Distributed Learning of Wardrop Equilibria*

Dominique Barth[2], Olivier Bournez[1], Octave Boussaton[1], and Johanne Cohen[1]

[1] LORIA/INRIA-CNRS-UHP, 615 Rue du Jardin Botanique, 54602 Villers-Lès-Nancy, FRANCE
{Olivier.Bournez,Octave.Boussaton,Johanne.Cohen}@loria.fr

[2] Laboratoire PRiSM Université de Versailles, 45, avenue des Etats-Unis, 78000 Versailles, FRANCE
Dominique.Barth@prism.uvsq.fr

Abstract. We consider the problem of learning equilibria in a well known game theoretic traffic model due to Wardrop. We consider a distributed learning algorithm that we prove to converge to equilibria. The proof of convergence is based on a differential equation governing the global macroscopic evolution of the system, inferred from the local microscopic evolutions of agents. We prove that the differential equation converges with the help of Lyapunov techniques.

1 Introduction

We consider in this paper a well-known game theoretic traffic model due to Wardrop [34] (see also [30] for an alternative presentation). This model was conceived to represent road traffic with the idea of an infinite number of agents being responsible for an infinitesimal amount of traffic each. A network equipped with non-decreasing latency functions mapping flow on edges to latencies is given. For each of several commodities a certain amount of traffic, or flow demand, has to be routed from a given source to a given destination via a collection of paths. A flow in which for all commodities the latencies of all used paths are minimal with respect to this commodity is called a Wardrop equilibrium of the network. Whereas this is well-known that such equilibria can be solved by centralized algorithms in polynomial time, as in [31] we are interested in *distributed* algorithms to compute Wardrop equilibria.

Actually, we consider in this paper a slightly different setting from the original Wardrop model [34] (similar to the one considered in [31]): we consider that the flow is controlled by a finite number N of agents only, each of which is responsible for a fraction of the entire flow of one commodity. Each agent has a set of admissible paths among which it may distribute its flow. Each agent aims at balancing its own flow such that the jointly computed allocation will be a Wardrop equilibrium.

We consider for these networks a dynamics for learning Nash equilibria in multiperson games presented in [28]. This dynamics was proved to be such that

* This work was supported by French Ministry of Research through ANR Project SOGEA.

all stable stationary points are Nash equilibria for general games. Whereas for general games, the dynamic is not necessarily convergent [28], we prove that the dynamics is convergent for linear Wardrop networks. We call linear Wardrop networks the case where latency functions are affine.

Our motivation behind this study is twofold. On one hand, we want to understand if, how and when equilibria can be learned in games. The dynamics considered here has both the advantage of being decentralized and of requiring partial and very limited informations. It is indeed a discrete stochastic dynamics played by the N players, each of which chooses between a finite number of strategies (paths) they can use at each instant. After each play, players are rewarded with random payoffs. In order for players to learn optimal strategies, the game is played repeatedly. Hence, after each play, each player updates his strategy based solely on his current action and payoff. Our interest is in learning equilibria in games, through distributed algorithms and with minimal informations for players.

On the other hand, our interest in this dynamics comes from a general plan of one of the authors here behind his study of computational properties of continuous time dynamical systems: see e.g. [6, 5, 3], or survey [4]. As we noticed in the introduction of this latter survey, continuous time systems arise in experimental sciences as soon as a huge population of agents (molecules, individuals,. . .) is abstracted into real quantities. Wardrop networks constitute a clear and nice example where this holds for systems coming from road traffic [34], or from computer network traffic [30]. One strong motivation behind the current work is also to discuss the efficiency attained by such networks, and more generally by distributed systems. Our approach is based on a macroscopic abstraction of the microscopic rules of evolutions of the involved agents, in terms of a differential equation governing the global state of the system. This differential equation is proved to converge for linear Wardrop networks, using Lyapunov techniques. For general games the considered dynamics is not always convergent [28].

2 Related Work

For a survey on continuous time systems, and on their computational properties, we refer to [4].

In the history of game theory, various algorithms for learning equilibrium states have been proposed: centralized and decentralized (or distributed) algorithms, games with perfect, complete or incomplete information, with a restricted number of players, etc... See e.g. [23] for an introduction to the learning automata model, and the general references in [28] for specific studies for zero-sum games, N-person games with common payoff, non-cooperative games, etc...

Wardrop traffic model was introduced in [34] to apprehend road traffic. More recently, it has often been considered as a model of computer network traffic. The price of anarchy, introduced by [22] in order to compare costs of Nash equilibria to costs of optimal (social) states has been intensively studied on these games: see e.g. [30, 29, 7, 16, 8].

There are a few works considering dynamical versions of these games, where agents try to learn equilibria, in the spirit of this paper.

In [13], extending [14] and [15], Fischer and al. consider a game in the original Wardrop settings, i.e. a case where each user carries an infinitesimal amount of traffic. At each round, each agent samples an alternative routing path and compares the latency on its current path with the sampled one. If an agent observes that it can improve its latency, then it switches with some probability that depends on the improvement offered by the better paths, otherwise, it sticks to its current path. Upper bounds on the time of convergence were established for asymmetric and symmetric games.

In [31] Fischer and al. consider a more tractable version of this learning algorithm, considering a model with a finite number of players, similar to ours. The considered algorithm, based on a randomized path decomposition in every communication round, is also very different from ours.

Nash equilibria learning algorithms for other problems have also been considered recently, in particular for load balancing problems.

First, notice that the proof of existence of a pure Nash equilibria for the load balancing problem of [22] can be turned into a dynamics: players play in turn, and move to machines with a lower load. Such a strategy can be proved to lead to a pure Nash equilibrium. Bounds on the convergence time have been investigated in [10, 11]. Since players play in turns, this is often called the *Elementary Step System*. Other results of convergence in this model have been investigated in [17, 25, 27].

Concerning models that allow concurrent redecisions, we can mention the followings works. In [12], tasks are allowed in parallel to migrate from overloaded to underloaded resources. The process is proved to terminate in expected $O(\log \log n + \log m)$ rounds.

In [2] is considered a distributed process that avoids that latter problem: only local knowledge is required. The process is proved to terminate in expected $O(\log \log n + m^4)$ rounds. The analysis is also done only for unitary weights, and for identical machines. Techniques involved in the proof, relying on martingale techniques, are somehow related to techniques for studying the classical problem of allocating balls into bins games as evenly as possible.

The dynamics considered in our present paper has been studied in [28] for general stochastic games where Thathachar & al. proved that the dynamics is weakly convergent to some function solution of an ordinary differential equation. This ordinary differential equation turns out to be a replicator equation. While a sufficient condition for convergence is given, no error bounds are provided and no Lyapunov function is established for systems similar to the ones considered in this paper.

Replicator equations have been deeply studied in evolutionary game theory [20, 35]. Evolutionary game theory isn't restricted to these dynamics but considers a whole family of dynamics that satisfy a so called folk theorem in the spirit of Theorem 2.

Bounds on the rate of convergence of fictitious play dynamics have been established in [18], and in [21] for the best response dynamics. Fictitious play has been reproved to be convergent for zero-sum games using numerical analysis methods or, more generally, stochastic approximation theory: fictitious play can be proved to be a Euler discretization of a certain continuous time process [20].

A replicator equation for allocation games has been considered in [1], where authors establish a potential function for it. Their dynamics is not the same as ours : we have a replicator dynamics where fitnesses are given by true costs, whereas for some reason, marginal costs are considered in [1].

3 Wardrop's Traffic Model

A routing game [34] is given by a graph $G = (V, E)$. To each edge $e \in E = (v_1, v_2)$, where $v1, v2 \in V$, is associated a continuous and non decreasing latency function $\ell_e : [0, 1] \to \mathbb{R}^+$. We are given $[k] = \{1, 2, ..., k\}$ a set of commodities, each of which is specified by a triplet consisting in: a source-destination pair (s_i, t_i), $G_i = (V_i, E_i)$ a directed acyclic sub-graph of G connecting s_i to t_i, and a flow demand $r_i \geq 0$. The total demand is $r = \sum_{i\in[k]} r_i$. We assume without loss of generality that $r = 1$. Let $\mathcal{P}_i$ denote the admissible paths of commodity i, i.e. all paths connecting s_i and t_i in G_i. We may assume that the sets $\mathcal{P}_i$ are disjoint and define i_P to be the unique commodity to which path P belongs.

A non-negative path flow vector $(f_P)_{P\in\mathcal{P}}$ is *feasible* if it satisfies the flow demands $\sum_{P\in\mathcal{P}_i} f_P = r_i$, for all $i \in [k]$. A path flow vector $(f_P)_{P\in\mathcal{P}}$ induces an edge flow vector $f = (f_{e,i})_{e\in E, i\in[k]}$ with $f_{e,i} = \sum_{P\in\mathcal{P}_i : e\in P} f_P$. The total flow on edge e is $f_e = \sum_{i\in[k]} f_{e,i}$. The latency of an edge e is given by $\ell_e(f_e)$ and the latency of a path P is given by the sum of the edge latencies $\ell_P(f) = \sum_{e\in P} \ell_e(f_e)$.

A flow vector in this model is considered stable when no fraction of the flow can improve its latency by moving unilaterally to another path. It is easy to see that this implies that all used paths must have the same (minimal) latency.

Definition 1 (Wardrop Equilibrium). *A feasible flow vector f is at a Wardrop equilibrium if for every commodity $i \in [k]$ and paths $P_1, P_2 \in \mathcal{P}_i$, with $f_{P_1} > 0$, $\ell_{P_1}(f) \leq \ell_{P_2}(f)$ holds.*

We now extend the original Wardrop model [34] to an N player game as follows (a similar setting has been considered in [31]). We assume that we have a finite set $[N]$ of players. Each player is associated to one commodity, and is supposed to be in charge of a fraction w_i of the total flow r_i of a fixed commodity.

Each player (agent) aims at balancing its own flow in such a way that its latency becomes minimal.

In this present work, we will narrow down our investigations to the case of linear cost functions: we assume that for every edge e, there are some constants α_e, and β_e such that $\ell_e(\lambda) = \alpha_e \lambda + \beta_e$.

4 Game Theoretic Settings

We assume that players distribute their flow selfishly without any centralized control and only have a local view of the system. All players know how many paths are available. We suppose that the game is played repeatedly. At each elementary step t, players know their cost and the path they chose at step $t' < t$. Each one of them selects a path at time step t according to a mixed strategy $q_j(t)$, with $q_{j,s}$ denoting the probability for player j to select path s at step t.

Any player associated to commodity i has the finite set of actions $\mathcal{P}_i$. We assume that paths are known by and available to all of the players. An element of $\mathcal{P}_i$, is called a *pure strategy*.

Define payoff functions $d_i : \prod_{j=1}^{N} \mathcal{P} \to [0,1], 1 \leq i \leq N$, by:

$$d_i(a_1, a_2, ..., a_N) = cost\ for\ i \mid \text{player } j \text{ chose action } a_j \in \mathcal{P}, 1 \leq j \leq N \quad (1)$$

where $(a_1, ..., a_N)$ is the set of pure strategies played by all the players.

In our case,

$$d_i(a_1, a_2, ..., a_N) = \ell_{a_i}(f),$$

where f is the flow induced by $a_1, a_2, ..., a_N$.

We call it the payoff function, or utility function of player i and the objective of all players is to minimize their payoff.

Now, we want to extend the payoff function to mixed strategies. To do so, let $\mathcal{S}_p$ denote the simplex of dimension p which is the set of p-dimensional probability vectors:

$$\mathcal{S}_p = \{q = (q_1, \cdots, q_p) \in [0,1]^p : \sum_{s=1}^{p} q_s = 1\}. \quad (2)$$

For a player associated to commodity i, we write abusively $\mathcal{S}$ for $\mathcal{S}_{|\mathcal{P}_i|}$, i.e. the set of its mixed strategies.

We denote by $K = \mathcal{S}^N$ the space of mixed strategies.

Payoff functions d_i defined on pure strategies in equation (1) can be extended to functions $\overline{d}_i$ on the space of mixed strategies K as follows:

$$\begin{aligned}\overline{d}_i(q_1, ..., q_N) &= E[cost\ for\ i \mid \text{player } z \text{ employs strategy } q_z, 1 \leq z \leq N] \\ &= \textstyle\sum_{j_1,...,j_N} d_i(j_1, ..., j_N) \times \prod_{z=1}^{N} q_{z,j_z}\end{aligned} \quad (3)$$

where $(q_1, ..., q_N)$ is the set of mixed strategies played by the set of players and E denotes a conditiona expectation.

Definition 2. *The N-tuple of mixed strategies $(\tilde{q}_1, ..., \tilde{q}_N)$ is said to be a Nash equilibrium (in mixed strategies), if for each i, $1 \leq i \leq N$, we have:*

$$\overline{d}_i(\tilde{q}_1, ..., \tilde{q}_{i-1}, \tilde{q}_i, \tilde{q}_{i+1}, ..., \tilde{q}_N) \leq \overline{d}_i(\tilde{q}_1, ..., \tilde{q}_{i-1}, q, \tilde{q}_{i+1}, ..., \tilde{q}_N)\ \forall q \in \mathcal{S} \quad (4)$$

It is well known that every n-person game has at least one Nash equilibrium in mixed strategies [26].

We define $K^* = (\mathcal{S}^*)^N$ where $\mathcal{S}^* = \{q \in \mathcal{S} |\ q$ is a p-dimensional probability vector with 1 component unity$\}$ as the corners of the strategy space K. Clearly, K^* can be put in one-to-one correspondence with pure strategies. A N-tuple of actions $(\tilde{a}_1, ..., \tilde{a}_N)$ can be defined to be a pure Nash Equilibrium similarly.

Now the learning problem can be stated as follows: Assume that we play a stochastic repeated game with incomplete information. $q_i[k]$ is the strategy employed by the i^{th} player at instant k. Let $a_i[k]$ and $c_i[k]$ be the action selected and the payoff obtained by player i respectively at time k $(k = 0, 1, 2, \dots)$. Find a decentralized learning algorithm T_i, where $q_i[k+1] = T_i(q_i[k], a_i[k], c_i[k])$, such that $q_i[k] \to \tilde{q}_i$ as $k \to +\infty$ where $(\tilde{q}_1, ..., \tilde{q}_N)$ is a Nash equilibrium of the game.

5 Distributed Algorithm

We consider the following learning algorithm, already considered in [23, 28], and also called the Linear Reward-Inaction (L_{R-I}) algorithm.

Definition 3 (Considered Algorithm).

1. *At every time step, each player chooses an action according to its current Action Probability Vector (APV). Thus, the i^{th} player selects path $s = a_i(k)$ at instant k with probability $q_{i,s}(k)$.*
2. *Each player obtains a payoff based on the set of all actions. We note the reward to player i at time k: $c_i(k) = \ell_{a_i}(f(k))$.*
3. *Each player updates his APV according to the rule:*

$$q_i(k+1) = q_i(k) + b \times (1 - c_i(k)) \times (e_{a_i(k)} - q_i(k)), i = 1, ..., n, \quad (5)$$

where $0 < b < 1$ is a parameter and $e_{a_i(k)}$ is a unit vector of dimension m with $a_i(k)^{th}$ component unity.

It is easy to see that decisions made by players are completely decentralized, at each time step, player i only needs c_i and a_i, respectively its payoff and last action, to update his APV.

Notice that, componentwise, Equation (5) can be rewritten:

$$q_{i,s}(k+1) = \begin{cases} q_{i,s}(k) - b(1 - c_i(k))q_{i,s}(k) & \text{if } a_i \neq s \\ q_{i,s}(k) + b(1 - c_i(k))(1 - q_{i,s}(k)) & \text{if } a_i = s \end{cases} \quad (6)$$

Let $Q[k] = (q_1(k), ..., q_N(k)) \in K$ denote the state of the player team at instant k. Our interest is in the asymptotic behavior of $Q[k]$ and its convergence to a Nash Equilibrium. Clearly, under the learning algorithm specified by (5), $\{Q[k], k \geq 0\}$ is a Markov process.

Observe that this dynamic can also be put in the form

$$Q[k+1] = Q[k] + b \cdot G(Q[k], a[k], c[k]), \quad (7)$$

where $a[k] = (a_1(k), ..., a_N(k))$ denotes the actions selected by the player team at k and $c[k] = (c_1(k), ..., c_N(k))$ their resulting payoffs, for some function $G(., ., .)$ representing the updating specified by equation (5), that does not depend on b.

Consider the piecewise-constant interpolation of $Q[k], Q^b(.)$, defined by

$$Q^b(t) = Q[k], t \in [kb, (k+1)b], \tag{8}$$

where b is the parameter used in (5).

$Q^b(.)$ belongs to the space of all functions from $\mathbb{R}$ into K. These functions are right continuous and have left hand limits. Now consider the sequence $\{Q^b(.) : b > 0\}$. We are interested in the limit $Q(.)$ of this sequence as $b \to 0$.

The following is proved in [28]:

Proposition 1 ([28]). *The sequence of interpolated processes $\{Q^b(.)\}$ converges weakly, as $b \to 0$, to $Q(.)$, which is the (unique) solution of Cauchy problem*

$$\frac{dQ}{dt} = \phi(Q), Q(0) = Q_0 \tag{9}$$

where $Q_0 = Q^b(0) = Q[0]$, and $\phi : K \to K$ is given by

$$\phi(Q) = E[G(Q[k], a[k], c[k])|Q[k] = Q],$$

where G is the function in Equation (7).

Recall that a family of random variable $(Y_t)_{t \in \mathbb{R}}$ weakly converges to a random variable Y, if $E[h(X_t)]$ converges to $E[h(Y)]$ for each bounded and continuous function h. This is equivalent to convergence in distributions.

The proof of Proposition 1 in [28], that works for general (even with stochastic payoffs) games, is based on constructions from [24], in turn based on [32], i.e. on weak-convergence methods, non-constructive in several aspects, and does not provide error bounds.

It is actually possible to provide a bound on the error between $Q(t)$ and the expectation of $Q^b(.)$ in some cases.

Theorem 1. *Let $Q(.)$ be a process defined by an equation of type (7), and let $Q^b(.)$ be the corresponding piecewise-constant interpolation, given by (8).*

Assume that $E[G(Q[k], a[k], c[k])] = \phi(E[Q[k]])$ for some function ϕ of class $\mathcal{C}^1$.

Let $\epsilon(t)$ be the error in approximating the expectation of $Q^b(t)$ by $Q(t)$:

$$\epsilon(t) = ||E[Q^b(t)] - Q(t)||,$$

where $Q(.)$ is the (unique) solution of the Cauchy problem

$$\frac{dQ}{dt} = \phi(Q), Q(0) = Q_0, \tag{10}$$

where $Q_0 = Q^b(0) = Q[0]$.

We have

$$\epsilon(t) \leq Mb\frac{e^{\Lambda t} - 1}{2\Lambda},$$

for t of the form $t = kb$, where $\Lambda = \max_{i,\ell} ||\frac{\partial \phi}{\partial q_{i,\ell}}||$, and M is a bound on the norm of $Q''(t) = \frac{d\phi(Q(t))}{dt}$.

Proof. The general idea of the proof is to consider the dynamic (7), as an Euler discretization method of the ordinary differential equation (10), and then use some classical numerical analysis techniques to bound the error at time t.

Indeed, by hypothesis we have

$$\begin{aligned} E[Q[k+1]] &= E[Q[k]] + b \cdot E[G(Q[k], a[k], c[k])] \\ &= E[Q[k]] + b\phi(E[Q[k]]). \end{aligned}$$

Suppose that $\phi(.)$ is Λ-Lipschitz:

$$||\phi(x) - \phi(x')|| \leq \Lambda ||x - x'||,$$

for some positive Λ. From Taylor-Lagrange inequality, we can always take $\Lambda = \max_{i,\ell} ||\frac{\partial \phi}{\partial q_{i,\ell}}||$, if ϕ is of class $\mathcal{C}^1$.

We can write,

$$\begin{aligned} \epsilon((k+1)b) &= ||E[Q^b((k+1)b)] - Q((k+1)b)|| \\ &\leq ||E[Q^b((k+1)b)] - E[Q^b(kb)] - b\phi(Q(kb))|| \\ &\quad + ||E[Q^b(kb)] - Q(kb)|| + ||Q(kb) - Q((k+1)b) + b\phi(Q(kb))|| \\ &= ||b\phi(E[Q^b(kb)]) - b\phi(Q(kb))|| + \epsilon(kb) + ||b\phi(Q(kb)) - \int_{kb}^{(k+1)b} \phi(Q(t'))dt'|| \\ &\leq \Lambda b ||E[Q^b(kb)] - Q(kb)|| + \epsilon(kb) + e(kb) \\ &\leq (1 + \Lambda b)\epsilon(kb) + e(kb) \end{aligned}$$

where $e(kb) = ||b\phi(Q(kb)) - \int_{kb}^{(k+1)b} \phi(Q(t'))dt'||$.

From Taylor-Lagrange inequality, we know that $e(kb) \leq K = M\frac{b^2}{2}$, where M is a bound on the norm of $Q''(t) = \frac{d\phi(Q(t))}{dt}$.

From an easy recurrence on k, (sometimes called Discrete Gronwall's Lemma, see e.g. [9]), using inequality $\epsilon((k+1)b) \leq (1 + \Lambda b)\epsilon((k+1)b) + K$, we get that

$$\begin{aligned} \epsilon(kb) &\leq (1 + \Lambda b)^k \epsilon(0) + K\frac{(1+\Lambda b)^k - 1}{1 + \Lambda b - 1} \\ &\leq K\frac{e^{k\Lambda b} - 1}{\Lambda b} \\ &= Mb\frac{e^{k\Lambda b} - 1}{2\Lambda} \end{aligned}$$

using that $(1+u)^k \leq e^{ku}$ for all $u \geq 0$, and $\epsilon(0) = 0$. This completes the proof.

□

Using (6), we can rewrite $E[G(Q[k], a[k], c[k])]$ in the general case as follows.

$$\begin{aligned} E[G(Q[k], a[k], c[k])]_{i,s} &= q_{i,s}(1 - q_{i,s})(1 - E[c_i|Q(k), a_i = s]) \\ &\quad - \sum_{s' \neq s} q_{i,s'} q_{i,s}(1 - E[c_i|Q(k), a_i = s']) \\ &= q_{i,s}[\sum_{s' \neq s} q_{i,s'}(1 - E[c_i|Q(k), a_i = s]) \\ &\quad - \sum_{s' \neq s} q_{i,s'}(1 - E[c_i|Q(k), a_i = s']] \\ &= -q_{i,s} \sum_{s'} (E[c_i|Q(k), a_i = s] - q_{i,s'} E[c_i|Q(k), a_i = s']), \end{aligned} \quad (11)$$

using the fact that $1 - q_{i,s} = \sum_{s' \neq s} q_{i,s'}$.

Let $h_{i,s}$ be the expectation of the payoff for i if player i plays pure strategy s, and players $j \neq i$ play (mixed) strategy q_j. Formally,

$$h_{i,s}(q_1, \cdots, q_{i-1}, s, q_{i+1}, \cdots, q_n) = E[cost\ for\ i\ |Q(k), a_i = s].$$

Let $\overline{h_i}(Q)$ the mean value of $h_{i,s}$, in the sense that

$$\overline{h_i}(Q) = \sum_{s'} q_{i,s'} h_{i,s'}(Q).$$

We obtain from (11),

$$E[G(Q[k], a[k], c[k])]_{i,s} = -q_{i,s}(h_{i,s}(Q) - \overline{h_i}(Q)). \tag{12}$$

Hence, the dynamics given by Ordinary Differential Equation (9) is componentwise:

$$\frac{dq_{i,s}}{dt} = -q_{i,s}(h_{i,s}(Q) - \overline{h_i}(Q)). \tag{13}$$

This is a replicator equation, that is to say a well-known and studied dynamics in evolutionary game theory [20, 35]. In this context, $h_{i,s}(Q)$ is interpreted as player i's fitness for a given game, and $\overline{h_i}(Q)$ is the mean value of the expected fitness in the above sense.

In particular, solutions are known to satisfy the following theorem (sometimes called Evolutionary Game Theory Folk Theorem) [20, 28].

Theorem 2 (see e.g. [20, 28]). *The following are true for the solutions of the replicator equation* (13)*:*

- *All corners of space K are stationary points.*
- *All Nash equilibria are stationary points.*
- *All strict Nash equilibria are asymptotically stable.*
- *All stable stationary points are Nash equilibria.*

From this theorem, we can conclude that the dynamics (13), and hence the learning algorithm when b goes to 0, will never converge to a point in K which is not a Nash equilibrium.

However, for general games, there is no convergence in the general case [28].

We will now show that for linear Wardrop games, there is always convergence. It will then follow that the learning algorithm we are considering here converges towards Nash equilibria, i.e. solves the learning problem for linear Wardrop games.

First, we specialize the dynamics for our routing games. We have

$$\ell_{a_i}(f) = \sum_{e \in a_i} \ell_e(\lambda_e) = \sum_{e \in a_i} [\beta_e + \alpha_e w_i + \alpha_e \sum_{j \neq i} \mathbf{1}_{e \in a_j} w_j] \tag{14}$$

where $\mathbf{1}_{e \in a_j}$ is 1 whenever $e \in a_j$, 0 otherwise. Let us also introduce the following notation:

$$prob(e,Q)_i = \sum_{P \in \mathcal{P}_i} q_{i,P} \times \mathbf{1}_{e \in P} \tag{15}$$

which denotes the probability for player i to use edge e, for his given probability vector q_i.

Using expectation of utility for player i using path s, we get it as:

$$h_{i,s}(Q) = \sum_{e \in s} [\beta_e + \alpha_e w_i + \alpha_e \sum_{j \neq i} \sum_{P \in \mathcal{P}_j} q_{j,P} \times \mathbf{1}_{e \in P} w_j]$$

That we can also write (from (15))

$$h_{i,s}(Q) = \sum_{e \in s} [\beta_e + \alpha_e w_i + \alpha_e \sum_{j \neq i} prob(e,Q)_j w_j]$$

We claim the following.

Theorem 3 (Extension of Theorem 3.3 from [28]). *Suppose there is a non-negative function*

$$F : K \to \mathbb{R}$$

such that for some constants $w_i > 0$, for all i, s, Q,

$$\frac{\partial F}{\partial q_{i,s}}(Q) = w_i \times h_{i,s}(Q). \tag{16}$$

Then the learning algorithm, for any initial condition in $K - K^*$, always converges to a Nash Equilibrium.

Proof. We claim that $F(.)$ is monotone along trajectories.

We have:

$$\begin{aligned} \frac{dF(Q(t))}{dt} &= \sum_{i,s} \frac{\partial F}{\partial q_{i,s}} \frac{dq_{i,s}}{dt} \\ &= -\sum_{i,s} \frac{\partial F}{\partial q_{i,s}}(Q) q_{i,s} \sum_{s'} q_{i,s'} [h_{i,s}(Q) - h_{i,s'}(Q)] \\ &= -\sum_{i,s} w_i h_{i,s}(Q) q_{i,s} \sum_{s'} q_{i,s'} [h_{i,s}(Q) - h_{i,s'}(Q)] \\ &= -\sum_i w_i \sum_s \sum_{s'} q_{i,s} q_{i,s'} [h_{i,s}(Q)^2 - h_{i,s}(Q) h_{i,s'}(Q)] \\ &= -\sum_i w_i \sum_s \sum_{s' > s} q_{i,s} q_{i,s'} [h_{i,s}(Q) - h_{i,s'}(Q)]^2 \\ &\leq 0 \end{aligned} \tag{17}$$

Thus F is decreasing along the trajectories of the ODE and, due to the nature of the ODE (13), for initial conditions in K will be confined to K.

Hence from the Lyapunov Stability theorem (see e.g. [19] page 194), if we note Q^* an equilibrium point, we can define $L(Q) = F(Q) - F(Q^*)$ as a Lyapunov function of the game. Asymptotically, all trajectories will be in the set $K' = \{Q^* \in K : \frac{dF(Q^*)}{dt} = 0\}$.

From (17), we know that $\frac{dF(Q^*)}{dt} = 0$ implies $q_{i,s} q_{i,s'} [h_{i,s}(Q) - h_{i,s'}(Q)] = 0$ for all i, s, s'. Such a Q^* is, thus, a stationary point of the dynamics.

Since from Theorem 2, all stationary points that are not Nash equilibria are unstable, the theorem follows.

We claim that such a function exists for linear Wardrop games.

Proposition 2. For our definition we gave earlier of linear Wardrop games, the following function F satisfies the hypothesis of the previous theorem:

$$F(Q) = \sum_{e\in E}\Bigg[\ \beta_e\left(\sum_{j=1}^{N} w_j \times prob(e,Q)_j\right)+ \frac{\alpha_e}{2}\left(\sum_{j=1}^{N} w_j \times prob(e,Q)_j\right)^2 + \alpha_e\left(\sum_{j=1}^{N} w_j^2 \times prob(e,Q)_j \times (1-\frac{prob(e,Q)_j}{2})\right)\Bigg] \tag{18}$$

Notice that the hypothesis of affine cost functions is crucial here.

Proof. We use the fact that $F(Q)$ is of the form $\sum_{e\in E} expr(e,Q)$ in order to lighten the next few lines.

$\frac{\partial F}{\partial q_{i,s}}(Q) = \frac{\partial \sum_{e\in E} expr(e,Q)}{\partial q_{i,s}} = \sum_{e\in E} \frac{\partial expr(e,Q)}{\partial q_{i,s}}$ which can be rewritten as

$\frac{\partial F}{\partial q_{i,s}}(Q) = \sum_{e\in E} \frac{\partial expr(e,Q)}{\partial prob(e,Q)_i} \times \frac{\partial prob(e,Q)_i}{\partial q_{i,s}}$.

Note that, from (15), $\frac{\partial prob(e,Q)_i}{\partial q_{i,s}} = \mathbf{1}_{e\in s}$, we then get

$$\frac{\partial F}{\partial q_{i,s}}(Q) = \sum_{e\in E} \frac{\partial expr(e,Q)}{\partial prob(e,Q)_i} \times \mathbf{1}_{e\in s} = \sum_{e\in s} \frac{\partial expr(e,Q)}{\partial prob(e,Q)_i} \tag{19}$$

Let us now develop the derivative of each term of the sum and come back to (19) in the end, we have

$\frac{\partial expr(e,Q)}{\partial prob(e,Q)_i} = \beta_e\times w_i + \alpha_e \times w_i(\sum_{j=1}^{N} w_j \times prob(e,Q)_j) + \alpha_e(w_i^2(1-prob(e,Q)_i))$

$= \beta_e \times w_i + \alpha_e \times w_i(\sum_{j\neq i} w_j \times prob(e,Q)_j) + \alpha_e w_i^2$.

This finally leads to:

$\sum_{e\in s} \frac{\partial expr(e,Q)}{\partial prob(e,Q)_i} = \sum_{e\in s} \beta_e \times w_i + \alpha_e \times w_i(\sum_{j\neq i} w_j \times prob(e,Q)_j) + \alpha_e w_i^2$

$\frac{\partial F}{\partial q_{i,s}}(Q) = w_i \times h_{i,s}(Q)$

We showed that Equation (16) holds, which ends the proof and confirms that F is a good potential function for such a game.

□

Proposition 3. *Suppose for example that cost functions were quadratic :*

$$\ell_e(\lambda_e) = \alpha_e\lambda_e^2 + \beta_e\lambda_e + \gamma_e,$$

with $\alpha_e, \beta_e, \gamma_e \geq 0$*,* $\alpha_e \neq 0$*.*

There can not exist a function F *of class* $\mathcal{C}^2$ *that satisfies* (16) *for all* i*,* s*,* Q*, and general choice of weights* $(w_i)_i$*.*

Proof. By Schwartz theorem, we must have

$$\frac{\partial}{\partial q_{i',s'}}(\frac{\partial F}{\partial q_{i,s}}) = \frac{\partial}{\partial q_{i,s}}(\frac{\partial F}{\partial q_{i',s'}}),$$

and hence

$$W_i \frac{\partial h_{i,s}}{\partial q_{i',s'}} = W_{i'} \frac{\partial h_{i',s'}}{\partial q_{i,s}},$$

for all i, i', s, s', for some constants $W_i, W_{i'}$. It is easy to see that this doesn't hold for general choice of Q and weights $(w_i)_i$ in this case.

□

Coming back to our model (with affine costs), we obtain the following result:

Theorem 4. *For linear Wardrop games, for any initial condition in* $K - K^*$*, the considered learning algorithm converges to a (mixed) Nash equilibrium.*

6 Conclusion

In this paper we considered the classical Wardrop traffic model but where we introduced some specific dynamical aspects.

We considered an update algorithm proposed by [28] and we proved that the learning algorithm depicted is able to learn mixed Nash equilibria of the game, extending several results of [28].

To do so, we proved that the learning algorithm is asymptotically equivalent to an ordinary differential equation, which turns out to be a replicator equation. Using a folk theorem from evolutionary game theory, one knows that if the dynamics converges, it will be towards some Nash equilibria. We proved using a Lyapunov function argument that the dynamics converges in our considered settings.

We established some time bounds on the time required before convergence, based on the analysis of the dynamics, and numerical analysis arguments in some special cases.

We are also investigating the use of this dynamics over other games which are known to have some potential function, such as load balancing problems [22, 33].

We also believe that this paper yields a very nice example of distributed systems whose study is done through a macroscopic view of a set of distributed systems defined by microscopic rules: whereas the microscopic rules are quite simple, and based on local views, the macroscopic evolution computes global equilibria of the system.

We also intend to pursue our investigations on the computational properties of distributed systems through similar macroscopic continuous time dynamical system views.

References

1. E. Altman, Y. Hayel, and H. Kameda. Evolutionary Dynamics and Potential Games in Non-Cooperative Routing. In *Wireless Networks: Communication, Cooperation and Competition (WNC3 2007)*, 2007.
2. Petra Berenbrink, Tom Friedetzky, Leslie Ann Goldberg, Paul Goldberg, Zengjian Hu, and Russell Martin. Distributed Selfish Load Balancing. In *SODA '06: Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm*, pages 354–363, New York, NY, USA, 2006. ACM.
3. Olivier Bournez. *Modèles Continus. Calculs. Algorithmique Distribuée.* Hdr, Institut National Polytechnique de Lorraine, 7 Décembre 2006.
4. Olivier Bournez and Manuel L. Campagnolo. *New Computational Paradigms. Changing Conceptions of What is Computable*, chapter A Survey on Continuous Time Computations, pages 383–423. Springer-Verlag, New York, 2008.
5. Olivier Bournez, Manuel L. Campagnolo, Daniel S. Graça, and Emmanuel Hainry. Polynomial Differential Equations Compute All Real Computable Functions on Computable Compact Intervals. *Journal of Complexity*, 23(3):317–335, June 2007.
6. Olivier Bournez and Emmanuel Hainry. Recursive Analysis Characterized as a Class of Real Recursive Functions. *fundinform*, 74(4):409–433, 2006.
7. Richard Cole, Yevgeniy Dodis, and Tim Roughgarden. How much can taxes help selfish routing? In *Proceedings of the 4th ACM Conference on Electronic Commerce (EC-03)*, pages 98–107, New York, June 9–12 2003. ACM Press.
8. R. Cominetti, J.R. Correa, and N.E. Stier-Moses. Network Games with Atomic Players. *Automata, Languages and Programming: Proceedings of the 33rd International Colloquium, Venice, Italy*, 4051:525–536, 2006.
9. Jean-Pierre Demailly. *Analyse Numérique et Equations Différentielles.* Presses Universitaires de Grenoble, 1991.
10. E. Even-Dar, A. Kesselman, and Y. Mansour. Convergence Time to Nash Equilibria. *30th International Conference on Automata, Languages and Programming (ICALP)*, pages 502–513, 2003.
11. Eyal Even-Dar, Alexander Kesselman, and Yishay Mansour. Convergence Time to Nash equilibrium in Load Balancing. *ACM Transactions on Algorithms*, 3(3), 2007.
12. Eyal Even-Dar and Yishay Mansour. Fast Convergence of Selfish Rerouting. In *SODA '05: Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 772–781, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics.
13. S. Fischer, H. Räcke, and B. Vöcking. Fast Convergence to Wardrop Equilibria by Adaptive Sampling Methods. *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 653–662, 2006.
14. S. Fischer and B. Vocking. On the Evolution of Selfish Routing. *Algorithms–ESA 2004: 12th Annual European Symposium, Bergen, Norway, September 14-17, 2004, Proceedings*, 2004.
15. S. Fischer and B. Vöcking. Adaptive Routing with Stale Information. *Proceedings of the twenty-fourth annual ACM SIGACT-SIGOPS symposium on Principles of distributed computing*, pages 276–283, 2005.
16. L. Fleischer. Linear Tolls Suffice: New Bounds and Algorithms For Tolls in Single Source Networks. *Theoretical Computer Science*, 348(2-3):217–225, 2005.

17. Paul W. Goldberg. Bounds for the Convergence Rate of Randomized Local Search in a Multiplayer Load-Balancing Game. In *PODC '04: Proceedings of the twenty-third annual ACM symposium on Principles of distributed computing*, pages 131–140, New York, NY, USA, 2004. ACM.
18. C. Harris. On the Rate of Convergence of Continuous-Time Fictitious Play. *Games and Economic Behavior*, 22(2):238–259, 1998.
19. Morris W. Hirsch, Stephen Smale, and Robert Devaney. *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Elsevier Academic Press, 2003.
20. J. Hofbauer and K. Sigmund. Evolutionary Game Dynamics. *Bulletin of the American Mathematical Society*, 4:479–519, 2003.
21. J. Hofbauer and S. Sorin. Best Response Dynamics for Continuous Zero-Sum Games. *Discrete and Continuous Dynamical Systems-Series B*, 6(1), 2006.
22. E. Koutsoupias and C. Papadimitriou. Worst-case Equilibria. In *STACS99*, pages 404–413, Trier, Germany, 4–6March 1999.
23. M.A.L. Thathachar K.S. Narendra. Learning Automata: An Introduction. *Englewood Cliffs: Prentice Hall*, 1989.
24. H. J. Kushner. *Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory*. Cambridge, MA: MIT Press, 1984.
25. L. Libman and A. Orda. Atomic Resource Sharing in Noncooperative Networks. *Telecommunication Systems*, 17(4):385–409, 2001.
26. John F. Nash. Equilibrium Points in n-person Games. *Proc. of the National Academy of Sciences*, 36:48–49, 1950.
27. A. Orda, R. Rom, and N. Shimkin. Competitive Routing in Multi-user Communication Networks. *IEEE/ACM Transactions on Networking (TON)*, 1(5):510–521, 1993.
28. M.A.L. Thathachar P.S. Sastry, V.V. Phansalkar. Decentralized Learning of Nash Equilibria in Multi-Person Stochastic Games With Incomplete Information. *IEEE transactions on system, man, and cybernetics*, 24(5), 1994.
29. T. Roughgarden. How unfair is optimal routing? *Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 203–204, 2002.
30. Tim Roughgarden and Éva Tardos. How bad is selfish routing? *Journal of the ACM*, 49(2):236–259, March 2002.
31. L. Olbrich S. Fischer and B. Vöcking. Approximating Wardrop Equilibria with Finitely Many Agents. *DISC07*, 2007.
32. D.W. Stroock and SRS Varadhan. *Multidimensional Diffusion Processes*. Springer, 1979.
33. Berthold Vöcking. *Algorithmic Game Theory*, chapter Selfish Load Balancing. Cambridge University Press, 2007.
34. J. Wardrop. Some Theoretical Aspects of Road Traffic Research. *Proceedings of the Institution of Civil Engineers, Part II*, 1(36):352–362, 1952.
35. Jörgen W. Weibull. *Evolutionary Game Theory*. The MIT Press, 1995.