Convergence of Mirror Descent Dynamics in the Routing Game

Walid Krichene Syrine Krichene Alexandre Bayen

Abstract— **We consider a routing game played on a graph, in which different populations of drivers (or packet routers) iteratively make routing decisions and seek to minimize their delays. The Nash equilibria of the game are known to be the minimizers of a convex potential function, over the product of simplexes which represent the strategy spaces of the populations. We consider a class of population dynamics which only uses local loss information, and which can be interpreted as a mirror descent on the convex potential. We show that for vanishing, non-summable learning rates, mirror descent dynamics are guaranteed to converge to the set of Nash equilibria, and derive convergence rates as a function of the learning rate sequences of each population, and illustrate these results on a numerical example.**

I. INTRODUCTION

Routing games form a class of potential games used to model the interaction of players on a network. It has been extensively studied in transportation settings since the seminal work of Beckman [2], see for example [13] and the references therein. Routing games are also used to model congestion in communication networks [10], as well as job scheduling [12]. The one-shot routing game has played an important role in understanding the inefficiencies of networks (for example, the Braes paradox [5], and the price of anarchy [14]), and developing strategies to alleviate this inefficiency, either through network design or pricing [10].

Nash equilibria of the one-shot routing game are known to be the solutions to a convex problem: Rosenthal [11] proposed a potential function and proved that the set of Nash equilibria is exactly the set of minimizers of this potential.

Beyond characterizing the equilibria of the one-shot game, many studies have been concerned with the *dynamics of routing*, both in continuous-time [15], [7] and in discrete time [3], [8]. Modeling the dynamics of the game can be very informative, as it allows us to study stability and convergence rates to the equilibria, and is essential in designing control schemes.

In [15], Sandholm studies continuous-time population dynamics for potential games, which include routing games, and shows that if a positive correlation condition is satisfied between the dynamics vector field and the potential gradient vector field, the population strategies converge to the set of Nash equilibria. Fischer and Vöcking [7] study one particular example of dynamics, given by the replicator equation, a popular model in evolutionary game theory [16]. They show that replicator dynamics for the routing game are guaranteed to converge to the set of stationary points, a superset of Nash equilibria. In [3], Blum et al. consider a discrete-time setting, and study online learning dynamics. They show that if the regret of each population is sublinear, then the time-averaged strategies are guaranteed to converge to the set of Nash equilibria, and they give convergence rates. This result is very powerful, as it applies to a large class of algorithms. However, due to its generality, it only guarantees convergence in the sense of Cesàro means (in other words, convergence of the time-averages), not convergence of the sequence itself. In [8], Krichene et. al consider a sub-class of dynamics with sub-linear regret, which can be viewed as a stochastic approximation of the replicator dynamics. Under this restriction, the sequence of strategies is shown to asymptotically converge, however no convergence rate is known for this class of dynamics.

In this paper, we consider a general class of dynamics which can be viewed as a mirror descent iteration on the Rosenthal potential function. This class is described in detail in Section II. Algorithms in this class are known to have sub-linear regret, as discussed in Section III, which proves convergence in the sense of Cesàro means. We additionally show that under a mild assumption on the learning rates, the sequence is, in fact, guaranteed to converge, and we derive convergence rates. These results hold even for heterogeneous dynamics, i.e. dynamics such that each population obeys a different update equation. This class of algorithms includes, in particular, the Hedge algorithm, also known as the exponential weights algorithm, perhaps the most widely studied algorithm in the online learning literature, see for example [6]. Finally, we give a few numerical examples to illustrate these convergence results in Section IV, and we compare the empirical convergence rates to the theoretical bounds of Section III.

A. *The routing game*

The routing game is given by a directed graph $G=(V,E)$ with vertex set V and edge set $E \subset V \times V$, and a finite number of populations $\{P_k\}_{k\in\{1,\dots,K\}}$. Population P_k is characterized by a source vertex $s_k \in V$ and a destination vertex $d_k \in V$, and represents a set of players (drivers, or packet routers) commuting or sending network traffic from s_k to d_k. Formally, a population is a measurable set $P_k = (S_k, \mathscr{S}_k, m_k)$, where S_k is the set of players, $\mathscr{S}_k$ is a σ-algebra of measurable

Walid Krichene is with the department of Electrical Engineering and Computer Sciences at the University of California, Berkeley, CA 94720. walid@eecs.berkeley.edu

Syrine Krichene is with the ENSIMAG, St-Martain d'Hères, France. syrine.krichene@ensimag.grenoble-inp.fr

Alexandre Bayen is with the department of Electrical Engineering and Computer Sciences and the department of Civil and Environmental Engineering at the University of California, Berkeley, CA 94720. bayen@berkeley.edu

subsets of S_k, and m_k is a finite measure[1]. The action set of every player in S_k is the set of simple paths connecting the source s_k to the destination d_k, and will be denoted $\mathscr{P}_k$. In other words, each player of each population chooses a route (or a distribution over routes) from the origin to the destination. Let $\pi^k : S_k \to \Delta^{\mathscr{P}_k}$ be a measurable function, which determines, for each player $s \in S_k$, the strategy $\pi^k(s)$ of the player, an element of $\Delta^{\mathscr{P}_k}$ the simplex on the set of paths $\mathscr{P}_k$. We refer to π^k as the strategy profile of population P_k. The strategy profiles determine the mass distribution of the players over paths, defined as

$$\forall k,\ x^k = \int_{S_k} \pi^k(s) m_k(ds) \in m_k(S_k)\Delta^{\mathscr{P}_k}, \tag{1}$$

as well as the total mass of players on each edge, defined as

$$\forall e \in E,\ \phi_e = \sum_{k=1}^{K} \sum_{p \in \mathscr{P}_k : e \in p} x_p^k.$$

We observe that ϕ_e is a simple linear function of the mass distribution x, so we can write in vector form $\phi = Mx$, where

- $\phi \in \mathbb{R}^E$ is the vector of edge masses,
- $x \in m_1(S_1)\Delta^{\mathscr{P}_1} \times \cdots \times m_K(S_K)\Delta^{\mathscr{P}_K}$ is the vector of strategy profiles of all populations,
- $M \in \mathbb{R}^{E \times (\mathscr{P}_1 \cup \cdots \cup \mathscr{P}_K)}$ is an incidence matrix, such that for all $e \in E$, all $k \in \{1, \dots, K\}$ and all $p \in \mathscr{P}_k$, $M_{e,p} = 1$ if $e \in p$ and 0 otherwise.

The set of mass distributions $m_1(S_1)\Delta^{\mathscr{P}_1} \times \cdots \times m_K(S_K)\Delta^{\mathscr{P}_K}$ will be denoted Δ for convenience.

The edge masses determine the loss of each player: the edge loss on $e \in E$ is given by a positive, Lipschitz-continuous, increasing function of ϕ_e, denoted $c_e(\phi_e)$, and the loss on a path is simply the sum of edge losses along that path. The loss on a path $p \in \mathscr{P}_k$ induced by a mass distribution $x = (x^1, \dots, x^K)$ will be denoted

$$\ell_p^k(x) = \sum_{e \in p} c_e((Mx)_e) = M_p^T (c_e((Mx)_e))_{e \in E}$$

Then the expected loss of a player $s \in S_k$ with strategy $\pi^k(s)$ is $\mathbb{E}_{p \sim \pi^k(s)}[\ell_p^k(x)] = \sum_{p \in \mathscr{P}_k} \pi_p^k(s) \ell_p^k(x)$, which will also be denoted $\left\langle \pi^k(s), \ell^k(x) \right\rangle$, where we use $\langle \cdot, \cdot \rangle$ to denote the Euclidean inner product.

The graph $G = (V, E)$, the set of populations $\{P_k\}_{k \in \{1,\dots,K\}}$ with their sources and destinations $\{(s_k, d_k)\}_{k \in \{1,\dots,K\}}$, and the edge loss functions $\{c_e\}_{e \in E}$ entirely determine the game.

B. Nash equilibria and the Rosenthal potential function

The set of Nash equilibria of the game is defined in terms of the mass distribution x, as follows.

Definition 1 (Nash equilibria of the one-shot routing game)**.** *A mass distribution $x \in \Delta$ is a Nash equilibrium if for all k, and for all $p \in support(x^k)$, $\ell_p^k(x) = \min_{q \in \mathscr{P}_k} \ell_q^k(x)$. The set of Nash equilibria will be denoted $\mathscr{N}$.*

[1]For example, one could take S_k to be a bounded interval, and m_k to be the Lebesgue measure on S_k.

In other words, a distribution is a Nash equilibrium if no path with positive mass is suboptimal under that distribution. Equivalently,

$$x \in \mathscr{N} \Leftrightarrow \forall y \in \Delta, \langle x - y, \ell(x) \rangle \leq 0. \tag{2}$$

The Rosenthal potential was first defined for finite player routing games, and later generalized to games with a continuum of players, see for example the analysis of Sandholm [15]. The potential function can be defined on the product of simplexes Δ as follows

$$f(x) = \sum_{e \in E} \int_0^{(Mx)_e} c_e(u) du$$

It can be viewed as the composition of the function $\bar{f} : \phi \mapsto \sum_e \int_0^{\phi_e} c_e(u) du$, and the linear function $x \mapsto Mx$. The function $\bar{f}$ has gradient $\nabla \bar{f}(\phi) = (c_e(\phi_e))_{e \in E}$, thus it is convex (the edge losses are increasing by assumption). Therefore f is convex (composition of a convex function and a linear function) and has gradient

$$\nabla_x f(x) = M^T \nabla \bar{f}(Mx) = M^T (c_e((Mx)_e))_{e \in E} = \ell(x)$$

In other words, the gradient of the potential f is exactly the loss vector field ℓ. This property is essential in the analysis.

As a first consequence, by first-order optimality for differentiable convex functions, $x \in \arg\min_{x \in \Delta} f(x)$ if and only if $\forall y \in \Delta$, $\langle x - y, \nabla f(x) \rangle \leq 0$ (in other words, the negative gradient $-\nabla f(x)$ defines a supporting hyperplane to the feasible set Δ, see 4.2.3 in [4] for a simple proof), which is exactly the characterization of Nash equilibria in equation (2). It follows that Nash equilibria are the minimizers of f over Δ. We observe that the minimizer is not unique in general, as the function f may be weakly convex.

C. Online learning, sublinear regret and Cesàro convergence

We now assume that the players make routing decisions at discrete time instants $t \in \mathbb{N}$. At each iteration t, each player in S_k chooses a distribution $\pi^{k^{(t)}}(s)$, which results in a mass distribution $x^{(t)}$, as defined in equation (1). This determines the losses of the players, and the loss vector $\ell^k(x^{(t)})$ is then revealed to the players of population P_k. When making a routing decision at time t, a player in P_k only has access to the history of losses $\ell^k(x^{(\tau)})$ and the mass distributions $x^{k^{(\tau)}}$ for her population, up to $\tau = t - 1$. In particular, the players do not know the underlying structure of the edge losses or path losses. The process is summarized in Algorithm 1.

Algorithm 1 Online learning dynamics for the routing game.

for $t \in \mathbb{N}$ **do**

For all k, every player $s \in S_k$ chooses a distribution $\pi^{k^{(t)}}(s) \in \Delta^{\mathscr{P}_k}$, as a function of the history of observed losses $\left(\ell^k(x^{(\tau)})\right)_{\tau \leq t-1}$ and mass distributions $\left(x^{k^{(\tau)}}\right)_{\tau \leq t-1}$.

The strategy profiles $\left(\pi^{k^{(t)}}\right)_{k \in \{1,\dots,K\}}$ determine the mass distributions $\left(x^{k^{(t)}}\right)_{k \in \{1,\dots,K\}}$ and the losses $\left(\ell^k(x^{(t)})\right)_{k \in \{1,\dots,K\}}$.

For each k, the loss vector $\ell^k(x^{(t)})$ is revealed to players in P_k.

Given this model of online learning, we can define, for each player $s \in S_k$, the discounted cumulative regret of the player, which provides a natural measure of performance for sequential decision problems, see for example [6].

Definition 2 (Discounted cumulative regret)**.** *Let $(\gamma_t)_{t\in\mathbb{N}}$ be a sequence of positive decreasing discount factors. Given a sequence of losses $(\ell^{k^{(t)}})_{t\in\mathbb{N}}$, and a sequence of routing decisions $(\pi^{k^{(t)}}(s))_{t\in\mathbb{N}}$, the regret of the player s with respect to a strategy $\nu \in \Delta^{\mathscr{P}_k}$ is defined as*

$$R_s^{(t)}(\nu) = \sum_{\tau=1}^{t} \gamma_\tau \left\langle \pi^{k^{(\tau)}}(s) - \nu, \ell^{k^{(\tau)}} \right\rangle$$

The discounted regret compares the discounted cumulative loss of the player $\sum_{\tau=1}^{t} \gamma_\tau \left\langle \pi^{k^{(\tau)}}(s), \ell^{k^{(\tau)}} \right\rangle$ to the discounted cumulative loss of the fixed strategy ν, $\left\langle \nu, \sum_{\tau=1}^{t} \gamma_\tau \ell^{k^{(\tau)}} \right\rangle$. We also define the population regret with respect to mass distribution $y \in m_k(S_k)\Delta^{\mathscr{P}_k}$ as the integral

$$R^{k^{(t)}}(y) = \int_{S_k} R_s^{(t)} \left(\frac{y}{m_k(S_k)} \right) m_k(ds) = \sum_{\tau=1}^{t} \gamma_\tau \left\langle x^{k^{(\tau)}} - y, \ell^{k^{(\tau)}} \right\rangle$$

Finally, the discounted regret is said to be sublinear if

$$\limsup_{t\to\infty} \sup_{y \in m_k(S_k)\Delta^{\mathscr{P}_k}} \frac{R^{k^{(t)}}(y)}{\sum_{\tau=1}^{t} \gamma_\tau} \leq 0.$$

(a sufficient condition is that the regret of m_k-almost all players in S_k is sublinear). In Lemma 1, we recall the fact that if all populations have sublinear regrets, then the sequence of mass distributions $x^{(t)}$ converges to the set of Nash equilibria in the sense of Cesàro means, as defined below:

Definition 3 (Convergence in the sense of Cesàro)**.** *A sequence $x^{(t)}$ of elements of Δ is said to converge to a set $L \subset \Delta$ in the sense of Cesàro with weights $(\gamma_t)_t$ (a positive non-increasing sequence), if*

$$\lim_{t\to\infty} d \left(\frac{\sum_{\tau=1}^{t} \gamma_\tau x^{(\tau)}}{\sum_{\tau=1}^{t} \gamma_\tau}, L \right) = 0$$

where $d(\cdot, L)$ is the Euclidean distance to the set L. We write $x^{(t)} \xrightarrow{(\gamma_t)} L$.

Lemma 1. *Suppose that for all k, the discounted cumulative regret $R^{k^{(t)}}$ is sublinear. Then $x^{(t)} \xrightarrow{(\gamma_t)} \mathscr{N}$.*

Proof. Since $\mathscr{N}$ is the set of minimizers of the potential f over Δ, and f is continuous and Δ is compact, it suffices to show that $f\left(\frac{\sum_{\tau=1}^{t} \gamma_\tau x^{(\tau)}}{\sum_{\tau=1}^{t} \gamma_\tau}\right) \to f^*$, the minimum of f over Δ.

By convexity of f, and the fact that $\nabla f(x) = \ell(x)$, we have for any $y \in \mathscr{N}$

$$\begin{aligned}
f\left(\frac{\sum_{\tau=1}^{t} \gamma_\tau x^{(\tau)}}{\sum_{\tau=1}^{t} \gamma_\tau}\right) - f^* &\leq \frac{\sum_{\tau=1}^{t} \gamma_\tau f(x^{(\tau)})}{\sum_{\tau=1}^{t} \gamma_\tau} - f(y) \\
&\leq \frac{\sum_{\tau=1}^{t} \gamma_\tau \left\langle \nabla f(x^{(\tau)}), x^{(t)} - y \right\rangle}{\sum_{\tau=1}^{t} \gamma_\tau} \\
&= \frac{1}{\sum_{\tau=1}^{t} \gamma_\tau} \sum_{k=1}^{K} \sum_{\tau=1}^{t} \gamma_\tau \left\langle \ell^k(x^{(\tau)}), x^{k^{(\tau)}} - y^k \right\rangle \\
&\leq \frac{1}{\sum_{\tau=1}^{t} \gamma_\tau} \sum_{k=1}^{K} R^{k^{(\tau)}}(y^k) \qquad (3)
\end{aligned}$$

which converges to zero if the population regrets are sublinear. ☐

We observe that by inequality (3), convergence rates of the population regrets directly translate into a convergence rate of the Cesàro means to $\mathscr{N}$. However general, this result remains limited in that it does not guarantee convergence of the actual sequence $x^{(t)}$ of mass distributions. In order to guarantee its convergence, we need to further restrict the class of dynamics, as discussed in the next section.

II. Mirror descent dynamics

Mirror descent is a general method for solving constrained convex optimization, proposed by Nemirovski and Yudin [9]. It can be interpreted, as observed by Beck and Teboule [1], as a gradient descent algorithm using a non-Euclidean projection. Consider the problem

$$\text{minimize}_{x\in\mathscr{X}} f(x)$$

where $\mathscr{X} \subseteq \mathbb{R}^n$ is a convex compact set, and f is convex subdifferentiable. The Mirror Descent method with Bregman divergence D_ψ and learning rates $(\eta_t)_t$ (a positive non-increasing sequence) can be summarized in Algorithm 2.

Algorithm 2 Mirror descent algorithm with Bregman divergence D_ψ and learning rates $(\eta_t)_t$.

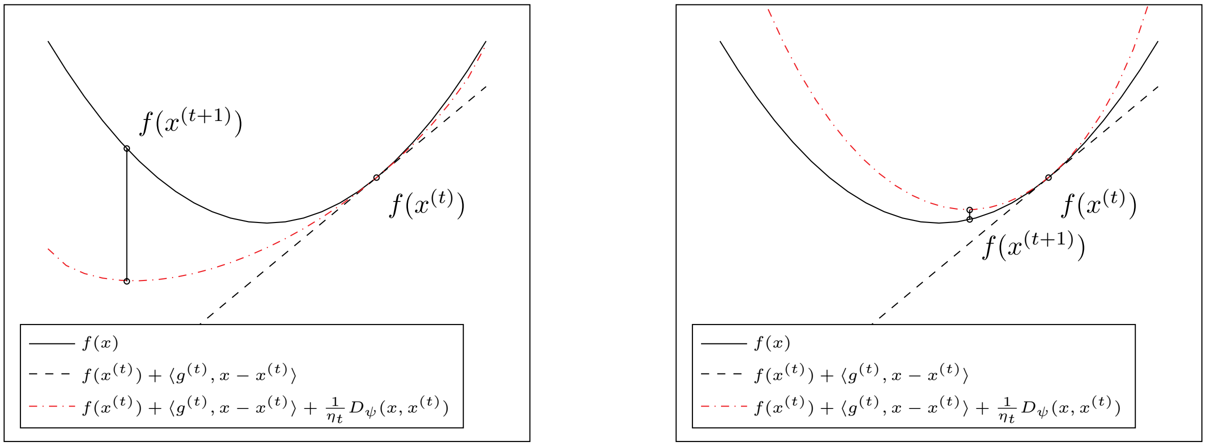
for $t \in \mathbb{N}$ **do**
  Query $g^{(t)} \in \partial f(x^{(t)})$, the sub differential of f at $x^{(t)}$.
  Update

$$x^{(t+1)} = \underset{x\in\mathscr{X}}{\arg\min}\, f(x^{(t)}) + \left\langle g^{(t)}, x - x^{(t)} \right\rangle + \frac{1}{\eta_t} D_\psi(x, x^{(t)}) \qquad (4)$$

Here, D_ψ is a Bregman divergence induced by a strongly convex function ψ, defined as follows: for all $x, y \in \mathscr{X}$,

$$D_\psi(x,y) = \psi(x) - \psi(y) - \langle \nabla\psi(y), x - y \rangle$$

The strong convexity assumption of ψ is equivalent to the existence of a positive constant ℓ_ψ such that $D_\psi(x,y) \geq \frac{\ell_\psi}{2}\|x-y\|^2$, where $\|\cdot\|$ is chosen to be the Euclidean norm. Note that by equivalence of norms, the choice of the norm does not affect the strong convexity of ψ, although it may affect the strong convexity constant ℓ_ψ. It follows that the Bregman divergence is positive definite, i.e. $D_\psi(x,y) \geq 0$ for all $x, y \in \mathscr{X}$, with equality if and only if $x = y$. When



(a) Large learning rate η_t.

(b) Small learning rate η_t.

Fig. 1: Illustration of a mirror descent update, using the KL-divergence as a Bregman divergence. The linear approximation of the function around the current iterate $x^{(t)}$ is given by $f(x^{(t)}) + \left\langle g^{(t)}, x - x^{(t)} \right\rangle$. The mirror descent update minimizes the linear approximation plus a Bregman divergence term $D_\psi(x, x^{(t)})$, which penalizes deviation from the current iterate $x^{(t)}$. The learning rate parameter η_t affects the relative importance of both terms, and the shape of the Bregman approximation to be minimized (dot-dashed, in red). Since the Bregman divergence is lower-bounded by a quadratic (by strong convexity of ψ), a smaller learning rate η_t results in a Bregman approximation with stronger curvature (b).

$\psi(x) = \frac{1}{2}\|x\|_2^2$, the Bregman divergence is $D_\psi(x,y) = \frac{1}{2}\|x - y\|_2^2$, and the mirror descent method reduces to projected subgradient descent; in this sense, a Bregman divergence is a generalization of the Euclidean distance, although in general, it is not symmetric and does not satisfy the triangle inequality. Figure 1 gives a geometric interpretation of the mirror descent update (4).

We now define population dynamics for the routing game, inspired from the mirror descent method for convex optimization. Suppose that each population P_k uses a Bregman divergence D_{ψ_k} and a sequence of learning rates $\left(\eta_t^k\right)_t$, and performs, at each iteration t, the following update, given the previous vector of losses $\ell^k(x^{(t)}) \in \mathbb{R}^{\mathscr{P}_k}$,

$$x^{k^{(t+1)}} = \underset{x^k \in m_k(S_k)\Delta^{\mathscr{P}_k}}{\arg\min} \left\langle \ell^k(x^{(t)}), x^k - x^{k^{(t)}} \right\rangle + \frac{1}{\eta_t^k} D_{\psi_k}(x^k, x^{k^{(t)}})$$

To simplify the analysis, we defined the dynamics in terms of the population mass distribution $x^{(t)}$, and not in terms of the individual player strategies $\pi^{k^{(t)}}(s)$, although it can be generalized to updates which depend on the individual players. The mirror descent dynamics can be summarized in Algorithm 3.

Algorithm 3 Mirror descent dynamics for the routing game.

for $t \in \mathbb{N}$ **do**
For each k, the loss vector $\ell^k(x^{(t)})$ is revealed to population P_k.
For each k, the mass distribution x^k is updated using the MD iteration with divergence D_{ψ_k} and learning rate $\left(\eta_t^k\right)$

$$x^{k^{(t+1)}} = \underset{x^k \in m_k(S_k)\Delta^{\mathscr{P}_k}}{\arg\min} \left\langle \ell^k(x^{(t)}), x^k - x^{k^{(t)}} \right\rangle + \frac{1}{\eta_t^k} D_{\psi_k}(x^k, x^{k^{(t)}})$$

III. Convergence rates of mirror descent dynamics

We now prove the convergence of mass distributions $x^{(t)}$ to the set of equilibria $\mathscr{N}$ under mirror descent dynamics, both in terms of Cesàro convergence and convergence of the actual sequence $x^{(t)}$.

A. A discounted regret bound

We first derive a general bound on the discounted population regret under mirror descent dynamics. A crucial observation is that since the loss function ℓ is continuous on the compact set Δ, it is bounded.

Lemma 2 (Discounted regret bound)**.** *Suppose that the mass distribution $(x^{k^{(t)}})_t$ of population P_k obeys the mirror descent dynamics defined in Algorithm 3, with Bregman divergence D_{ψ_k} (with strong convexity constant ℓ_{ψ_k}), and learning rates $(\eta_t^k)_t$. Let (γ_t) be a sequence of discount factors. Then for all t and all $x^k \in m_k(S_k)\Delta^{\mathscr{P}_k}$, the discounted regret with respect to x^k is bounded as follows*

$$R^{k^{(t)}}(x^k) \le \frac{L_k^2}{2\ell_{\psi_k}} \sum_{\tau=1}^{t} \eta_\tau^k \gamma_\tau + \frac{\gamma_1}{\eta_1^k} D_{\psi_k}(x^k, x^{k^{(1)}}) + \sum_{\tau=2}^{t} D_{\psi_k}(x^k, x^{k^{(\tau)}}) \left(\frac{\gamma_\tau}{\eta_\tau^k} - \frac{\gamma_{\tau-1}}{\eta_{\tau-1}^k} \right) \quad (5)$$

where L_k is an upper bound on the norm of the loss function $\|\ell^k\|$.

Proof. We seek to bound the sum $\sum_{\tau=1}^{t} \gamma_\tau \left\langle \ell^{k^{(\tau)}}, x^{k^{(\tau)}} - x^k \right\rangle$. We first decompose

$$\left\langle \ell^{k^{(\tau)}}, x^{k^{(\tau)}} - x^k \right\rangle = \left\langle \ell^{k^{(\tau)}}, x^{k^{(\tau+1)}} - x^k \right\rangle + \left\langle \ell^{k^{(\tau)}}, x^{k^{(\tau)}} - x^{k^{(\tau+1)}} \right\rangle$$

and bound each term separately.

For the first term, we have, by definition of the mirror descent update (4), $x^{k(\tau+1)}$ is the minimizer over $m_k(S_k)\Delta^{\mathscr{P}_k}$ of the convex function $h^{(\tau)}(x^k) = \left\langle \ell^{k(\tau)}, x^k - x^{k(\tau)} \right\rangle + \frac{1}{\eta_\tau^k} D_{\psi_k}(x^k, x^{k(\tau)})$. The gradient of this function is

$$\nabla h(x^k) = \ell^{k(\tau)} + \frac{1}{\eta_\tau^k}\left(\nabla\psi_k(x^k) - \nabla\psi_k(x^{k(\tau)})\right)$$

Thus the optimality conditions applied to $x^{k(\tau+1)}$ require that for any $x^k \in m_k(S_k)\Delta^{\mathscr{P}_k}$, $\left\langle \nabla h(x^{k(\tau+1)}), x^{k(\tau+1)} - x^k \right\rangle \le 0$, thus

$$\left\langle \ell^{k(\tau)} + \frac{1}{\eta_\tau^k}\left(\nabla\psi_k(x^{k(\tau+1)}) - \nabla\psi_k(x^{k(\tau)})\right), x^{k(\tau+1)} - x^k \right\rangle \le 0$$

then observing that

$$\left\langle \left(\nabla\psi_k(x^{k(\tau+1)}) - \nabla\psi_k(x^{k(\tau)})\right), x^{k(\tau+1)} - x^k \right\rangle = D_{\psi_k}(x^k, x^{k(\tau)}) - D_{\psi_k}(x^k, x^{k(\tau+1)}) - D_{\psi_k}(x^{k(\tau+1)}, x^{k(\tau)})$$

and using strong convexity of D_{ψ_k}, we obtain a bound on the first term

$$\left\langle \ell^{k(\tau)}, x^{k(\tau+1)} - x^k \right\rangle \le \frac{1}{\eta_\tau^k}\Big(D_{\psi_k}(x^k, x^{k(\tau)}) - D_{\psi_k}(x^k, x^{k(\tau+1)}) - \frac{\ell_{\psi_k}}{2}\|x^{k(\tau+1)} - x^{k(\tau)}\|^2\Big) \quad (6)$$

For the second term, we can use Young's inequality to obtain

$$\left\langle \ell^{k(\tau)}, x^{k(\tau)} - x^{k(\tau+1)} \right\rangle \le \frac{\eta_\tau^k}{2\ell_{\psi_k}}\|\ell^{k(\tau)}\|^2 + \frac{\ell_{\psi_k}}{2\eta_\tau^k}\|x^{k(\tau)} - x^{k(\tau+1)}\|^2 \quad (7)$$

Combining inequalities (6) and (7), and summing over τ, we have

$$\sum_{\tau=1}^t \gamma_\tau \left\langle \ell^{k(\tau)}, x^{k(\tau)} - x^k \right\rangle \le \frac{L_k^2}{2\ell_{\psi_k}}\sum_{\tau=1}^t \eta_\tau^k \gamma_\tau + \sum_{\tau=1}^t \frac{\gamma_\tau}{\eta_\tau^k}\left(D_{\psi_k}(x^k, x^{k(\tau)}) - D_{\psi_k}(x^k, x^{k(\tau+1)})\right)$$

and we can conclude by writing the Abel transformation

$$\sum_{\tau=1}^t \frac{\gamma_\tau}{\eta_\tau^k}\left(D_{\psi_k}(x^k, x^{k(\tau)}) - D_{\psi_k}(x^k, x^{k(\tau+1)})\right) = \frac{\gamma_1}{\eta_1^k} D_{\psi_k}(x^k, x^{k(1)}) + \sum_{\tau=2}^t D_{\psi_k}(x^k, x^{k(\tau)})\left(\frac{\gamma_\tau}{\eta_\tau^k} - \frac{\gamma_{\tau-1}}{\eta_{\tau-1}^k}\right) - \frac{\gamma_t}{\eta_t^k} D_{\psi_k}(x^k, x^{k(t+1)})$$

and bounding the last term by zero. ☐

B. Cesàro convergence

We now consider two particular cases in which the bound of Lemma 2 can be used to prove convergence in the Cesàro sense.

Theorem 1 (Cesàro convergence under identical learning rates)**.** *Suppose that all populations use the same sequence of learning rates (η_t). Then for all k, the regret, discounted by (η_t), is bounded as follows*

$$\sup_{x^k \in m_k(S_k)\Delta^{\mathscr{P}_k}} R^{k(t)}(x^k) \le \frac{L_k^2}{2\ell_{\psi_k}}\sum_{\tau=1}^t (\eta_\tau)^2 + D_{\psi_k}(x^k, x^{k(1)}) \quad (8)$$

This follows immediately from Lemma 2 by taking $\gamma_t = \eta_t$. In particular, if (η_t) converges to 0 and $\sum_{\tau=1}^t \eta_\tau \to \infty$ as $t \to \infty$, then $\sum_{\tau=1}^t \eta_\tau^2 = o(\sum_{\tau=1}^t \eta_\tau)$, and

$$\limsup_{t\to\infty} \sup_{x^k \in m_k(S_k)\Delta^{\mathscr{P}_k}} \frac{R^{k(t)}(x^k)}{\sum_{\tau=1}^t \eta_\tau} \le 0.$$

In other words, the regret discounted by (η_t) is sublinear, and, by Lemma 1, it follows that

$$f(x^{(t)}) \xrightarrow{(\eta_t)} \mathscr{N}$$

where the convergence rate is $O\left(\frac{\sum_{\tau=1}^t \eta_\tau^2}{\sum_{\tau=1}^t \eta_\tau}\right)$. For example, if $\eta_t = \theta\left(\frac{1}{t}\right)$, then the convergence rate is $O\left(\frac{1}{\log t}\right)$. If $\eta_t = \theta(t^{-\alpha})$ with $\alpha \in (\frac{1}{2}, 1)$, then $\frac{\sum_{\tau=1}^t \eta_\tau^2}{\sum_{\tau=1}^t \eta_\tau} = O\left(\frac{1}{t^{1-\alpha}}\right)$. If $\eta_t = \theta\left(\frac{1}{\sqrt{t}}\right)$, then the convergence is $O\left(\frac{\log t}{\sqrt{t}}\right)$, and if $\eta_t = \theta(t^{-\alpha})$ with $\alpha \in (0, \frac{1}{2})$, then $\frac{\sum_{\tau=1}^t \eta_\tau^2}{\sum_{\tau=1}^t \eta_\tau} = O\left(\frac{t^{1-2\alpha}}{t^{1-\alpha}}\right) = O(\frac{1}{t^\alpha})$.

Theorem 2 (Cesàro convergence under bounded Bregman divergence)**.** *Suppose that for all k,*

(i) The Bregman divergence D_{ψ_k} is bounded over $m_k(S_k)\Delta^{\mathscr{P}_k}$, i.e. there exists $D_k > 0$ such that for all $x^k, y^k \in m_k(S_k)\Delta^{\mathscr{P}_k}$, $D_{\psi_k}(x^k, y^k) \le D_k$

(ii) The sequence of learning rates (η_t) is decreasing.

Then taking the discount sequence (γ_t) to be constant equal to 1, we have the following regret bound: for all k and all t,

$$\sup_{x^k \in m_k(S_k)\Delta^{\mathscr{P}_k}} R^{k(t)}(x^k) \le \frac{L_k^2}{2\ell_{\psi_k}}\sum_{\tau=1}^t \eta_\tau^k + \frac{D_k}{\eta_t^k} \quad (9)$$

Proof. Applying Lemma 2 with $\gamma_t = 1$ and observing that $\frac{1}{\eta_\tau^k} - \frac{1}{\eta_{\tau-1}^k} \ge 0$ by assumption on the learning rates, we have

$$R^{k(t)}(x^k) \le \frac{L_k^2}{2\ell_{\psi_k}}\sum_{\tau=1}^t \eta_\tau^k + \frac{D_k}{\eta_1^k} + D_k\sum_{\tau=2}^t\left(\frac{1}{\eta_\tau^k} - \frac{1}{\eta_{\tau-1}^k}\right)$$

where the telescoping sum is equal to $\frac{1}{\eta_t} - \frac{1}{\eta_1}$. This proves the claim. ☐

In particular, if for all k, $\sum_{\tau=1}^t \eta_\tau^k = o(t)$ and $\frac{1}{\eta_t^k} = o(t)$, then the populations all have sublinear regret, and by Lemma 1, $f(x^{(t)}) \xrightarrow{(1)} \mathscr{N}$, with convergence rate $O\left(\frac{\sum_{\tau=1}^t \eta_\tau^k}{t} + \frac{1}{t\eta_t^k}\right)$. For example, if $\eta_t^k = \theta(t^{-\alpha_k})$ with $\alpha_k \in (0,1)$, then $\frac{\sum_{\tau=1}^t \eta_\tau^k}{t} = O(t^{-\alpha_k})$ and $\frac{1}{t\eta_t^k} = O(t^{-(1-\alpha_k)})$, thus the regret is sublinear, and the upper bound is $O(t^{-\min(\alpha_k, 1-\alpha_k)})$.

C. Convergence of $x^{(t)}$

We now turn to the harder question of proving convergence of $x^{(t)}$, as opposed to Cesàro convergence. We start from the following simple observation: if the sequence $f(x^{(t)})$ is eventually non-increasing, then convergence in the sense of Cesàro implies convergence. Indeed, if there exists t_0 such that for all $t \geq t_0$, $f(x^{(t+1)}) \leq f(x^{(t)})$, then for any positive sequence (γ_t) with $\sum_{\tau=1}^t \gamma_\tau \to \infty$,

$$\begin{aligned} f(x^{(t)}) - f^* &\leq \frac{\sum_{\tau=t_0+1}^t \gamma_\tau (f(x^{(\tau)}) - f^*)}{\sum_{\tau=t_0+1}^t \gamma_\tau} \\ &\leq \frac{\sum_{\tau=1}^t \gamma_\tau}{\sum_{\tau=t_0+1}^t \gamma_\tau} \frac{\sum_{\tau=1}^t \gamma_\tau (f(x^{(\tau)}) - f^*)}{\sum_{\tau=1}^t \gamma_\tau} \end{aligned}$$

where the first term is bounded and the second term is the Cesàro mean. This allows us to immediately extend convergence rates of the Cesàro means to convergence rates of the actual sequence, whenever we can show the potential values $f(x^{(t)})$ are eventually monotone.

In the following Lemma, we argue that this is indeed the case for mirror descent dynamics, whenever the sequence of learning rates is vanishing and the potential f has Lipschitz gradient.

Lemma 3 (Smooth potentials are eventually decreasing under vanishing learning rates)**.** *Consider the mirror descent dynamics defined in Algorithm 2, and suppose that*

(i) The function f is differentiable, and has L-Lipschitz gradient,

(ii) For all k, the sequence of learning rates (η_t^k) is decreasing and converges to 0.

Let $t_0 = \min\left\{t : \forall k, \eta_t^k \leq \frac{L}{\ell_{\psi_k}}\right\}$. Then for all $t \geq t_0$, the mirror descent update guarantees

$$f(x^{(t+1)}) \leq f(x^{(t)})$$

Proof. First, since f is assumed to have L-Lipschitz gradient, then for all $x, y \in \Delta$

$$f(x) \leq f(y) + \langle \nabla f(y), x - y \rangle + \frac{L}{2} \|x - y\|^2 \tag{10}$$

In other words, f is upper-bounded by a quadratic. The argument of the proof is as follows: the mirror descent dynamics are obtained by minimizing, at each iteration, an approximation $\tilde{f}$ of the function around the current iterate $x^{(t)}$, given by the linear part of f plus a Bregman divergence term (see Figure 1). By strong convexity of the Bregman divergence, this approximation $\tilde{f}$ is lower-bounded by a quadratic, and when the learning rates are small enough, the curvature is such that $\tilde{f}$ dominates f, with $f(x^{(t)}) = \tilde{f}(x^{(t)})$. Thus minimizing $\tilde{f}$ is guaranteed to decrease the potential value f.

More precisely, we can write the joint dynamics of the mass distributions as follows: for all t,

$$x^{(t+1)} = \arg\min_{x \in \Delta} f(x^{(t)}) + \left\langle x - x^{(t)}, \nabla f(x^{(t)}) \right\rangle + \sum_{k=1}^K \frac{1}{\eta_t^k} D_{\psi_k}(x^k, x^{k(t)})$$

where Δ is the product of scaled simplexes, $\Delta = m_1(S_1)\Delta^{\mathscr{P}_1} \times \cdots \times m_K(S_K)\Delta^{\mathscr{P}_K}$, and $\nabla f(x^{(t)}) = \ell(x^{(t)})$ is the vector of losses, as discussed in Section I-B. Let

$$\tilde{f}(x) = f(x^{(t)}) + \left\langle x - x^{(t)}, \nabla f(x^{(t)}) \right\rangle + \sum_{k=1}^K \frac{1}{\eta_t^k} D_{\psi_k}(x^k, x^{k(t)})$$

Then by strong convexity of each Bregman divergence, we have

$$\begin{aligned} \tilde{f}(x) &\geq f(x^{(t)}) + \left\langle \nabla f(x^{(t)}), x - x^{(t)} \right\rangle + \sum_{k=1}^K \frac{\ell_{\psi_k}}{2\eta_t^k} \|x^k - x^{k(t)}\|^2 \\ &\geq f(x^{(t)}) + \left\langle \nabla f(x^{(t)}), x - x^{(t)} \right\rangle + \frac{L}{2} \sum_{k=1}^K \|x^k - x^{k(t)}\|^2 \quad \forall t \geq t_0 \\ &= f(x^{(t)}) + \left\langle \nabla f(x^{(t)}), x - x^{(t)} \right\rangle + \frac{L}{2} \|x - x^{(t)}\|^2 \\ &\geq f(x) \end{aligned}$$

where the last inequality follows from the quadratic upper bound (10). Therefore for all $t \geq t_0$, $\tilde{f}$ dominates f everywhere, and since $x^{(t+1)} = \arg\min_{x \in \Delta} \tilde{f}(x)$,

$$f(x^{(t+1)}) \leq \tilde{f}(x^{(t+1)}) \leq \tilde{f}(x^{(t)}) = f(x^{(t)})$$

which proves the claim. $\square$

In the routing game, assumption *(i)* of Lemma 3 holds, since the gradient of the potential function f is exactly the loss $\ell(\cdot)$, which is, by definition, Lipschitz continuous as a linear combination of Lipschitz edge losses. Therefore, for any mirror descent dynamics with vanishing learning rates, the mass distributions are guaranteed to converge to $\mathcal{N}$ with the convergence rates given in Theorem 1 and 2.

IV. Numerical Example

We now illustrate some of the convergence results of Section III, on one instance of the routing game. Consider the simple network in Figure 2.

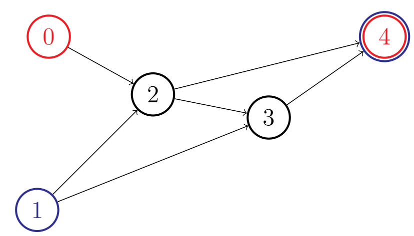


Fig. 2: Routing game example. Population P_1 travels from node 0 to 4, with paths $\mathscr{P}_1 = \{(0,2,4), (0,2,3,4)\}$, and population P_2 travels from node 1 to 4 with paths $\mathscr{P}_2 = \{(1,2,4), (1,3,4), (1,2,3,4)\}$.

We simulate the following mirror descent dynamics: population P_1 uses learning rates $\eta_t = \theta(t^{-\alpha_1})$, $\alpha_1 = .5$ with a Euclidean Bregman divergence $D_{\psi_1}(x,y) = \|x - y\|_2^2$, and population P_2 uses learning rates $\eta_t^2 = \theta(t^{-\alpha_2})$, $\alpha_2 = .2$, with the Bregman divergence $D_{\psi_2}(x,y) = \frac{1}{2}\|x - y\|_1^2$. The trajectories of the mass distributions $x^{k(t)}$ and the path losses $\ell^k(x^{(t)})$ are given in Figure 3. We observe that for

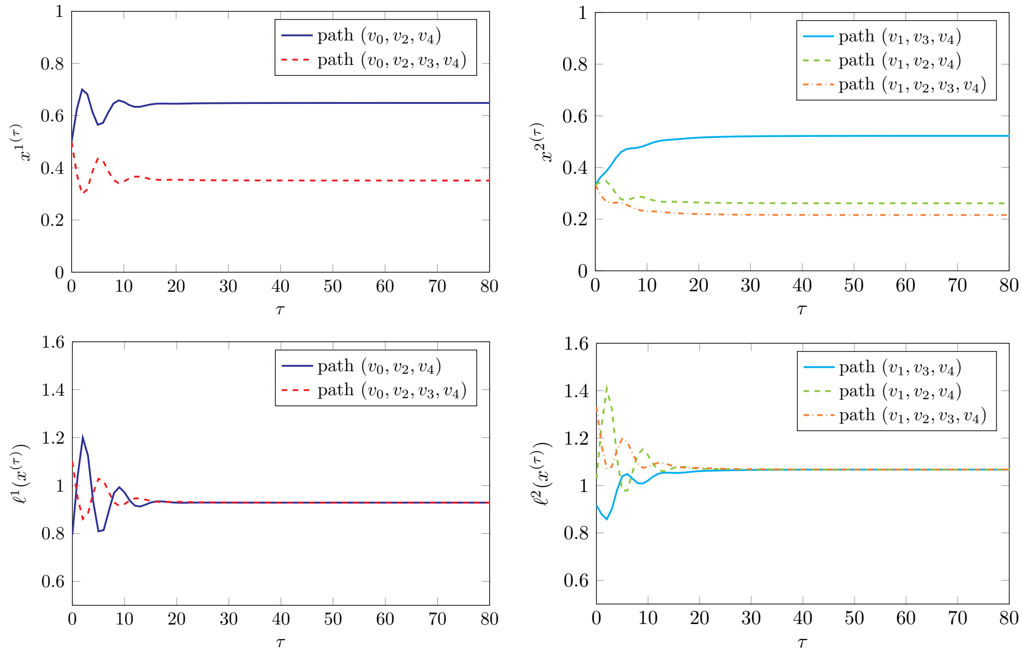


Fig. 3: Simulation results: mass distributions (top) and path losses (bottom) for population P_1 (left) and P_2 (right).

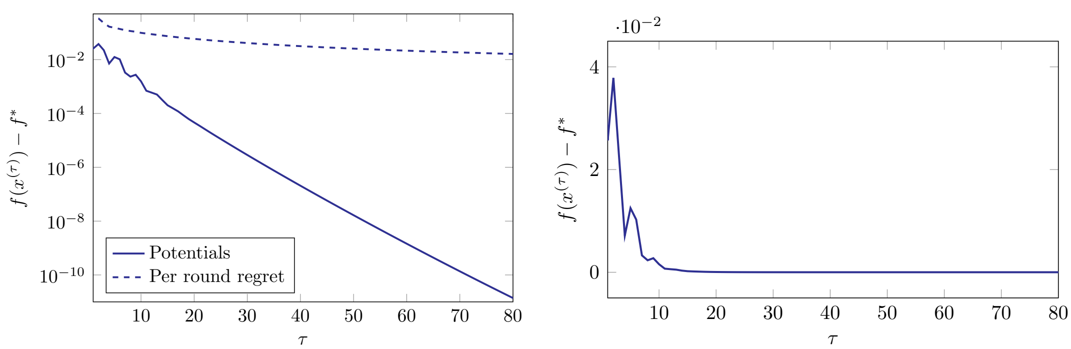


Fig. 4: Potential values $f(x^{(t)}) - f^*$ in log scale (left) and linear scale (right).

each population, the losses converge to a common limit for all paths, which confirms convergence to the set of Nash equilibria of the one-shot game. By Lemma 3, we expect the potentials $f(x^{(t)})$ to be eventually decreasing, which is confirmed by Figure 4. Finally, Figure 5 illustrates that the regrets of both populations converge to 0, which is consistent with the results of Theorem 2.

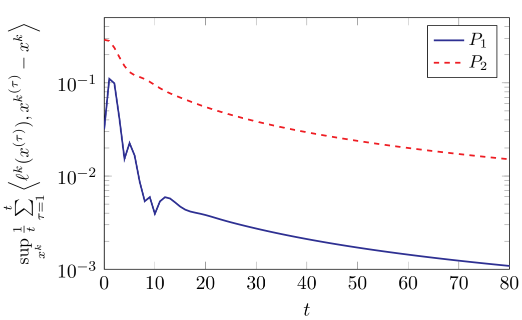


Fig. 5: Cumulative regret of the two populations.

In this simple example, the potential function is, in fact, strongly convex, since for each population, the adjacency matrix is injective. As a result, the regret and the potentials converge faster than the upper bounds provided by Theorem 2.

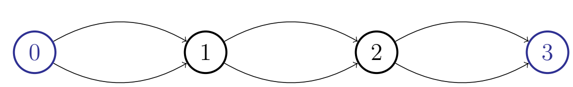


Fig. 6: Second routing game example. Population P_1 travels from node 0 to node 3.

We give a second example in which the potential function is not strongly convex. To simplify, we consider a routing game with a single population on the network of Figure 6. Here, the incidence matrix is non-injective and the Nash equilibrium is non-unique. We simulate dynamics with $\eta_t = \theta(t^{-\alpha})$, $\alpha = .6$, and $D_\psi(x,y) = \sum_{p\in\mathscr{P}} x_p \ln \frac{x_p}{y_p}$, the KL divergence of the distribution x with respect to y,

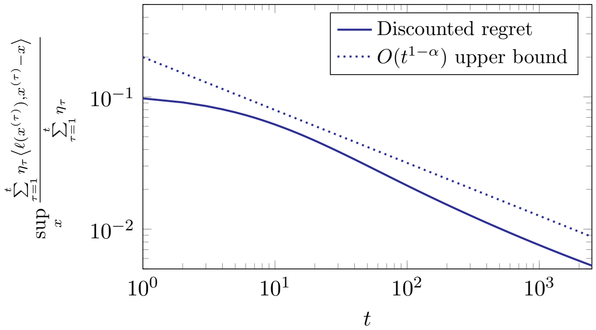


Fig. 7: Discounted per-round regret and corresponding upper bound provided by Theorem 1.

which corresponds to taking ψ to be the negative entropy. By Theorem 1, we have a $O(t^{1-\alpha})$ upper bound on the discounted per-round regret, $\sup_x \frac{\sum_{\tau=1}^{t} \eta_\tau \langle \ell(x^{(\tau)}), x^{(\tau)} - x \rangle}{\sum_{\tau=1}^{t} \eta_\tau}$. The results are given in Figure 7, where the regret decay rate seems to match the rate given by the upper bound.

V. Conclusion

We considered a class of online learning dynamics for the routing game, in which each population updates its mass distribution by applying a mirror descent update using its vector of losses from the previous iteration. We derived a bound on the discounted population regret, and applied it to show that the mass distributions converge in the sense of Cesàro means, and derived convergence rates under different assumptions on the Bregman divergences and the learning rates. We then argued that whenever the populations use vanishing sequences of learning rates, the potentials $f(x^{(t)})$ are eventually decreasing, which proves convergence of $x^{(t)}$ to the set of equilibria, with the same convergence rates as the Cesàro means.

While we derived these results in the context of the routing game, they hold for any potential game in which the potential function is convex. This defines a broad class of online learning dynamics which are guaranteed to converge, together with upper bounds on the convergence rates. However, the bound on the convergence of $\left(f(x^{(t)})\right)_t$ may not be tight in general, since a sequence may converge faster than its Cesàro means. We believe that a more refined analysis may provide tighter bounds on the convergence rate of the potentials.

References

[1] Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Oper. Res. Lett.*, 31(3):167–175, May 2003.

[2] Martin J Beckmann, Charles B McGuire, and Christopher B Winsten. Studies in the economics of transportation. 1955.

[3] Avrim Blum, Eyal Even-Dar, and Katrina Ligett. Routing without regret: on convergence to nash equilibria of regret-minimizing algorithms in routing games. In *Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*, PODC '06, pages 45–52, New York, NY, USA, 2006. ACM.

[4] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*, volume 25. Cambridge University Press, 2010.

[5] D. Braess. Uber ein paradoxon der verkehrsplanung. *Unternehmensforschung*, 12:258–268, 1968.

[6] Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.

[7] Simon Fischer and Berthold Vöcking. On the evolution of selfish routing. In *Algorithms–ESA 2004*, pages 323–334. Springer, 2004.

[8] Walid Krichene, Benjamin Drighès, and Alexandre Bayen. On the convergence of no-regret learning in selfish routing. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, pages 163–171. JMLR Workshop and Conference Proceedings, 2014.

[9] A. S. Nemirovsky and D. B. Yudin. *Problem complexity and method efficiency in optimization*. Wiley-Interscience series in discrete mathematics. Wiley, 1983.

[10] Asuman Ozdaglar and R Srikant. Incentives and pricing in communication networks. *Algorithmic Game Theory*, pages 571–591, 2007.

[11] Robert W Rosenthal. A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory*, 2(1):65–67, 1973.

[12] Tim Roughgarden. Stackelberg scheduling strategies. *SIAM Journal on Computing*, 33(2):332–350, 2004.

[13] Tim Roughgarden. Routing games. In *Algorithmic game theory*, chapter 18, pages 461–486. Cambridge University Press, 2007.

[14] Tim Roughgarden and Éva Tardos. How bad is selfish routing? *Journal of the ACM (JACM)*, 49(2):236–259, 2002.

[15] William H Sandholm. Potential games with continuous player sets. *Journal of Economic Theory*, 97(1):81–108, 2001.

[16] Jörgen W Weibull. *Evolutionary game theory*. MIT press, 1997.