocation Games with Repercussion Utilities are Potential Games

In this section, we show that, given an allocation game, the game with repercussion utilities (1) admits a potential function and (2) this potential equals the sum of the payoffs for all players in the initial game. This appealing property is exploited in the next section to show some strong results on the behavior of the well-known replicator dynamics on such games.

Consider an allocation $(\mathcal{N}, \mathcal{I}, \mathcal{U})$ and its companion game $(\mathcal{N}, \mathcal{I}, \mathcal{R})$. We first assume that players have *mixed* strategies. Hence a strategy for player n is a vector of probability $q_n = (q_{n,i})_{i \in \mathcal{I}_n}$, where $q_{n,i}$ is the probability for player n to take action i (i.e. $q_{n,i} = \mathbb{P}(s_n = i) \geq 0$ and $\sum_{i \in \mathcal{I}_n} q_{n,i} = 1$). The strategy domain for player n is $\Delta_n \stackrel{\text{def}}{=} \{0 \leq q_{n,i} \leq 1, s.t. \sum_{i \in \mathcal{I}_n} q_{n,i} = 1\}$ Then, the global domain is $\Delta = \prod_{n=1}^{N} \Delta_n$ and a global strategy is $q \stackrel{\text{def}}{=} (q_n)_{n \in \mathcal{N}}$. We say that q is a *pure strategy* if for any n and i, $q_{n,i}$ equals either 0 or 1.

We denote by $S \in \mathcal{I}$ the random vector whose entries S_n are all independent and whose distribution is $\forall n \in \mathcal{N}, \forall i \in \mathcal{I}_n$, $\mathbb{P}(S_n = i) = q_{n,i}$. The expected payoff for player n

when she takes action i is $f_{n,i}(q) \stackrel{\text{def}}{=} \mathbb{E}[r_n(\ell^i(S))|S_n = i]$. Then, her mean payoff is $\overline{f}_n(q) \stackrel{\text{def}}{=} \sum_{i\in\mathcal{I}_n} q_{n,i} f_{n,i}(q)$. We can notice that $f_{n,i}(q)$ only depends on $(q_{m,i})_{m\neq n}$ and it is a multi-linear function of $(q_{m,i})_{m\neq n}$.

The next theorem claims that the allocation game with repercussion utilities is a potential game. Potential games were first introduced in [9]. The notion was afterward extended to continuous set of players [10]. In our case, it refers to the fact that the expected payoff for each player derives from a potential function. More precisely, here we show that $f_{n,i}(q) = \frac{\partial F}{\partial q_{n,i}}(q)$, where

$$F(q) \stackrel{\text{def}}{=} \sum_{n\in\mathcal{N}} \sum_{i\in\mathcal{I}_n} q_{n,i}\mathbb{E}[u_n(\ell^i(S))|S_n = i]. \tag{3}$$

It is interesting to notice the connection between $f_{n,i}(q)$ which is the expected repercussion utility, and $F(q)$ which refers to the sum of expected payoffs in the initial game: if a strategy increases the expected repercussion utility of a player, then it increases the potential.

Theorem 1: The allocation game with repercussion utilities is a potential game, and its associated potential function is F, as defined in Eq. 3.

Proof: Let us first differentiate function F:

$$\frac{\partial F}{\partial q_{n,i}}(q) = \mathbb{E}[u_n(\ell^i(S))|S_n = i] + \sum_{m\neq n} q_{m,i} \frac{\partial \mathbb{E}[u_m(\ell^i(S))|S_m = i]}{\partial q_{n,i}}.$$

In fact, it is clear that $\frac{\partial \mathbb{E}[u_m(\ell^j(S))|S_m = j]}{\partial q_{n,i}} = 0$ if $j \neq i$, and $\frac{\partial \mathbb{E}[u_n(\ell^i(S))|S_n = i]}{\partial q_{n,i}} = 0$. To simplify the notations, we omit the index i. Then,

$$\frac{\partial F}{\partial q_n}(q)$$

$$= \mathbb{E}[u_n(\ell(S))|S_n = i] + \sum_{m\neq n} q_m \frac{\partial}{\partial q_n} \sum_{\ell} u_m(\ell)\mathbb{P}(\ell(S) = \ell|S_m = i)$$

$$= \mathbb{E}[u_n(\ell(S))|S_n = i] + \sum_{m\neq n} q_m \sum_{\ell} u_m(\ell)\Big(\mathbb{P}(\ell(S) = \ell|S_m = i, S_n = i) - \mathbb{P}(\ell(S) = \ell|S_m = i, S_n \neq i)\Big)$$

$$= \mathbb{E}[u_n(\ell(S))|S_n = i] + \sum_{m\neq n} q_m \sum_{\ell} u_m(\ell) \Big(\mathbb{P}(\ell(S) = \ell, S_m = i|S_n = i) - \mathbb{P}(\ell(S) = \ell + e_n, S_m = i|S_n = i)\Big)$$

$$= \mathbb{E}[u_n(\ell(S))|S_n = i] - \sum_{\substack{m\neq n:\\ S_m=S_n}} \Big(\mathbb{E}[u_m(\ell(S) - e_n)|S_n = i] - \mathbb{E}[u_m(\ell(S))|S_n = i]\Big)$$

$$= \mathbb{E}[r_n(\ell(S))|S_n = i]$$
$$= f_{n,i}(q).$$ ■

III. Replicator dynamics and algorithms

In this section, we show how to design a strategy update mechanism for all players in an allocation game with repercussion utilities that converges to pure Nash Equilibria. We will study in the next section (Section IV) their efficiency properties.

A. Replicator Dynamics.

We now consider that the player strategies vary over time, hence q depends on the time t: $q = q(t)$. The trajectories of the strategies are described below by a dynamics called *replicator dynamics*. We will see in section III-B that this dynamics can be seen as the limit of a learning mechanism.

Definition 2: The replicator dynamics [1][2] is ($\forall n \in \mathcal{N}, i \in \mathcal{I}_n$):

$$\frac{dq_{n,i}}{dt}(q) = q_{n,i}\left(f_{n,i}(q) - \overline{f_n}(q)\right). \tag{4}$$

We say that $\hat{q}$ is a *stationary point* (or equilibrium point) if ($\forall n \in \mathcal{N}, i \in \mathcal{I}_n$):

$$\frac{dq_{n,i}}{dt}(\hat{q}) = 0.$$

In particular, $\hat{q}$ is a stationary point implies $\forall n \in \mathcal{N}, i \in \mathcal{I}_n,\ \hat{q}_{n,i} = 0$ or $f_{n,i}(\hat{q}) = \overline{f}_n(\hat{q})$.

Intuitively, this dynamics can be understood as an update mechanism where the probability for each player to choose actions whose expected payoffs are above average will increase in time, while non profitable actions will gradually be abandoned.

Let us notice that the trajectories of the replicator dynamics remain inside the domain Δ. Also, from [10], the potential function F is a Lyapunov function for the replicator dynamics, hence is strictly increasing along the trajectories which are not stationary.

In this context, a closed set A is *Lyapunov stable* if, for every neighborhood B, there exists a neighborhood $B' \subset B$ such that the trajectories remain in B for any initial condition in B'. A is *asymptotically stable* if it is Lyapunov stable and is an attractor (i.e. there exists a neighborhood C such that all trajectories starting in C converge to A).

Finally, let us recall the Lasalle principle [11] which states that the existence of a (strictly increasing) Lyapunov function implies that all the trajectories converge to connected sets of equilibrium points.

Proposition 2: All the asymptotically stable sets of the replicator dynamics are faces of the domain. These faces are sets of equilibrium points for the replicator dynamics.

Proof: We show that any set which is not a face of the domain is not an attractor. This results from a property discovered by E. Akin [7] which states that the replicator dynamics preserves a certain form of volume.

Let A be a asymptotically stable set of the replicator dynamics. Since the domain Δ is polyhedral, A is included in a face F_A of Δ. The support of the face $S(F_A)$ is the set of subscripts (n,i) such that there exists $q \in A$ with $q_{n,i} \neq 0$ or 1. The relative interior of the face is $\mathrm{I}nt(F_A) = \{q \in F(A) s.t. \forall (n,i) \in S(F_A), 0 < q_{n,i} < 1\}$.

Furthermore, it should be clear that faces are invariant under the replicator dynamics. Hence on the face F_A, by using the transformation $v_{n,i} \stackrel{\text{def}}{=} log(\frac{q_{n,i}}{q_{n,i_n}}), \forall q \in \mathrm{I}nt(F_A)$, one can see that

$$\frac{\partial}{\partial v_{n,i}} \frac{dv_{n,i}}{dt} = 0, \forall n \in \mathcal{N}, i \in \mathcal{I}.$$

Up to this transformation, the divergence of the vector field is null on F_A. Using Liouville's theorem [7], we infer that the transformed dynamics preserves volume in $\mathrm{I}nt(F_A)$. The only possibility to preserve volume in the set $\mathrm{I}nt(F_A)$ containing an asymptotically stable set is that all points in $\mathrm{I}nt(F_A)$ are equilibrium points. By continuity of the vector field, all the points in face F_A are equilibria. Finally, since A is asymptotically stable, this means that $A = F_A$. ■

We say that $s = (s_n)_{n \in \mathcal{N}}$ is a *pure Nash Equilibrium* if $\forall n \in \mathcal{N}$, $\forall s'_n \neq s_n, U_n(s_1 \dots s_n \dots s_{|\mathcal{N}|}) \geq U_n(s_1 \dots s'_n \dots s_{|\mathcal{N}|})$.

Remark 2: Let q be a pure strategy. We denote by i_n the choice of player n such that $q_{n,i_n} = 1$. Then, q is a pure Nash equilibrium is equivalent to:

$$\forall n \in \mathcal{N}, \forall j \neq i_n, f_{i_n,n}(q) \geq f_{j,n}(q).$$

Proposition 3: All asymptotically stable points of the replicator dynamics are pure Nash equilibria.

Proof: Let $\hat{q}$ be an asymptotically stable point. Then $\hat{q}$ is a face of Δ by Proposition 2 (*i.e.* a 0-1 point), with, say $\hat{q}_{n,i} = 1$. Assume that $\hat{q}$ is not a Nash equilibrium. Then, there exists $j \neq i$ such that $f_{j,n}(\hat{q}) \geq f_{i,n}(\hat{q})$. Now, consider a point $q' = \hat{q} + \epsilon e_{n,j} - \epsilon e_{n,i}$. Notice that $f_{n,i}(q') = f_{n,i}(\hat{q})$ since q' and $\hat{q}$ only differ on components concerning user n. Then starting in q', the replicator dynamics is

$$\begin{aligned}
\tfrac{dq_{n,i}}{dt}(q') &= q'_{n,i}(f_{n,i}(q') - ((1-\epsilon) f_{n,j}(q') + \epsilon f_{n,i}(q')) \\
&= (1-\epsilon)(f_{n,i}(\hat{q}) - ((1-\epsilon) f_{n,j}(\hat{q}) + \epsilon f_{n,i}(\hat{q})) \\
&= -\epsilon(1-\epsilon)(f_{n,j}(\hat{q}) - f_{n,i}(\hat{q})) \\
&\leq 0,
\end{aligned}$$

and $\frac{dq_{n,j}}{dt}(q') = -\frac{dq_{n,i}}{dt}(q') \geq 0$.

For all users $m \neq n$, $\forall u \in \mathcal{I}_m, q'_{m,u} \in \{0,1\}$, then $\frac{dq_{m,k}}{dt}(q') = q'_{m,k}\left(f_{n,k}(q') - \sum_u q'_{m,u}(f_{n,u}(q'))\right) = 0.$

Therefore starting from q', the dynamics keeps moving in the direction $e_{n,j} - e_{n,i}$ (or stays still) and does not converge to $\hat{q}$. This contradicts the fact that $\hat{q}$ is asymptotically stable. ■

Proposition 4: Allocation games with repercussion utilities admit at least one pure Nash equilibrium.

Proof: Allocation games with repercussion utilities admit a potential that is a Lyapunov function of their replicator dynamics. Since the domain Δ is compact, the Lyapunov function reaches its maximal value inside Δ. The argmax of the Lyapunov function form an asymptotically stable sets A of equilibrium points. By Proposition 2, these sets are faces of the domain (hence contain pure points). All points in A are Nash equilibrium points by using a similar argument as in Proposition 3. This concludes the proof. ■

B. A Stochastic Approximation of the Replicator Dynamics.

In this section, we present an algorithmic construction of the players' strategies that select a pure Nash equilibrium for the game with repercussion utilities. A similar learning mechanism is proposed in [12]. We now assume a discrete time, in which at each epoch t, players take random decision $s_n(t)$ according to their strategy $q_n(t)$, and update their strategy profile according to their current payoff. We look at the following algorithm ($\forall n \in \mathcal{N}, i \in \mathcal{I}_n$):

$$q_{n,i}(t+1) = q_{n,i}(t) + \epsilon\, r_n(\ell^{s_n}(s))\,(1_{s_n=i} - q_{n,i}(t)), \quad (5)$$

where $s_n = s_n(t)$, $\epsilon > 0$ is the constant step size of the algorithm, and $1_{s_n=i}$ is equal to 1 if $s_n = i$, and 0 otherwise. Note that if ϵ is small enough, then $q_{n,i}$ remains in the interval $[0;1]$. Strategies are initialized with value $q(0) = q_0$. The step-size is chosen to be constant in order to have higher convergence speed than with decreasing step size.

One can notice that this algorithm is fully distributed, since for each player n, the only information needed is $r_n(\ell^{s_n}(s))$. Furthermore, at every iteration, each player only need the utility on one action (which is randomly chosen). In applicative context, this means that a player does not have to scan all the action before update her strategy, what would be costly.

Below, we provide some intuition on why the algorithm is characterized by a differential equation, and how it asymptotically follows the replicator dynamics (4). Note that we can re-write (5) as:

$$q_{n,i}(t+1) = q_{n,i}(t) + \epsilon\, b(q_{n,i}(t), s_n(t)).$$

Then, we can split b into its expected and martingale components:

$$\begin{aligned}
\bar{b}(q_{n,i}(t)) &= \mathbb{E}[b(q_{n,i}(t), s_n(t))] \\
\nu(t) &= b(q_{n,i}(t), s_n(t)) - \bar{b}(q_{n,i}(t)).
\end{aligned}$$

Again, (5) can be re-written as:

$$\frac{q_{n,i}(t+1) - q_{n,i}(t)}{\epsilon} = \overline{b}(q_{n,i}(t)) + \nu(t).$$

As $\nu(t)$ is a random difference between the update and its expectation, then by application of a law of large numbers, for small ϵ, this difference goes to zero. Hence, the trajectory of $q_{n,i}(t)$ in discrete time converges to the trajectory in continuous time of the differential equation:

$$\begin{cases} \dfrac{dq_{n,i}}{dt} = \overline{b}(q_{n,i}), \quad \text{and} \\ q(0) = q_0. \end{cases}$$

Let us compute $\overline{b}(q_{n,i})$ (for ease of notations, we omit the dependence on time t):

$$\begin{aligned} \overline{b}(q_{n,i}) &= \mathbb{E}[b(q_{n,i}, s_n)] \\ &= q_{n,i}(1 - q_{n,i}) f_{n,i}(q) - \sum_{j \neq i} q_{n,j} q_{n,i} f_{n,j}(q) \\ &= q_{n,i}(f_{n,i}(q) - \sum_{j} q_{n,j} f_{n,j}(q)) \\ &= q_{n,i}(f_{n,i}(q) - \overline{f}(q)). \end{aligned}$$

Then, $q_{n,i}(t)$ follows the replicator dynamics.

C. Properties of the algorithm.

The algorithm is designed so as to follow the well-known replicator dynamics. Furthermore, the stochastic aspect of the algorithm provides some stability to the solution: whereas the deterministic solution of a replicator dynamics may converge to a saddle point, this cannot happen with the stochastic algorithm. The use of repercussion utilities provides a potential to the companion game and it is known that the potential is a Lyapunov function for the replicator dynamics, hence the potential is increasing along the trajectories. The following theorem aggregates the main results about the algorithm applied on repercussion utilities.

Theorem 2: The algorithm (5) converges to a pure point which is locally optimal for the potential function, and a Nash equilibrium of the allocation game with repercussion utilities.

Proof: The algorithm is a stochastic algorithm with constant step size. From Theorem 8.5.1 of Kushner and Yin [13], we infer that the algorithm asymptotically weakly converges as $\epsilon \to 0$ to the solution of an ode, which is in our case the replicator dynamics (4) (it is a particular case of the theorem in which conditions of the theorem hold: all variables are in a compact set and the dynamics is continuous). Furthermore, since the replicator dynamics admits a Lyapunov function, that is increasing along the trajectories, then the set to which the sequence $q(t)$ converges is an asymptotically stable set of the replicator dynamics. From Proposition 2, the only asymptotically stable sets of the dynamics are faces which are sets of stationary points. Hence the algorithm converges to faces which are asymptotically stable. Let $q(t)$ be a trajectory of the algorithm. Suppose that it converges to a closed set $A \subset F_A$, where F_A is an asymptotically stable face of the domain Δ (hence the vector field is null on F_A). We further assume that A does not contain any pure point.

Let $\hat{q}(0) \in A$. Then, the trajectory $\hat{q}(t)$ following the algorithm stays in F_A. Furthermore, $\hat{q}(t)$ converges almost surely to a pure point. Indeed:

$$\begin{aligned} &\mathbb{E}[\hat{q}_{n,i}(t+1)|\hat{q}(t)] \\ &\quad = \hat{q}_{n,i}(t)(\hat{q}_{n,i}(t) + \epsilon f_{n,i}(\hat{q}(t))(1 - \hat{q}_{n,i}(t))) \\ &\quad\quad + \sum_{j \neq i} \hat{q}_{n,j}(\hat{q}_{n,i}(t) - \epsilon f_{n,j}(\hat{q}(t))\hat{q}_{n,i}(t)) \\ &\quad = \hat{q}_{n,i}(t) + \epsilon q_{n,i}(f_{n,i}(\hat{q}(t)) - \overline{f}_n(\hat{q}(t))). \end{aligned}$$

Since at a mixed stationary point $f_{n,i}(\hat{q}) = \overline{f}_n(\hat{q})$, then:

$$\mathbb{E}[\hat{q}_n(t+1)|\hat{q}(t)] = \hat{q}_n(t).$$

Hence the process $(\hat{q}_n(t))_t$ is a martingale, and is almost surely convergent. The process converges necessarily to a fixed point of the iteration $\hat{q}_{n,i}(t+1) = \hat{q}_{n,i}(t) + \epsilon\, r_n(\ell^{s_n}(s))\,(1_{s_n=i} - \hat{q}_{n,i}(t))$, and the sole fixed points are pure points (since the step size ϵ is constant).

We now show that if $q(t)$ and $\hat{q}(t)$ are defined on the same probability space and are coupled (driven by the same sequence ω) and if the distance $d(q(0), \hat{q}(0))$ is smaller than ν, then, with a probability p that goes to 1 when ν goes to 0, $q(t)$ goes away from A, which is a contradiction.

First, we denote by Q the pure point attained by $\hat{q}(t)$ under ω. We can separate Q from A since they are two closed sets: in particular, for ν sufficiently small, there exists a finite time T such that the distance $d(\hat{q}(T), A) > \nu$. Let us define $d_t = d(q(t), \hat{q}(t))$. Then, one can check that, under ω, $d_{t+1} \leq d_t(1 + \epsilon r)$, where $r \stackrel{\text{def}}{=} \max_n \max_s |r_n(\ell^{s_n}(s))|$, with probability at least $p(d_t) \stackrel{\text{def}}{=} 1 - d_t \sum_{n=1}^{N} I_n$. Indeed the vector of actions $s(q)$ is the same as $s(\hat{q})$ as long as ω picks the same choice for all players. This happens with probability

$$1 - \sum_{n=1}^{N} \sum_{i=1}^{I_n} |\sum_{k=1}^{i} q_{n,k}(t) - \hat{q}_{n,k}(t)|.$$

See Figure 1 for an illustration of this.

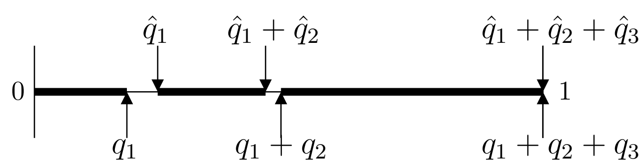

Fig. 1. The thick line shows the measure of the set of all ω corresponding to the same choices for player 1 (with 3 choices).

The lower bound follows from the inequality $|\sum_{k=1}^{i} q_{n,k}(t) - \hat{q}_{n,k}(t)| < d_t$.

Then, $d_T \leq \nu(1 + \epsilon r)^T$ with probability at least $p = \prod_{t=0}^{T} p(d_t)$. As ν goes to zero, d_T goes to 0, and p goes to 1. Hence, with a probability closed to 1, $d(q(T), A) > \nu$. This is a contradiction with the fact that $q(t)$ converges to A.

Finally, the fact that it is a Nash equilibrium follows from Proposition 3. ■

One can notice that the convergence of the algorithm to a pure point relies on the fact that the step size ϵ is constant. If it were decreasing, like in Chapter 2 of [14], the algorithm would converge to an equilibrium point in a stable face, that need not be pure.

The combination of both algorithm (5) and repercussion utilities provides an iterative method to select a pure allocation which is stable, and locally optimal. This can be called *selection algorithm.*

IV. Global Maximum vs Local Maximum for the Selection Algorithm.

In the previous section, we showed that the algorithm converges to a local maximum of the potential function. This induces that if there is only one local maximum, the algorithm attains the global maximum. This arises for instance if the potential function is concave. Without the uniqueness of the local maximum, there is no guaranty of convergence to the global maximum. Hence, assume there are multiple local maxima (that are pure points), which is common when the payoffs are random. Each of them is an attractor for the replicator dynamics. In this section, we investigate the following question: does the initial point of the algorithm belongs to the basin of attraction of the global maximum?

Since every player has no preference at the beginning of the algorithm, we assume that initially, $\forall n \in \mathcal{N}, i \in \mathcal{I}_n, q_{n,i}(0) = \frac{1}{|\mathcal{I}_n|}$. In the following sub-section we show that in the case of two players, both having two choices, $q(0)$ is in the basin of attraction of the global maximum. Then, in Subsections IV-B and IV-C, we give counter examples to show that the result does not extend to the general case of more than two players or more than two strategies.

A. Case of two players and two choices

Proposition 5: In a two players, two actions allocation game with repercussion utilities, the initial point of the algorithm is in the basin of attraction of the global maximum.

Proof: Both players 1 and 2 can either take action a or b. We denote by x the probability for player 1 to choose a, and by y the probability for 2 to choose a. We denote by $K = (k_{i,j})_{i,j\in\{0,1\}}$ the matrix such that $k_{i,j} \stackrel{\text{def}}{=} F(x = i, y = j)$, where $F(x, y)$ is the potential function. Then, the dynamics (4) can be rewritten:

$$\begin{cases} \dfrac{dx}{dt} = x(1-x)(k_{0,1} - k_{0,0} + Ky) \\ \dfrac{dy}{dt} = y(1-y)(k_{1,0} - k_{0,0} + Kx), \end{cases} \quad (6)$$

where $K = k_{1,1} + k_{0,0} - k_{0,1} - k_{1,0}$. Note that in a two-player two-action game, there are at most two local maxima. Suppose that in the considered game, there are two local maxima. They are necessarily attained either at points $(0, 0)$ and $(1, 1)$ or at points $(0, 1)$ and $(1, 0)$. Without loss of generality, we can assume the former case. Hence, $k_{0,0}$ and $k_{1,1}$ are local maxima, and $k_{1,1} > k_{0,0} + \epsilon$, where $\epsilon > 0$.

We now define set E and function V as follows:

$$\begin{aligned} V(x,y) &= |1-x| + |1-y|, \\ E &= \{(x,y) : x + y > 1,\ 0 < x,\ y < 1\}. \end{aligned}$$

(V is actually the distance of (x, y) to the point $(1, 1)$ for the 1-norm.) We next show that V is a Lyapunov function for the dynamics on the open set E. Toward that goal, it is sufficient to show that

$$L(x,y) \stackrel{\text{def}}{=} \frac{\partial V}{\partial x}(x,y)\frac{dx}{dt} + \frac{\partial V}{\partial y}(x,y)\frac{dy}{dt} < 0.$$

First, note that $\forall (x, y) \in E$, $V(x, y) = 2 - x - y$. Hence, from Eq. 6,

$$L(x,y) = -x(1-x)(k_{0,1} - k_{0,0} + Ky) - y(1-y)(k_{1,0} - k_{0,0} + Kx). \quad (7)$$

Let also be D the open segment $\{(x, y) : x + y = 1, 0 < x, y < 1\}$. Trivially,

$$\forall (x,y) \in D, \quad L(x, y = 1-x) = -x(1-x)(k_{1,1} - k_{0,0}) < 0. \quad (8)$$

Let us finally consider the segment

$$S(x_0) = \{(x,y) : x + y \geq 1, x = x_0, 0 \leq y \leq 1\}.$$

Figure 2 summarizes the different notations introduced.

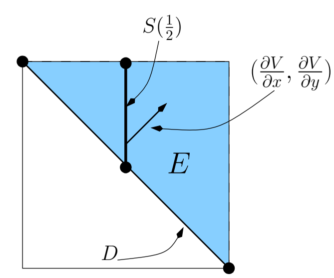

Fig. 2. Proof of Prop. 5: Summary of notations

Since $E \subset \bigcup_{0<x<1} S(x)$, it is sufficient to show the negativeness of L on $S(x)$ for all x. Let us denote by $L_x(y)$ the restriction of L on $S(x)$. From Eq. 8, we have $L_x(1 - x) < 0$. Furthermore, $L_x(y)$ is a quadratic function and its discriminant is $4(k_{1,0} - k_{0,0})(k_{1,1} - k_{0,1})$, hence is negative. So, for all x, $L_x(y)$ is negative (strictly). Finally, L is negative (strictly) in E and hence non-positive in a neighborhood of E.

Therefore, V is a Lyapunov function for the dynamics on a neighborhood of the open set E. More precisely, V is strictly decreasing on the trajectories of the dynamics starting in the set E, hence they converge to the unique minimum of V which is the point $(1, 1)$. This applies to the initial point $(0.5, 0.5)$. ■

Figure 3 illustrates this result: consider a two player (numbered 1 and 2), two strategy (denoted by A and B)

game. Let x (resp. y) be the probability for player 1 (resp. 2) to take action A. While two (local) maxima exist - namely $(1,1)$ and $(0,0)$ - the surface covered by the basin of attraction of the global optimum (which is $(1,1)$ in this example) is greater than those of the other one. A by-product is that the dynamics starting in point $(0.5, 0.5)$ converges to the global optimum.

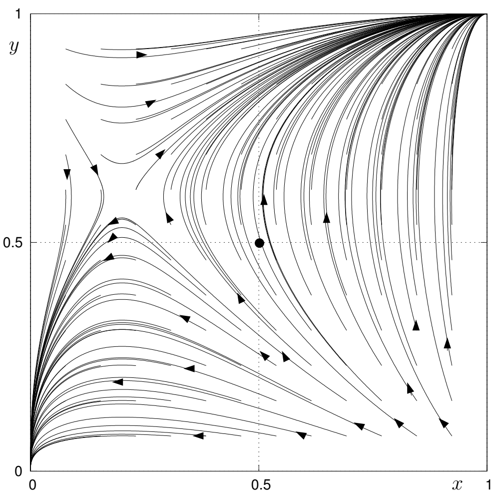

Fig. 3. An example with 2 players with 2 choices each. There are 2 maxima. The point $(\frac{1}{2}, \frac{1}{2})$ is inside the attracting basin of the global maximum.

Unfortunately, this appealing result cannot be generalized to more players or more actions, as exemplified in the following subsections.

B. Extension to more than two players

Example 2: Let us consider a three player game : $(\mathcal{N},\mathcal{I},\mathcal{U})$ with $\mathcal{N} = \{1,2,3\}$, $\mathcal{I} = \{A,B\}$, and $\mathcal{U} = (u_n(i,j,k))_{n\in\{1,2,3\},i,j,k\in\{A,B\}}$, where i (resp. j), denotes the choice of player 1 (resp. 2). The matrix representation of (u_1,u_2,u_3) are given below:

$$(u_1,u_2,u_3)(i,j,1) = \begin{pmatrix} (9,6,4) & (5,5,5) \\ (5,8,1) & (2,4,4) \end{pmatrix},$$

$$(u_1,u_2,u_3)(i,j,2) = \begin{pmatrix} (7,2,8) & (5,4,7) \\ (6,3,3) & (10,2,8) \end{pmatrix}.$$

Note that this game has no pure strategies Nash equilibrium and a single mixed strategies Nash equilibrium, which is $(x,y,z) = (1/3, 5/6, 0)$. The corresponding value of the potential function is $87/6 = 14.5$.

The repercussion utility matrices are:

$$(r_1,r_2,r_3)(i,j,1) = \begin{pmatrix} (10,9,10) & (6,5,5) \\ (5,5,6) & (1,1,4) \end{pmatrix},$$

$$(r_1,r_2,r_3)(i,j,2) = \begin{pmatrix} (6,4,8) & (5,3,7) \\ (1,3,4) & (9,11,14) \end{pmatrix}.$$

This game has two pure Nash equilibria, that are $(x,y,z) = (1,1,1)$ and $(x,y,z) = (0,0,0)$, corresponding to values of the potential function that are respectively 29 and 34.

Figure 4 shows that the trajectory starting at point $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ converges to the local maximum $(x,y,z) = (1,1,1)$ instead of the global maximum $(x,y,z) = (0,0,0)$. Note that the performance of the local maximum is way ahead that of the Nash Equilibrium in the original game.

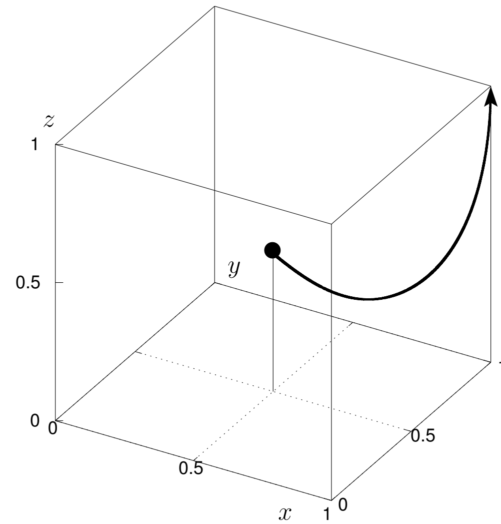

Fig. 4. Example with 3 players, with 2 choices each. The figure represents the dynamic trajectory starting from the point $(x,y,z)(0) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, with x (resp. y, z) the probability for player 1 (resp. 2, 3) to adopt action A. The dynamics converges to the point $(1,1,1)$ whereas the global maximum is $(0,0,0)$.

C. Extension to more than two choices

Example 3: Let us now consider the two player game $(\mathcal{N},\mathcal{I},\mathcal{U})$ with $\mathcal{N} = \{1,2\}$, $\mathcal{I} = \{A,B,C\}$, $\mathcal{U} = (u_n(i,j))_{n\in\{1,2\},i\in\{A,B\},j\in\{A,B,C\}}$. (Note that in this example, only the second player has three possible choices).

The payoff matrix is:

$$(u_1,u_2)(i,j) = \begin{pmatrix} (6,3) & (-3,11) & (-3,10) \\ (0,2) & (-1,1) & (0,10) \end{pmatrix}.$$

The companion game is:

$$(r_1,r_2)(i,j) = \begin{pmatrix} (7,12) & (-3,11) & (-3,10) \\ (0,2) & (-11,0) & (0,10) \end{pmatrix}.$$

The original game has one single pure Nash equilibria which is (B,C) resulting in the value 10 for the potential function and no mixed strategies equilibria exists.

The companion game has two pure Nash equilibria that are (A,A) and (B,C), corresponding to values of the potential function of 9 and 10 respectively.

Denote x the probability for player 1 to choose action A and y_1 (resp. y_2) the probability for player 2 to choose action A (resp. B). Then, the global maximum of the potential function is 10, and is attained when $x = y_1 = y_2 = 0$. Figure 5 shows that the trajectory starting at point $(\frac{1}{2}, \frac{1}{3}, \frac{1}{3})$ converges to the local maximum $(1,1,0)$, corresponding to Nash equilibrium (A,A) of the companion game, which is inefficient. Interestingly in this example, the unique Nash equilibrium of the original game corresponds to the global maximum of the game.

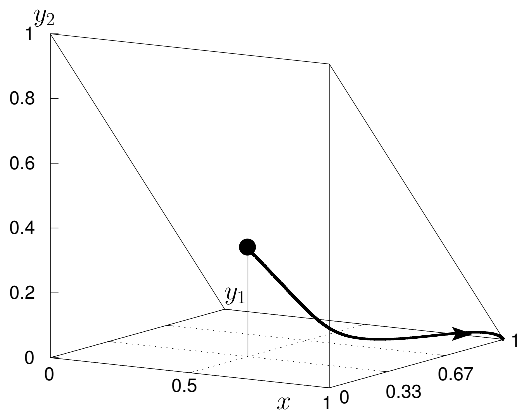

Fig. 5. Example with 2 players. The first one has 2 choices and the second one has 3 choices. Here we display the 3-dimensional plot of y_1 vs x and y_2 vs x. The dynamics starting in $(1/2, 1/3, 1/3)$ converges to the point $(1, 1, 0)$ whereas the global maximum is $(0, 0, 0)$.

V. Conclusion

This paper presents a distributed stochastic algorithm that selects efficient pure allocations in game by introducing the concept of repercussion utilities. The construction of repercussion utilities insures that the game becomes a potential game and the algorithm is an approximation of the replicator dynamics for this game. The trajectories are shown to converge to pure allocations that are Nash equilibria for the later game and locally socially optimal for the original payoffs.

References

[1] J. W. Weibull, *Evolutionary Game Theory*. MIT Press, 1995.

[2] J. Hofbauer and K. Sigmund, “Evolutionary game dynamics,” *Bull. Amer. Math. Soc.*, vol. 40, pp. 479–519, 2003.

[3] S. Shakkottai, E. Altman, and A. Kumar, “The Case for Non-cooperative Multihoming of Users to Access Points in IEEE 802.11 WLANs,” in *Proc. of INFOCOM*, 2005.

[4] S. Fischer, N. Kammenhuber, and A. Feldmann, “Replex : Dynamic Traffic Engineering Based on Wardrop Routing Policies,” in *CoNext*, 2006.

[5] R. W. Rosenthal, “A Class of Games Possessing Pure-Strategy Nash Equilibria,” *Int. J. Game Theory*, vol. 2, pp. 65–67, 1973.

[6] J. Hofbauer, “Evolutionary Dynamics for Bimatrix Games : A Hamiltonian System?” *Journal of mathematical biology*, vol. 34, pp. 675–688, 1996.

[7] I. Eshel and E. Akin, “Coevolutionary Instability of Mixed Nash Solutions,” *Journal of Mathematical Biology*, vol. 18, pp. 123–133, 1983.

[8] P. Coucheney, C. Touati, and B. Gaujal, “Fair and efficient user-network association algorithm for multi-technology wireless networks,” in *Infocom*, 2009.

[9] D. Monderer and L. S. Shapley, “Potential Games,” *Games and Economic Behavior*, vol. 14, pp. 124–143, 1996.

[10] W. H. Sandholm, “Potential Games with Continuous Player Sets,” *Journal of Economic Theory*, vol. 24, pp. 81–108, 2001.

[11] J.-P. LaSalle, “Some Extensions of Liapunov’s Second Method,” *IRE Transactions on Circuit Theory*, vol. 7, pp. 520–527, 1960.

[12] P. Sastry, V. Phansalkar, and A. Thathachar, “Decentralized Learning of Nash Equilibria in Multi-person Stochastic Games with Incomplete Information,” *IEEE Transactions on System, man, and cybernetics*, vol. 24, no. 5, pp. 769–777, 1994.

[13] Kushner, J. Harold, and G. G. Yin, *Stochastic Approximation Algorithms and Applications*. Springer-Verlag, New-York, 1997.

[14] V. S. Borkar, *Stochastic Approximation: a Dynamical Systems Viewpoint*, Hardback, Ed. Cambridge Univ Pr, 2008.