On the Convergence of Regret Minimization Dynamics in Concave Games

Eyal Even-Dar* Yishay Mansour† Uri Nadav‡

November 17, 2008

Abstract

We study a general sub-class of concave games, which we call socially concave games. We show that if each player follows any no-external regret minimization procedure then the dynamics converges in the sense that both the average action vector converges to a Nash equilibrium and that the utility of each player converges to her utility in that Nash equilibrium.

We show that many natural games are socially concave games. Specifically, we show that linear Cournot competition and linear resource allocation games are socially-concave games, and therefore our convergence result applies to them. In addition, we show that a simple best response dynamic might diverge for linear resource allocation games, and is known to diverge for a linear Cournot competition. For the TCP congestion games we show that "near" the equilibrium these games are socially-concave, and using our general methodology we show convergence of specific regret minimization dynamics.

*Google Research, 76 Ninth Avenue, New York, NY 10011. Email: evendar@google.com

†Google Research, 76 Ninth Avenue, New York, NY 10011 and Blavatnik School of Computer Science, Tel Aviv University. Email: mansour@tau.ac.il

‡Blavatnik School of Computer Science, Tel Aviv University. Email: urinadav@cs.tau.ac.il

1 Introduction

Equilibrium is a widely acceptable solution concept in economics, in general, and in game theory in particular. The basic definition of an equilibrium is static: describing a steady state of a system where no agent has an incentive to unilaterally deviate from. While an equilibrium can be viewed as a steady state, an important conceptual issue is how such an equilibrium can be reached, and which *dynamics* lead to an equilibrium (see, [16, 28]). Understanding the dynamics has two different perspectives. The first perspective is developing general dynamics that will always reach an equilibrium (for any game). This line of research includes computationally efficient procedures for correlated equilibrium [12, 20, 7, 5] and inherently exponential time procedures for Nash equilibrium [13, 14, 28, 11, 21, 17] (see, [19] for communication complexity lower bounds). The second perspective is analyzing specific simple and natural behavior and its resulting dynamics in concrete games, which is the focus of our work.

When considering a dynamics, an important and critical question is to what extent are the dynamics "natural". We are interested in natural dynamics, since they will support the belief that the system naturally reaches an equilibrium. We would like the individual agent procedures to be *rational*, where the agent can be viewed as trying to maximize her long-term utility. We would like the dynamics to be *distributed*, where each agent has very little knowledge about other agents' utilities, and is mainly concerned with her own utility. We would like the dynamics to be *simple*, which would support the idea that they naturally occur. Finally, we would like the dynamics to be *flexible* in the sense that different agents may behave differently (and will not have to follow specific prescribed procedure). Examples of such widely studied dynamics are *best response*, where the agent selects an action which is a best response to the current action of the other players, and *fictitious play* where an agent selects an action which is best response to the average action profile of the other players (see, [16]).

Regret minimization procedures match all the above requirements. In a regret minimization setting each agent focuses on her own utility (and conceptually is oblivious to the other agents utilities). By *regret* we compare the agent's average utility to that of the best procedure from a given limited class of procedures. Having *no-regret* means that no procedure in that class would improve significantly the agent's utility. The *no-external regret* compares the agent's utility to that of a fixed constant action. The *no-internal regret* uses a larger comparison class, where the agent can swap each time it did a specific action a_1 by another action a_2. The main results are that there are no-regret algorithm and that their average regret vanishes at the rate of $O(T^{-1/2})$, where T is the number of time steps. In this work we focus on no-external regret procedures.

Regret minimization procedures prescribe to most of the requirement we mentioned. It is rational, in the sense that the agent has a guarantee on her own utility regardless of how the other agents act. It is distributed, since an agent needs to be aware only of her own utility. Many of the no-external regret procedures are very simple, and they share the idea that an agent increases the weight on actions that have been doing well. There is a large variety of no-external regret procedures that have been studied, and more conceptually, the assumption is not tied to any specific procedure, but describes the payoffs of the agents in retrospect.

The regret minimization dynamics have attracted significant attention in recent years. First, the no-internal regret dynamics converges to the set of approximate correlated equilibria [12, 20, 7, 5]. However, for continuous games, no general efficient no-internal regret algorithm is known. Second, there have been works addressing the properties of no-external regret dynamics in specific games. In $z$ero-sum games it is known that the no regret dynamics will converge to the min-max solution

[15]. However, for general games, it is known that the no external regret does not necessarily converge to equilibrium [30]. Blum et al. [3] analyzed the no-external regret dynamics in routing games showing that most of the agents, most of the time, are playing near equilibrium strategies. In a subsequent work, Blum et al. [4] defined the *price of total anarchy* as the ratio of the optimum to that of the worst no-external regret dynamics, and showed that it can be similar to the price of anarchy even in cases where the dynamics do not lead to an equilibrium.

In this work we will be studying no-external regret dynamics of various concave games. In a concave game the utility function of each player is concave in her own action. Rosen [26] showed that a Nash equilibrium in pure strategies always exists in such games. We concentrate on a sub-class of concave games, which we call *socially concave games*. We show that many interesting games are socially-concave. including linear Cournot competition [25], and linear resource allocation games [23].[1] We also show that TCP congestion control [24] games are socially-concave "near" their equilibrium.

We derive a general convergence result, showing that if each agent follows a no-external regret procedure, then the dynamics, in any socially-concave game reaches an equilibrium. The convergence is both of the average action vector, and the payoff of the individual agents. On the other hand, we show that a best response dynamics might diverge for linear resource allocation games, and it is known to diverge for linear Cournot competition [27]. For the TCP congestion control setting we study two different games, depending on how the network handles overflow. The first is related to router's tail-drop policies, was proposed in [24], and studied in the context of competitive online algorithms [24, 1]. Our second is motivated by router's policy of Random Early Discard (RED) [10]. In both cases, although the games are not socially-concave, we show that gradient based no-external regret procedures (such as [29]) guarantee the desired convergence.

All proofs, excluding the proof of our main theorem (Theorem 3.1) are deferred to the appendix, for lack in space.

2 Model

An n-person game is denoted by the tuple $\Gamma = (N, (S_i)_{i\in N}, (u_i)_{i\in N})$, where $N = \{1, \ldots, n\}$ is the set of n players, S_i is the set of actions of player i and $u_i : S \to \mathbb{R}$ is player's i utility function, where $S = S_1 \times S_2 \times \cdots \times S_n$ is the joint action set.

For $s \in S$ let s_{-i} denote the strategy combination of all players except i, *i.e.*, $s_{-i} = (s_j)_{j\neq i}$. An action $x_i \in S_i$ is an ϵ-best response to s_{-i} if for every action $y_i \in S_i$, $u_i(x_i, s_{-i}) \geq u_i(y_i, s_{-i}) - \epsilon$. A joint strategy $s \in S$ is an ϵ-*Nash equilibrium*(NE), if for each player i we have that $s_i \in S_i$ is an ϵ best response to s_{-i}. For $\epsilon = 0$ a 0-Nash equilibrium is a Nash equilibrium and a 0-best response is a best response and is denoted by BR.

For the rest of the paper we assume that S_i is a closed, convex and bounded action set and that $S_i \subset \mathbb{R}^{m_i}$ for some $m_i \in \mathbb{N}$. Also, we assume that the utility functions u_i are twice differentiable and bounded from above by 1.

In his seminal paper, Rosen [26] considered the class of *concave games*, where every player's utility function u_i is concave in her own action $s_i \in S_i$. In this paper we consider a closely related class of *socially concave* games we now define.

[1] In a linear Cournot competition multiple firms compete by setting their production levels, and the price is a linear function of the overall production level. In a linear resource allocation game where the agent receive a share of the resource (e.g., bandwidth or a market share), as a function of their investment.

Definition 2.1. *A game is* socially concave *if the following holds:*

A1 There exists a strict convex combination of the payoff functions which is a concave function. Formally, there exists an n-tuple $(\lambda_i)_{i \in N}$, $\lambda_i > 0$, and $\sum_{i \in N} \lambda_i = 1$, such that $g(x) = \sum_{i \in N} \lambda_i u_i(x)$ is a concave function in x.

A2 The utility function of each player i, is convex in the actions of the other players, i.e., for every $s_i \in S_i$ the function $u_i(s_i, x_{-i})$ is convex in $x_{-i} \in S_{-i}$.

Next, we show that the class of socially concave games is a sub-class of concave games.

Lemma 2.2. *If u_i is twice differentiable for every $i \in N$, then a socially concave game is also a concave game.*

Rosen [26] showed that every concave game has a Nash equilibrium point in S. Therefore, every socially concave game has a Nash equilibrium.

External Regret Minimization and Online Concave Programming

Definition 2.3. *A concave programming problem, C, consists of a convex feasible set $F \subset \mathbb{R}^m$ and an infinite sequence $\{f^1, f^2, \ldots\}$ where each $f^\tau : F \to \mathbb{R}$ is a concave function.*

An online concave programming algorithm *selects, at each time step t, a vector $x^t \in F$, given the history. After x^t is selected, the algorithm observes the payoff function $f^t(\cdot)$, and receives a payoff of $f^t(x^t)$.*

After defining the basic setting, let us define the online problem.

Definition 2.4. *Let A be an online concave programming algorithm, and let $x \in F$ be the optimal static solution, i.e., $x = \arg\min_y \sum_{\tau=1}^t f^t(y)$, then the regret of algorithm A is defined by*

$$\mathcal{R}_A(T) = \sum_{t=1}^{T} f^t(x^t) - \min_x \sum_{t=1}^{T} f^t(x).$$

An algorithm A has no external regret, if $\mathcal{R}_A(T) = o(T)$.

Regret Minimization bounds for concave games: Zinkevich [29] provides a simple procedure that guarantees a regret of at most $\mathcal{R}(T) = O(\sqrt{T})$, when F is a closed, convex, bounded and non-empty set, and that every f^t is differentiable with a bounded first derivative. Hazan *et al.*[22], give algorithms that guarantee an upper bound of $O(\log(T))$ on the external regret, when f^t are strictly concave functions and twice differentiable. Other gradient based procedures include [9, 2].
Regret and repeated games: The regret of player i in a game Γ following the sequence of joint actions $(x)_{t=1}^T$ is $\mathcal{R}_i((x)_{t=1}^T) = \sum_{t=1}^T u_i(x^t) - \max_{y \in S_i} u_i(y, x_{-i}^t)$. We define an upper bound $\mathcal{R}_i(T) = \max_{(x)_{t=1}^T \in S^T} \mathcal{R}_i(T)$.

3 Main Result

We show that for the class of socially concave games, no-external regret behavior leads to a Nash equilibrium. For every t, let $\hat{x}^t$ be the average of the n-tuples of strategies played up to time t, that is $\hat{x}^t = \frac{1}{t}\sum_{\tau=1}^t x^\tau$. Let $\hat{u}^t$ denote the average payoff vector up to time t, that is, for every player i, $\hat{u}_i^t = \frac{1}{t}\sum_{\tau=1}^t u_i(x^\tau)$.

Theorem 3.1 (Main Theorem)**.** *Let $\Gamma = (N, (S_i)_{i\in N}, (u_i)_{i\in N})$ be an n-players socially concave game. If every player i plays according to a procedure with external regret bound $\mathcal{R}_i(t)$, then at time t,*

(i) The average strategy vector $\hat{x}^t$ is an ϵ^t-Nash equilibrium, where $\epsilon^t = \frac{1}{\lambda_{\min}} \sum_{j\in N} \frac{\lambda_j \mathcal{R}_j(t)}{t}$ and $\lambda_{\min} = \min_{j\in N} \lambda_j$.

(ii) The average utility of each player i is close to her utility at $\hat{x}^t$, the average vector of strategies. Formally,

$$|\hat{u}_i^t - u_i(\hat{x}^t)| \le \frac{1}{\lambda_i} \sum_{j\in N} \frac{\lambda_j \mathcal{R}_j(t)}{t}.$$

Proof. The proof is constructed in the following six steps:

Step I: By definition of a regret minimization algorithm, for any period $\{1, \dots, t\}$, each player $i \in N$ has low regret to any action $s_i \in S_i$. Specifically, we can apply it to the best response to $\hat{x}^t_{-i} = \frac{1}{t}\sum_{\tau=1}^t x_{-i}$, *i.e.*, $\mathrm{BR}_i(\hat{x}^t_{-i}) \in S_i$. Therefore,

$$\hat{u}_i^t = \frac{1}{t}\sum_{\tau=1}^t u_i(x^\tau) \ge \max_{s_i\in S_i} \frac{1}{t}\sum_{\tau=1}^t u_i\left(s_i, x^\tau_{-i}\right) - \mathcal{R}_i(t) \ge \frac{1}{t}\sum_{\tau=1}^t u_i\left(\mathrm{BR}_i(\hat{x}^t_{-i}), x^\tau_{-i}\right) - \frac{\mathcal{R}_i(t)}{t}. \tag{3.1}$$

Step II: Fix an action $y_i \in S_i$ of player i. Property A2 states that $u_i(y_i, x_{-i})$ is convex in its second argument x_{-i}, which implies,

$$\frac{1}{t}\sum_{\tau=1}^t u_i(y_i, x^\tau_{-i}) \ge u_i(y_i, \frac{1}{t}\sum_{\tau=1}^t x^\tau_{-i}) = u_i(y_i, \hat{x}^t_{-i}). \tag{3.2}$$

Step III: By definition of best response, for every n-tuple of strategies $y \in S$,

$$u_i(\mathrm{BR}_i(y_{-i}), y_{-i}) \ge u_i(y). \tag{3.3}$$

Step IV: Property A1 states that for some $\lambda = (\lambda_i)_{i\in N}$, $g(x) = \sum_{i\in N} \lambda_i u_i(x)$ is concave. By the concavity of $g(x)$ we have that,

$$\sum_{i\in N} \lambda_i u_i(\hat{x}^t) = \sum_{i\in N} \lambda_i u_i(\frac{1}{t}\sum_{\tau=1}^t x^\tau) \ge \sum_{i\in N} \lambda_i \sum_{\tau=1}^t \frac{1}{t} u_i(x^\tau) = \sum_{i\in N} \lambda_i \hat{u}_i^t. \tag{3.4}$$

Step V: Combining Inequalities (3.1)—(3.4) we get the following chain of inequalities:

$$\begin{aligned}
\sum_{i\in N} \lambda_i \hat{u}_i^t = \sum_{i\in N} \lambda_i \left(\frac{1}{t}\sum_{\tau=1}^t u_i(x^\tau)\right) &\overset{(a)}{\ge} \sum_{i\in N} \lambda_i \left(\frac{1}{t}\sum_{\tau=1}^t u_i\left(\mathrm{BR}_i(\hat{x}^t_{-i}), x^\tau_{-i}\right) - \frac{\mathcal{R}_i(t)}{t}\right) \\
&\overset{(b)}{\ge} \sum_{i\in N} \lambda_i \left(u_i(\mathrm{BR}_i(\hat{x}^t_{-i}), \hat{x}^t_{-i}) - \frac{\mathcal{R}_i(t)}{t}\right) \\
&\overset{(c)}{\ge} \sum_{i\in N} \lambda_i \left(u_i(\hat{x}^t) - \frac{\mathcal{R}_i(t)}{t}\right) \\
&\overset{(d)}{\ge} \sum_{i\in N} \lambda_i \left(\frac{1}{t}\sum_{\tau=1}^t u_i(x^\tau)\right) - \sum_{i\in N} \lambda_i \frac{\mathcal{R}_i(t)}{t},
\end{aligned}$$

where (a) follows from (3.1), (b) follows from (3.2) with $y_i = \mathrm{BR}_i(\hat{x}^t_{-i})$, (c) follows from (3.3) with $y = \hat{x}^t$, and (d) follows from (3.4).

The above inequalities imply that

$$|\sum_{i\in N}\lambda_i u_i(\hat{x}^t) - \sum_{i\in N}\lambda_i u_i(\mathrm{BR}_i(\hat{x}^t_{-i}), \hat{x}^t_{-i})| \leq \sum_{i\in N}\lambda_i \frac{\mathcal{R}_i(t)}{t},$$

since given any set of a_i's such that $a_1 \geq a_2 \geq a_3 \geq a_4 \geq a_5$ we have that $|a_3 - a_2| \leq a_1 - a_5$. Hence,

$$\sum_{i\in N}\lambda_i u_i(\hat{x}^t) \geq \sum_{i\in N}\lambda_i u_i(\mathrm{BR}_i(\hat{x}^t_{-i}), \hat{x}^t_{-i}) - \sum_{i\in N}\lambda_i \frac{\mathcal{R}_i(t)}{t}. \tag{3.5}$$

By definition, for every i, $u_i(\hat{x}^t) \leq u_i(\mathrm{BR}_i(\hat{x}^t_{-i}), \hat{x}^t_{-i})$. Therefore, $\sum_{i\neq j}\lambda_i u_i(\hat{x}^t) \leq \sum_{i\neq j}\lambda_i u_i(\mathrm{BR}_i(\hat{x}^t_{-i}), \hat{x}^t_{-i})$. Combining this with (3.5) we have,

$$\lambda_j u_j(\hat{x}^t) \geq \lambda_j u_j(\mathrm{BR}_j(\hat{x}^t_{-j}), \hat{x}^t_{-j}) - \sum_{i\in N}\frac{\lambda_i \mathcal{R}_i(t)}{t}.$$

Thus $\hat{x}^t$ is an ϵ^t-Nash equilibrium for $\epsilon^t = \frac{1}{\lambda_{\min}}\sum_{i\in N}\lambda_i \frac{\mathcal{R}_i(t)}{t}$.

Step VI: Similar to Step V, from Inequalities (3.1), (3.2) and (3.3),

$$\begin{aligned}
\hat{u}^t_i = \left(\frac{1}{t}\sum_{\tau=1}^{t} u_i(x^\tau)\right) &\overset{(a)}{\geq} \frac{1}{t}\sum_{\tau=1}^{t} u_i\left(BR(\hat{x}^t_{-i}), x^\tau_{-i}\right) - \frac{\mathcal{R}_i(t)}{t} \\
&\overset{(b)}{\geq} u_i(\mathrm{BR}_i(\hat{x}^t_{-i}), \hat{x}^t_{-i}) - \frac{\mathcal{R}_i(t)}{t} \\
&\overset{(c)}{\geq} u_i(\hat{x}^t) - \frac{\mathcal{R}_i(t)}{t}
\end{aligned}$$

where (a) follows from (3.1), (b) follows from (3.2), and (c) follows from (3.3). Therefore, it follows that the average utility of player i, $\hat{u}^t_i$, is at least her utility when the average strategy vector is played, $u_i(\hat{x}^t)$, minus her own average regret,

$$\hat{u}^t_i \geq u_i(\hat{x}^t) - \frac{\mathcal{R}_i(t)}{t}. \tag{3.6}$$

From Inequality (3.4) we have that $\sum_{i\in N}\lambda_i u_i(\hat{x}^t) \geq \sum_{i\in N}\lambda_i \hat{u}^t_i$. Therefore,

$$\hat{u}^t_i - u_i(\hat{x}^t) \leq \frac{1}{\lambda_i}\sum_{j\neq i|j\in N}\lambda_j\left(u_j(\hat{x}^t) - \hat{u}^t_j\right) \leq \frac{1}{\lambda_i}\sum_{j\neq i|j\in N}\lambda_j \frac{\mathcal{R}_j(t)}{t} \leq \frac{1}{\lambda_i}\sum_{j\in N}\lambda_j \frac{\mathcal{R}_j(t)}{t}, \tag{3.7}$$

where the second inequality follows from (3.6).

Combining the two inequalities (3.6) and (3.7), we bound the difference between player's i average utility and its utility at $\hat{x}^t$, *i.e.*,

$$|\hat{u}^t_i - u_i(\hat{x}^t)| \leq \frac{1}{\lambda_i}\sum_{j\in N}\lambda_j \frac{\mathcal{R}_j(t)}{t}.$$

□

The following are immediate consequences of Theorem 3.1. First, if every player employs a no-regret algorithm, then the average strategy vector converges to a Nash equilibrium, and the average payoff of each player converges to her payoff at that Nash equilibrium. Second, if every player employs the *generalized infitisimal gradient ascent* algorithm [29], which has regret $O(\sqrt{t})$, then after t steps the average strategy vector is an $O(n/\sqrt{t})$-Nash equilibrium and the average payoff of each player differs from its payoff at that Nash equilibrium by at most $O(n/\sqrt{t})$, assuming that $\lambda_{\min}$ is bounded away from zero and that the utility's values are in the range $[0, 1]$.

4 Cournot competition

Cournot competition [8] is a fundamental economic model used to describe competition between firms. The model considers multiple firms (oligopoly), which produce the same good. The main interaction between the firms is due to their influence on the good market price. Specifically, each firm decides on its production level (the quantity it produces from the good), and incurs an associated cost (which depends on the quantity, and may be different for different firms). The revenue of a firm is the product of its quantity and the market price, where market price depends on the aggregate quantity produced by all firms. Let us first define formally a Cournot competition.

Definition 4.1. *A* Cournot competition *is a game* $\Gamma^C = (N, (S_i)_{i\in N}, (c_i)_{i\in N}, p, (u_i)_{i\in N})$, *where* N *is the set of firms,* $S_i = \mathbb{R}_+$ *is the quantity firm* i *decides to produces,* $c_i : S_i \to \mathbb{R}_+$ *is the cost of firm* i *to produce a quantity* q_i, $p : S \to \mathbb{R}_+$ *is the price of the good (where* $p(s)$ *is a function of the aggregate quantity, i.e.,* $\sum_{i\in N} s_i$*), and* $u_i : S \to \mathbb{R}$ *is firm* i *utility such that* $u_i(s) = s_i p(s) - c_i(s_i)$. *A* linear Cournot competition *is a Cournot competition where* $p(s) = a - b(\sum_{i\in N} s_i)$, *where* a *and* b *are some positive constants and the cost function* c_i *are convex.*

A well known result by Cournot (cf. [16, pp. 11]), is that if two firms participate in linear Cournot competition, and both play simultaneously the best response dynamics, then their joint play converges to equilibrium Theocharis [27] showed that the best response dynamics fails to converge in linear Cournot competition, whenever there are four or more firms participating.

We show that in a linear Cournot competition, if each firm bases its production level on a no-external regret algorithm, then the firms' average payoffs and quantities produced converges to the unique NE of the linear Cournot competition.

Lemma 4.2. *A linear Cournot competition is a socially concave game.*

Proof. First, the sum of utilities is

$$g(s) = \sum_{i\in N} u_i(s) = \sum_{i\in N} s_i(a - b(\sum_{j\in N} s_j)) - \sum_{i\in N} c_i(s_i),$$

which is concave in s, hence Assumption A1 in Definition 2.1 holds. Since $u_i(s_i, s_{-i}) = a - b(\sum_{j\in N} s_j) - c_i(s_i)$ is a linear function in s_{-i} it is also convex in s_{-i}, hence, Assumption A2 in Definition 2.1 holds as well. $\square$

Since a linear Cournot competition is a socially concave game, Theorem 3.1 implies the convergence of the no-external regret dynamics.

Theorem 4.3. *In a linear Cournot competition, if each firm $i \in N$ follows a regret minimization algorithm with $\mathcal{R}_i(T) = O(\sqrt{T})$, then after T steps the average strategy vector will be $O(n/\sqrt{T})$-Nash equilibrium, and the payoff of each firm will be within $O(n/\sqrt{T})$ of its payoff at the average strategy vector.*

Remark: We can show convergence for a larger class of Cournot competition. Namely, consider the case that the cost functions c_i are convex, $xp(x)$ is concave, and $p(x)$ is convex, where $x = \sum_{i \in N} s_i$. Since $p(x)$ is convex, it implies that $u_i(s)$ is convex in s_{-i}, and this satisfies Assumption A2 in Definition 2.1. Since the function $g(s) = \sum_{i \in N} s_i p(x) - c_i(s_i) = xp(x) - \sum_{i \in N} c_i(s_i)$ is concave, Assumption A1 in Definition 2.1 holds. Therefore, this is a socially concave game, and Theorem 3.1 guarantee the convergence.

5 Linear Resource Allocation Games

In a resource allocation game [18, 23] n users share a communication link of capacity $C > 0$ (we assume without loss of generality $C = 1$). Let d_i denote the rate allocated to user i. We assume that user i receives a value $\varphi_i(d_i)$ if the allocated rate is d_i; we assume that this value is measured in monetary units.

Each user i submits a "bid" w_i to the network from a bid space $S_i = [b_{\min}, 1]$, where $b_{\min} > 0$ [2]. The network accepts these submitted bids and determines the share of capacity each user is allocated, according to an allocation function $M : S^n \to [0, C]^n$, mapping the bids to a feasible allocation (*i.e.,* for any $w \in S$ we have $\sum_{i \in N} M_i(s) = 1$ and $M_i(w) \geq 0$ for every i). This makes the model a game Γ^{R} between the n users, $\Gamma^{\mathrm{R}} = (N, (S_i)_{i \in N}, (u_i)_{i \in N}, M, (\varphi_i)_{i \in N})$, where user i utility function is $u_i(w) = \varphi_i(M_i(w)) - w_i$ and φ_i is its value function.

Hajek and Gopalakrishnan [18] studied resource allocation games with the proportional allocation function $M_i(w) = \frac{w_i}{\sum_{j \in N} w_j}$. They showed that when the value function, $\varphi_i(\cdot)$, of each user is concave, a unique Nash equilibrium of the Γ^{R} exists. The existence of a pure Nash equilibrium for a more general class of quasi proportional allocation functions where $M_i(w) = w_i^c / (\sum_{j \in N} w_j^c)$, for some $c \in (0, 1]$, follows from [26].

We will concentrate on the following sub-class of resource allocation games.

Definition 5.1. *A* linear resource allocation game *is a resource allocation game* $\Gamma^{\mathrm{R}} = (N, (S_i)_{i \in N}, (u_i)_{i \in N}, M, (\varphi_i)_{i \in N})$ *such that* $\varphi_i(d_i) = \alpha_i d_i$ *and the allocation mechanism is quasi-proportional,* $M_i(w) = \frac{w_i^c}{\sum_{j \in N} w_j^c}$*, where* $c \in (0, 1]$.

Theorem 5.2. *In a linear resource allocation game, if every player employs a procedure with no external regret, then the average action of the players will converge to a Nash equilibrium, and the average payoff of each player will converge to its payoff in that Nash equilibrium.*

Theorem 5.2 follows immediately from Theorem 3.1, once we establish in the following lemma that a linear resource allocation game is a socially concave game.

Lemma 5.3. *A linear resource allocation game is socially concave game.*

[2] We require that $b_{\min} > 0$ in order to avoid the technicalities when all the players select a zero bid.

5.1 Divergence of the best response dynamics in resource allocation

In the *best response dynamics*, every player optimizes her decision for the next step assuming all other players will play the same as they did in the previous step. Namely, at time t, player i plays $x^t = \mathrm{BR}_i(x_{-i}^{t-1})$. Clearly, for this dynamics, if the joint vector of actions converges, then it must be that it converges to a Nash equilibrium. However, unlike the no regret dynamics, we show that the best response dynamics is not guaranteed to converge.

Consider an n-player linear resource allocation game with $c = 1$, where all n players have identical utility for resource $\varphi_i(d_i) = d_i$. Namely, the utility of player i, in terms of her bid and the bids of the other players is,

$$u_i(x) = \frac{x_i}{\sum_{j \in N} x_j} - x_i.$$

The best response of player $i \in N$ to $x_{-i} \in S_{-i}$ is $\max(b_{min}, \sqrt{\sum_{j \neq} x_j} - \sum_{j \neq} x_j) \in S_i$.

Theorem 5.4. *Consider a linear allocation game with identical utilities $\varphi_i(s_i) = s_i$. For $n \geq 4$, the best response dynamics does not necessarily converge.*

6 Congestion Control Protocols

In this section we present an application of Theorem 3.1 in the design of protocols for congestion control. We consider multiple connections sharing a network path, with a common bottleneck. Time is divided into successive rounds, *i.e.,* time is discrete and events happen only at time $t = 1, 2, 3, \ldots$. It is assumed that all connections have the same round trip time. At the start of a round each connection transmits a window of packets and at the end of a round each connection receives a feedback with the number of packets that were actually delivered. Pending data is always available for sending at every source.

Each connection is associated with a single selfish player. A player benefits from delivering packets, and suffers a penalty for dropped packets, due to retransmission delays and overhead. We view the actions of the players as the 'load' they introduce (or alternatively, the bandwidth they consume). Let x_i^t denote the load imposed by the i'th player at time t, and denote by b_i^t her actual load *i.e.,* the fraction of packets that were not dropped. Player i utility is $b_i^t - \alpha_i(x_i^t - b_i^t)$, namely, α_i is a parameter that reflects i's cost for losing a packet (or more precisely, one 'unit' of bandwidth).

The shared bottleneck is controlled by a router with finite capacity C assumed to be 1 unit of bandwidth. The router receives packets from each connection and decides to forward or discard each packet [3]. Different scheduling policies would share the capacity in different ways. And generally, a scheduling policy maps transmission rate vectors $x = (x_i)_{i \in N} \in \mathbb{R}_+^n$ to feasible bandwidth allocations $\{b \mid \sum_i b_i \leq C = 1\}$. Clearly, no mechanism can assign more than the link capacity, but some mechanisms might be more restrictive. We consider several scheduling policies here,

1. *Tail drop* (TD). Load is accepted while the channel is not full; packets are dropped when the total transmission exceeds the capacity.

[3] No queueing is assumed in our model. The technical difficulty with introducing a queue in our model is that a queue is essentially a state, and this will be a miss-match to both the repeated nature of the game and the regret minimization.

2. *Random early discard (RED).* Packets are randomly dropped with a dropping probability that increases with the offered load.

The policies, TD and RED, to be formally defined in later sections, when combined with a vector $\alpha = (\alpha_i)_{i \in N}$, admits a game $\Gamma^{\text{TD}}, \Gamma^{\text{RED}}$, respectively. Unfortunately, none of these games is socially concave, and therefore Theorem 3.1 cannot be applied directly. Luckily, in a region near the NE of every such game, these games becomes a socially concave game. We use this fact, to show that the *generalized infitisimal gradient ascent* (GIGA) procedure attains the convergence properties guaranteed for socially concave games in Theorem 3.1.

Definition 6.1 (GIGA)**.** *For player i with utility function $u_i(\cdot)$, and strategy space S_i, GIGA sets i's action at time t, x_i^t in the following method:*

$$y_i^t = x_i^{t-1} + \eta^t \frac{\partial}{\partial x_i} u_i(x^{t-1}) \ ; \ x_i^t = \pi(S_i, y_i^t)$$

where $\pi(S_i, y_i^t)$ is the projection of y_i^t into the set S_i,[4] *and η^t is a learning rate where we assume that: (1) η^t is non-increasing in t, i.e., $\eta^t \geq \eta^{t+1}$, (2) η^t vanishes, i.e., for every $\epsilon > 0$ there is a time t^ϵ such that $\eta^{t^\epsilon} < \epsilon$, and (3) that the sum of η^t diverges, i.e., for any ρ there is at time t^ρ, such that $\sum_\tau^{t^\rho} \eta^\tau > \rho$.*[5]

Assuming that a player learns her own payoff as a feedback in each of the games Γ^{TD} and Γ^{RED}, it is possible for her to compute her own gradient $\frac{\partial}{\partial x_i} u_i(x^t)$, at time t, and calculate her next action in the GIGA procedure accordingly.

6.1 Tail Drop

When a router employs a *tail drop* policy, packets are accepted as long as the overall load does not exceed the link capacity. This is model as follows. While $S(x) \leq 1$, every player gets her transmission rate, x_i. But, when $S(x) > 1$ player i gets only a share of the capacity, proportional to her transmission rate. This implies the following utility function:

$$u_i^{\text{TD}}(x) = \begin{cases} x_i & S(x) \leq 1 \\ \frac{x_i}{S(x)} - \alpha_i(x_i - \frac{x_i}{S(x)}) & S(x) > 1 \end{cases}$$

Let $\alpha_{\min} = \min_{j \in N} \alpha_j$; $\alpha_{\max} = \max_{j \in N} \alpha_j$. We assume that the penalty per packet loss parameters of the players are bounded as follows, $\frac{3}{n-1} \leq \alpha_{\min} \leq \alpha_{\max} \leq 1$. We also assume that a single player load never exceeds the channel capacity, *i.e.*, $x_i \in [b_{\min}, 1]$, where $b_{\min} > 0$.[6]

One can verify that the game $\Gamma^{\text{TD}} = (N, ([b_{\min}, 1])_{i \in N}, (u_i^{\text{TD}})_{i \in N})$ is not a socially concave game. However, a slight modification of Γ^{TD} results in a socially concave game, $\Gamma^{\text{TD}'}$. $\Gamma^{\text{TD}'}$ has the same set of players and the same strategy space, and a different utility functions. Player i utility function is modified to

$$q_i(x) = \frac{x_i}{S(x)} - \alpha_i(x_i - \frac{x_i}{S(x)}).$$

[4] In our setting the projection only means that if $y_i^t < 0$ then we set $x_i^t = 0$.

[5] In case that the derivative $\frac{\partial}{\partial x_i} u_i(x^{t-1})$ is not continuous we define it to be the limit of the derivatives $x' < x$, which will be always well define in our settings.

[6] In the final version we extend the proof to handle the case that $b_{\min} = 0$.

Lemma 6.2. *The game* $\Gamma^{\mathrm{TD}'} = (N, ([b_{\min}, 1])_{i\in N}, (q_i)_{i\in N})$*, is a socially concave game for any* $b_{\min} > 0$*.*

Theorem 3.1 does not apply to Γ^{TD}, as it is not a socially concave game. Even so, using the fact that the game $\Gamma^{\mathrm{TD}'}$ is a socially concave game, we can show that the GIGA dynamics (*i.e.,* when all players act according to the GIGA procedure), attains similar convergence properties as general no-regret dynamics in socially concave games.

Theorem 6.3. *Assuming there are at least* $n \geq 4$ *players in the game, if every player in a tail-drop game* Γ^{TD} *plays according to the GIGA procedure, then the average strategy vector will converge to a NE and the payoff of each player will converge to her payoff at that NE.*
Furthermore, if every player runs GIGA with $\eta^t = 1/\sqrt{t}$*, then at every time* $t > t^*$*, where* $t^* = O(n^4)$*,*

(i) The average profile of actions will be an ϵ*-NE, where* $\epsilon = O(\frac{1}{\sqrt{t}})$

(ii) The average payoff of each player will differ from its payoff at that NE by at most $O(\frac{1}{\sqrt{t}})$*.*

The main step in the proof of Theorem 6.3 is to show that after sufficient number of time steps the players total offered load is always greater than the channel capacity.

Lemma 6.4. *Assuming there are at least* $n \geq 4$ *players in the game, if every player in a tail-drop game* Γ^{TD} *plays according to the GIGA procedure, then there exists a time* t^* *such that* $S(x^t) > 1$ *for every* $t > t^*$*. Furthermore, if the learning rate* $\eta^t = \frac{1}{\sqrt{t}}$*, then* $t^* = O(n^4)$*.*

6.2 RED-like models

In random early discard the router drops packets as a function of the queue size. Since we do not have a queue size in our models, we model this by dropping fraction of the rate at each time step as a function of the offered load $S(x)$.

Assume that the router drops packets at a rate of $\beta S(x)$. This implies that user i would have an effective bandwidth of $x_i(1 - \beta S(x))$, and the total effective bandwidth is $S(x)(1 - \beta S(x))$. Since the capacity of the link is $C = 1$ we need that $S(x)(1 - \beta S(x)) \leq 1$, which always holds for $\beta \geq 1/4$ (also, we can not dropping more than $S(x)$ units of rate, so $\beta \leq 1$).

In this case, for $\beta \geq 1/4$ we derive the RED utility function,

$$u_i^{\mathrm{RED}}(x) = \begin{cases} x_i(1 - \beta S(x)) - \alpha_i \beta x_i S(x) & S(x) \leq 1/\beta \\ -\alpha_i x_i & S(x) > 1/\beta \end{cases} \tag{6.1}$$

Notice that $u_i^{\mathrm{RED}}(1/\beta, x_{-i}) > u_i^{\mathrm{RED}}(x_i, x_{-i})$, for every $x_i > \frac{1}{\beta}$, $x_{-i} \in \mathbb{R}_+^{n-1}$, since $\frac{\partial}{\partial x_i} u^{\mathrm{RED}}(x) < 0$ for every such x. We therefore set a player's strategy space to $[0, \frac{1}{\beta}]$. The game associated with the RED policy will be denoted by $\Gamma^{\mathrm{RED}} = (N, ([0, \frac{1}{\beta}])_{i\in N}, (u_i^{\mathrm{RED}})_{i\in N}])$

Theorem 6.5. *If every player in a* Γ^{RED} *game plays according to the GIGA procedure, then the average strategy vector will converge to a NE and the payoff of each player will converge to her payoff at that NE.*

(i) The average profile of actions will be an ϵ*-NE, where* $\epsilon = O(\frac{1}{\sqrt{t}})$

(ii) The average payoff of each player will differ from its payoff at that NE by at most $O(\frac{1}{\sqrt{t}})$.

Furthermore, if every player runs GIGA with $\eta^t = 1/\sqrt{t}$, then at every time $t > t^*$, where $t^* = O\left(\left(\frac{n}{\alpha_{\min}}\right)^2\right)$ and $\alpha_{\min} = \min_{i \in N} \alpha_i$,

References

[1] S. Arora and W. Brinkman. A randomized online algorithm for bandwidth utilization. In *SODA*, 2002.

[2] Peter L. Bartlett, Elad Hazan, and Alexander Rakhlin. Adaptive online gradient descent. In *Advances in Neural Information Processing Systems (NIPS)*, 2008.

[3] A. Blum, E. Even-Dar, and K. Ligett. Routing without regret: On convergence to nash equilibria of regret-minimizing algorithms in routing games. In *PODC*, pages 45–52, 2006.

[4] A. Blum, K. Ligett M. Hajiaghay and, and A. Roth. Regret minimization and the price of total anarchy. In *STOC*, pages 373–382, 2008.

[5] A. Blum and Y. Mansour. From external to internal regret. *Journal of Machine Learning Research (JMLR)*, 8:1307–1324, 2007. (Preliminary version in COLT 2005.).

[6] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[7] N. Cesa-Bianchi and G. Lugosi. Potential-based algorithms in on-line prediction and game theory. *Machine Learning*, 51(3):239–261, 2003.

[8] A. A. Cournot. Recherches sur les principes mathmatiques de la thorie des richesses.(english translation: Researches into the mathematical principles of the theory of wealth). 1838.

[9] Abie Flaxman, Adam Tauman Kalai, and Brendan McMahan. Online convex optimization in the bandit setting: Gradient descent without a gradient. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 385–394, 2005.

[10] S. Floyd and Jacobson V. Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*, 1:397–413, 1993.

[11] D. Foster and S. M. Kakade. Deterministic calibration and nash equilibrium. In *COLT*, pages 33–48, 2004.

[12] D. Foster and R. Vohra. Regret in the on-line decision problem. *Games and Economic Behavior*, 21:40–55, 1997.

[13] D. Foster and H. P. Young. Learning, hypothesis testing, and Nash equilibrium. *Games and Economic Behavior*, 45:73–96, 2003.

[14] D. Foster and H. P. Young. Regret testing: Learning to play Nash equilibrium without knowing you have an opponent. *Theoretical Economics*, 1:341–367, 2006.

[15] Y. Freund and R. Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29:79103, 1999.

[16] Drew Fudenberg and David K. Levine. *The theory of learning in games*. MIT press, 1998.

[17] F. Germano and G. Lugosi. Global Nash convergence of Foster and Young's regret testing. *Games and Economic Behavior*, 60:135–154, 2007.

[18] B. Hajek and G. Gopalakrishnan. do greedy autonomous systems make for a sensible internet?, 2002. presented at the Conference on Stochastic Networks, Stanford University.

[19] S. Hart and Y. Mansour. The communication complexity of uncoupled Nash equilibrium procedures. In *STOC*, pages 345–353, 2007.

[20] S. Hart and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68:1127–1150, 2000.

[21] S. Hart and A. Mas-Colell. Stochastic uncoupled dynamics and Nash equilibrium. *Games and Economice Behavior*, 57(2):286–303, 2006.

[22] E. Hazan, A. Agarwal, and S. Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.

[23] R. Joahri and J. Tsitsiklis. Efficiency loss in resource allocation games. *Mathematics of Operations Research*, 29(3):407435, 2004.

[24] R. M. Karp, E. Koutsoupias, C. H. Papadimitriou, and S. Shenker. Optimization problems in congestion control. In *FOCS*, pages 66–74, 200.

[25] A. Mas-Colell, J. Green, and M. D. Whinston. *Microeconomic Theory*. Oxford University Press, 1995.

[26] J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*, 33(3):520–534, 1965.

[27] R. D. Theocharis. On the stability of the cournot solution on the oligopoly problem. *The Review of Economic Studies.*, 27(2):133–134, 1960.

[28] H. P. Young. *Strategic Learning and Its Limits*. Oxford University Press, 2004.

[29] M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *ICML*, pages 928–936, 2003.

[30] M. Zinkevich. Theoretical guarantees for algorithms in multi-agent settings. Technical Report CMU-CS-04-161,, Carnegie Mellon University, 2004.

A Proof from Section 2

Proof of Lemma 2.2. Fix a player $i \in N$. Following property A1, the second derivative

$$\frac{\partial}{\partial^2 x_i} \sum_{j \in N} \lambda_j u_j(x) \leq 0.$$

However, following property A2, for every $j \neq i$, $u_j(\cdot)$ is convex in x_i and therefore the second derivative $\frac{\partial}{\partial^2 x_i} \sum_{j \neq i} \lambda_j u_j(x) > 0$ and hence $\frac{\partial}{\partial^2 x_i} u_i(x)$ must be negative. Therefore, the payoff function u_i is thus concave in x_i. $\square$

B Proofs from Section 5

Proof of 5.3. In a linear resource allocation game, the utility of player i is

$$u_i(w) = \varphi_i(w) - w_i = \alpha_i M_i(w) - w_i = \alpha_i \frac{w_i^c}{\sum_{j \in N} w_j^c} - w_i,$$

for $0 < c \leq 1$. To show Property A1 in Definition 2.1 holds, set $\lambda_i = \frac{\frac{1}{\alpha_i}}{\sum_{j \in N} \frac{1}{\alpha_j}}$ observe that

$$g(w) = \sum_{i \in N} \lambda_i u_i(w) = \frac{1}{\sum_{j \in N} \frac{1}{\alpha_j}} \left(1 - \sum_{i \in N} \frac{w_i}{\alpha_i} \right),$$

is a linear function and thus concave.

To show that property A2 in Definition 2.1 holds, note that $f(x_{-i}) = u(x_i, x_{-i}) + w_i = \frac{x_i^c}{x_i^c + \sum_{j \in N, j \neq i} w_j^c}$ is convex iff $u(x_i, x_{-i})$ is convex in x_{-i}. Let x be a vector in $\mathbb{R}^{n-1}$, and consider the function $f(x) = g(h(x))$, where $h(x) = \sum_{i=1}^{n-1} {x_i}^c$ and $g(y) = \alpha/(\alpha + y)$ for some positive α. Note that h is concave function if $c \in (0, 1]$ and that $g(y)$ is a convex non increasing function. Therefore, we obtain that f is convex, cf. [6, pp. 87], and thus $u(x_i, x_{-i})$ is convex in x_{-i}. $\square$

Proof of Theorem 5.4. In equilibrium, each player bids $x_i^{\mathrm{NE}} = \frac{n-1}{n^2}$. Consider the dynamics where initially all players bid equally $x^0 \neq \frac{n-1}{n^2}$. Due to the symmetry, the best response dynamics would keep the bids of the players equal. Let x^t be the bids of the players time t (they are all identical), then

$$x^{t+1} = \sqrt{(n-1)x^t} - (n-1)x^t \tag{B.1}$$

Let ϵ^t be the difference between the players bids at time t and the equilibrium, *i.e.*, $\epsilon^t = x^t - \frac{n-1}{n^2}$. By substituting x^t by $\frac{n-1}{n^2} + \epsilon^t$ in Equation B.1 we get

$$\frac{n-1}{n^2} + \epsilon^{t+1} = \sqrt{(n-1)(\frac{n-1}{n^2} + \epsilon^t)} - (n-1)(\frac{n-1}{n^2} + \epsilon^t) \tag{B.2}$$

We will show that the sequence $|\epsilon^t|$ does not converge to zero. We can describe ϵ^{t+1} as a function of ϵ^t,

$$\begin{aligned}
\epsilon^{t+1} &= \sqrt{\left(\frac{n-1}{n}\right)^2 + \epsilon^t(n-1)} - \left(\frac{n-1}{n}\right)^2 - \frac{n-1}{n^2} - \epsilon^t(n-1) \\
&= \sqrt{\left(\frac{n-1}{n}\right)^2 + \epsilon^t(n-1)} - \sqrt{\left(\frac{n-1}{n}\right)^2} - \epsilon^t(n-1) \\
&= \frac{\epsilon^t(n-1)}{\sqrt{\left(\frac{n-1}{n}\right)^2 + \epsilon^t(n-1)} + \sqrt{\left(\frac{n-1}{n}\right)^2}} - \epsilon^t(n-1) \\
&= \epsilon^t(n-1)\left(\frac{1}{\sqrt{\left(\frac{n-1}{n}\right)^2 + \epsilon^t(n-1)} + \sqrt{\left(\frac{n-1}{n}\right)^2}} - 1\right).
\end{aligned}$$

For $\epsilon^t \in [0, 3\frac{n-1}{n^2}]$ we have that

$$\epsilon^{t+1} \leq \epsilon^t(n-1)\left(\frac{1}{3\frac{n-1}{n}} - 1\right) = \epsilon^t\frac{3-2n}{3},$$

and for $n \geq 4$ we have $\epsilon^{t+1} \leq -(5/3)\epsilon^t < 0$.

For $\epsilon^t \in [-(7/16)\frac{n-1}{n^2}, 0]$ we have that

$$\epsilon^{t+1} \geq \epsilon^t(n-1)\left(\frac{1}{(7/4)\frac{n-1}{n}} - 1\right) = \epsilon^t\frac{7-4n}{7},$$

and for $n \geq 4$ we have $\epsilon^{t+1} \geq -(9/7)\epsilon^t > 0$. This implies that $|\epsilon^t|$ does not converge to zero. $\square$

C Proof from Section 6

C.1 Analysis of the GIGA dynamics for the tail-drop game

Proof of 6.2. The utility function $q_i(\cdot)$ is now exactly the utility function of a resource allocation game, with proportional allocation mechanism, i.e., $c = 1$. Hence, the same arguments made in the proof of Lemma 5.3 show that $\Gamma^{\mathrm{TD}'}$ is a socially concave game. $\square$

Proof of Lemma 6.4. Let $\Delta_i^t(x) = \frac{\partial}{\partial x_i}u_i^{\mathrm{TD}}(x^{t-1})$ and let $\Delta(x) = \sum_{i\in N}\Delta_i(x)$. The proof proceeds in a number of steps.

Step I: *If x is such that $0 \leq S(x) < 1+\epsilon$, where $\epsilon = (1-\frac{2}{n+1})\alpha_{\min} - \frac{2}{n+1}$, then $\Delta(x) > 1$.*

First notice that following our assumption that $\alpha_{\min} \geq 3/(n-1)$,

$$\epsilon = (1-\frac{2}{n+1})\alpha_{\min} - \frac{2}{n+1} > (1-\frac{2}{n+1})\frac{3}{n-1} - \frac{2}{n+1} = \frac{1}{n+1} > 0.$$

Now, if $S(x) \leq 1$ then $\Delta(x) = n$ and the claim follows. Otherwise, $1 < S(x) < 1+\epsilon$, and in this case,

$$\begin{aligned}
\Delta(x) &= \sum_{i\in N}\left((1+\alpha_i)\frac{S(x)-x_i}{(S(x))^2}-\alpha_i\right) && \text{(C.1)}\\
&\geq \sum_{i\in N}\left((1+\alpha_{\min})\frac{S(x)-x_i}{(S(x))^2}-1\right) && \text{(C.2)}\\
&= \frac{(1+\alpha_{\min})(n-1)}{S(x)}-n && \text{(C.3)}\\
&> \frac{(1+\alpha_{\min})(n-1)}{(1+\alpha_{\min})\frac{n-1}{n+1}}-n && \text{(C.4)}\\
&= 1\,. && \text{(C.5)}
\end{aligned}$$

Thus, for every x such that $0 \leq S(x) < 1+\epsilon$, we have $\Delta(x) > 1$.

Step II: *For every time t there exists a time $t' > t$ such that $S(x^{t'}) \geq 1+\epsilon$.*

By definition $S(x^t) \geq 0$ for every x^t. In step I we show that $\Delta(x) > 1$ when $0 \leq S(x) < 1+\epsilon$. Now, combining with the fact that the learning rate is such that $\sum_{\tau=t'}^{t}\eta^\tau \to \infty$ as $t \to \infty$, the claim follows.

Step III: *There exists a time $\hat{t}$ such that $\eta^t\Delta(x^t) > -\epsilon$ for every $t > \hat{t}$.*

Notice that $\Delta(x^t)$ is bounded from below:

$$\Delta(x^t) = \sum_i\left((1+\alpha_i)\frac{S(x^t)-x_i^t}{(S(x^t))^2}-\alpha_i\right) > -\sum_i \alpha_i \geq -n,$$

where the last inequality follows our assumption that $\alpha_i \leq 1$, for every $i \in N$. By our assumption on the learning rate there exists at time $\hat{t}$ such that $\eta^{\hat{t}} \leq \epsilon/n$, and the claim follows.

Step IV: *If $t_1 > \hat{t}$ and $S(x^{t_1}) > 1$ then $S(x^t) > 1$ for every $t > t_1$.*

The proof is by induction on t. For the induction base $t = t_1$, the claim holds trivially. Assume the induction hypothesis holds for some $t > t_1$. If $S(x^t) \in (1, 1+\epsilon]$ then $\Delta(x^t) > 0$ and

$$S(x^{t+1}) = S(x^t) + \Delta(x^t)\eta^t \geq S(x^t) > 1,$$

Otherwise, $S(x^t) > 1+\epsilon$, but then $\eta^t\Delta(x^t) > -\epsilon$ and

$$S(x^{t+1}) = S(x^t) + \Delta(x^t)\eta^t \geq S(x^t) - \epsilon > 1.$$

and the claim follows.

Step V: *There exists a time t^* such that $S(x^t) > 1$ for every $t > t^*$.*

It follows from step III that after time $\hat{t}$, $\Delta(x^t) > -\epsilon$ for every $t > \hat{t}$. From step II it follows that there exists a time $t^* > \hat{t}$ such that $S(x^{t^*}) \geq 1+\epsilon$. And from step IV, it follows that for every $t > t^*$, $S(x^t) > 1$.

Step VI: *If $\eta^t = 1/\sqrt{t}$ then there exists $t^* = O(n^4)$, such that $S(x^t) > 1$ for every $t > t^*$.*

If $\eta^t = \frac{1}{\sqrt{t}}$, then $\eta^t \Delta(x^t) > -\epsilon$ for every $t > \hat{t}$ where $\hat{t} = (n/\epsilon)^2$, and

$$\begin{aligned} \frac{n}{\epsilon} &= \frac{n}{\frac{n-1}{n+1}\alpha_{\min} - \frac{2}{n+1}} && \text{(C.6)} \\ &\leq \frac{n}{\frac{3}{n+1} - \frac{2}{n+1}} && \text{(C.7)} \\ &= n(n+1) && \text{(C.8)} \\ &< 2n^2 && \text{(C.9)} \end{aligned}$$

Hence, $\hat{t} < 4n^4$. Note that if $\alpha_{\min} = \Omega(1)$ then $\epsilon = \Omega(1)$ and $\hat{t} = O(n^2)$.

Now, if $S(x^{\hat{t}}) > 1$, then $S(x^t) > 1$ for every $t > \hat{t}$, as it follows from step IV. If not, then set $t^* = \hat{t} + 4n^2$. If $S(x^t) > 1$ for some $\hat{t} < t \leq t^*$ then we are done. Otherwise, we have that $\Delta(S(x^t)) = n$ for every such t. In this case

$$S(x^{t^*}) = S(x^{\hat{t}}) + \sum_{t=\hat{t}}^{t^*-1} \frac{1}{\sqrt{t}} \Delta(x^t) \geq 0 + 4n^2 \frac{1}{\sqrt{4n^4 + 4n^2}} n > 2,$$

since $n \geq 4$. But, this is in contradiction to the assumption that $S(x^t) \leq 1$ for $\hat{t} < t \leq t^*$. Thus, at some time $t' < t^* = O(n^4)$, $S(x^{t'}) > 1$ and therefore, following Step IV, for every $t > t^*$, $S(x^t) > 1$. □

Proof of Theorem 6.3. By Lemma 6.4 there exists time t^*, such that $S(x^t) > 1$, for every $t > t^*$. Using Lemma 6.2, we obtain that from t^* on, the players are a posteriori playing the socially concave game, $\Gamma^{\mathrm{TD}'}$.

The regret of a player i at a time $t > t^*$ is at most

$$\mathcal{R}_i(t) = O(\sqrt{t - t^*}) + O(t^*) = O(t^* + \sqrt{t}),$$

where the first term is the regret accumulated after t^* and $O(t^*)$ is an upper bound on the difference between the utility from the best fixed transmission rate (+1) and the worst possible loss (-1).

Thus, following Theorem 3.1, we conclude that at time $t > t^*$, the average strategy profile is an $O(n/\sqrt{t})$-Nash equilibrium, and that the average payoff of each player differs by at most $O(n/\sqrt{t})$ from her payoff in that ϵ-NE. □

C.2 Analysis of the GIGA dynamics for the RED game

The game Γ^{RED} is not necessarily a socially concave game (for once, the utility of a player is generally not concave). However, when we alter the utility function so that it captures the first condition in 6.1, we get a socially concave game. We define here a new game denoted $\Gamma^{\mathrm{RED}'} = (N, ([0,1])_{i \in N}, (q_i^{\mathrm{RED}})_{i \in N})$, where

$$q_i^{\mathrm{RED}}(x) = x_i(1 - \beta S(x)) - \alpha_i \beta x_i S(x).$$

Lemma C.1. *The game $\Gamma^{\mathrm{RED}'}$ is a socially concave game.*

Proof of Lemma C.1. Fix x_i, and consider the function

$$f(z) = x_i(1 - \beta(x_i + z)) - \alpha_i \beta x_i (x_i + z).$$

$f(\cdot)$ is linear in z and consequently, $q_i^{\mathrm{RED}}(x)$ is linear in x_{-i}, as a composition of two linear functions.

Now, consider the function $g(\cdot)$ which is the sum of utilities,

$$g(x) = \sum_{i\in N} q_i^{\mathrm{RED}}(x) = \sum_{i\in N} x_i - \sum_{i\in N} \beta(1+\alpha_i)x_i S(x) = S(x) - S^2(x)\sum_{i\in N}\beta(1+\alpha_i).$$

The function $g(\cdot)$ can be composed as $g(x) = h(S(x))$, where $h(z) = z - z^2\sum_{i\in N}\beta(1+\alpha_i)$. Hence, $g(\cdot)$ is a concave function as composition of a concave function with a linear function.

Thus, $\Gamma^{\mathrm{RED}'}$ is socially concave, as both conditions A1 and A2 of Definition 2.1 hold. $\square$

Lemma C.2. *If every player in a Γ^{RED} game, plays according to the GIGA procedure, then there exists a time t^* such that for every $t > t^*$, $S(x^t) < 1/\beta$. Furthermore, if the learning rate $\eta^t = \frac{1}{\sqrt{t}}$, then $t^* = O((\frac{n}{\alpha_{min}})^2)$.*

Proof of Lemma C.2. Let $\Delta_i^t(x) = \frac{\partial}{\partial x_i}u_i^{\mathrm{RED}}(x^{t-1})$ and let $\Delta(x) = \sum_{i\in N}\Delta_i(x)$. The proof proceeds in a number of steps.

Step I: *If x is such that $\frac{1}{\beta(\alpha_{\min}+1)} \le S(x) < \frac{1}{\beta}$, then $\Delta(x) \le -1$.*

If $S(x) \le 1/\beta$ then

$$\frac{\partial}{\partial x_i}u_i^{\mathrm{RED}}(x^t) = 1 - (\beta+\alpha_i\beta)\sum_{j\neq i} x_j - 2(\beta+\alpha_i\beta)x_i = 1 - (\beta+\alpha_i\beta)(S(x)+x_i)$$

and

$$\begin{aligned}
\Delta(x) &= n - \sum_{i\in N}(\beta+\alpha_i\beta)(S(x)+x_i) && \text{(C.10)}\\
&\le n - (n+1)\beta(1+\alpha_{\min})S(x) && \text{(C.11)}\\
&\le n - (n+1)\beta(1+\alpha_{\min})\frac{1}{\beta(\alpha_{\min}+1)} && \text{(C.12)}\\
&= -1. && \text{(C.13)}
\end{aligned}$$

Step II: *For every time t there exists a time $t' > t$ such that $S(x^{t'}) \le 1/\beta$.*

For x such that $S(x) > 1/\beta$, we have $\frac{\partial}{\partial x_i}u_i^{\mathrm{RED}}(x^t) = -\alpha_i$, and $\Delta(x) = \sum_{i\in N} -\alpha_i < 0$. Combining with the fact that the learning rate is such that $\sum_{\tau=t'}^{t}\eta^\tau \to \infty$ as $t\to\infty$, the claim follows.

Step III: *There exists a time $\hat{t}$ such that $\eta^t\Delta(x^t) < \epsilon$ for every $t > \hat{t}$, where $\epsilon = \frac{1}{\beta} - \frac{1}{\beta(\alpha_{\min}+1)}$.*

If x^t is such that $S(x^t) > \frac{1}{\beta}$ then $\Delta(x^t) < 0$, as we showed in Step II. Otherwise, as observed clearly in Equation (C.10), $\Delta(x^t) < n$. By our assumption on the learning rate there exists at time $\hat{t}$ such that $\eta^{\hat{t}} \le \epsilon/n$, and the claim follows.

Step IV: *If $t_1 > \hat{t}$ and $S(x^{t_1}) < 1/\beta$ then $S(x^t) < 1/\beta$ for every $t > t_1$.*

The proof is by induction on t. For the induction base $t = t_1$, the claim holds trivially. Assume the induction hypothesis holds for some $t > t_1$. If $S(x^t) \in [\frac{1}{\beta(1+\alpha_{\min})}, \frac{1}{\beta}]$ then $\Delta(x^t) < -1$ and

$$S(x^{t+1}) = S(x^t) + \Delta(x^t)\eta^t \le S(x^t) < \frac{1}{\beta},$$

Otherwise, $S(x^t) < \frac{1}{\beta(1+\alpha_{\min})}$, but then $\eta^t \Delta(x^t) < \epsilon$ and

$$S(x^{t+1}) = S(x^t) + \Delta(x^t)\eta^t < \frac{1}{\beta(1+\alpha_{\min})} + \left(\frac{1}{\beta} - \frac{1}{\beta(1+\alpha_{\min})}\right) = \frac{1}{\beta}.$$

and the claim follows.

Step V: *There exists a time t^* such that $S(x^t) > 1$ for every $t > t^*$.*

It follows from step III that $\Delta(x^t) < \epsilon$, for every $t > \hat{t}$. From step II it follows that there exists a time $t^* > \hat{t}$ such that $S(x^{t^*}) < \frac{1}{\beta}$. And, from step IV, it follows that for every $t > t^*$, $S(x^t) < \frac{1}{\beta}$. This completes the proof of the first part of the claim.

Step VI: *If $\eta^t = 1/\sqrt{t}$ then there exists $t^* = O(n^2(\frac{1}{\alpha_{\min}})^2)$, such that $S(x^t) > 1$ for every $t > t^*$.*

Since $\eta^t = \frac{1}{\sqrt{t}}$, then $\eta^t \Delta(x^t) < \epsilon$ for every $t > \hat{t}$ where

$$\begin{aligned}
\hat{t} &= (n/\epsilon)^2 && \text{(C.14)}\\
&= \left(\frac{n}{\frac{1}{\beta} - \frac{1}{\beta(1+\alpha_{\min})}}\right)^2 && \text{(C.15)}\\
&= \left(\frac{n}{\alpha_{\min}}\beta(1+\alpha_{\min})\right)^2 && \text{(C.16)}
\end{aligned}$$

Now, if $S(x^{\hat{t}}) < 1/\beta$, then $S(x^t) < 1/\beta$ for every $t > \hat{t}$, as it follows from step IV. If not, then set $t^* = 64(n/\epsilon)^2 + 1$. If $S(x^t) < 1/\beta$ for some $\hat{t} < t \le t^*$ then we are done. Otherwise, we have that $\Delta(S(x^t)) < -n \cdot \alpha_{\min}$ for every such t. Recall that no player ever transmits with a rate greater than $\frac{1}{\beta}$, hence at time $\hat{t}$, $S(x^{\hat{t}}) \le \frac{n}{\beta} \le 4n$. Therefore,

$$\begin{aligned}
S(x^{t^*}) &= S(x^{\hat{t}}) + \sum_{t=\hat{t}}^{t^*-1} \frac{1}{\sqrt{t}}\Delta(x^t) && \text{(C.17)}\\
&\le 4n + (\sum_{t=\hat{t}}^{t^*-1} \frac{1}{\sqrt{t}})(-n\alpha_{\min}) && \text{(C.18)}\\
&\le 4n + (\sum_{t=\hat{t}}^{t^*-1} \frac{1}{\sqrt{t^*-1}})(-n\alpha_{\min}) && \text{(C.19)}\\
&= 4n - (t^* - \hat{t})(n\alpha_{\min})\frac{1}{\sqrt{t^*}} && \text{(C.20)}\\
&= 4n - 63(n/\epsilon)^2(n\alpha_{\min})\frac{1}{8(n/\epsilon)} && \text{(C.21)}\\
&= 4n - \frac{63}{8}n^2(1+\alpha_{\min})\beta && \text{(C.22)}\\
&= 4n - \frac{63}{32}n^2(1+\alpha_{\min}) && \text{(C.23)}\\
&\overset{(a)}{\le} \frac{1}{8} && \text{(C.24)}\\
&< \frac{1}{\beta}, && \text{(C.25)}
\end{aligned}$$

where (a) holds for every $n \geq 2$, which is our case. But, this is in contradiction to the assumption that $S(x^t) \geq \frac{1}{\beta}$ for every $\hat{t} < t \leq t^*$. Thus, at some time t, $\hat{t} < t \leq t^*$, $S(x^t) \leq \frac{1}{\beta}$ and therefore, following Step IV, for every $t > t^*$, $S(x^t) > 1$. This completes the proof of the second part of the claim. □

Proof of Theorem 6.5. From Lemma C.2 we learn that after some time t^*, the profile of actions is such that $S(x^t) \leq \frac{1}{\beta}$, for every $t > t^*$. It follows Lemma C.1, that from t^* on, the players are a posteriori playing the socially concave game, $\Gamma^{\mathrm{TD}'}$.

The regret of a player i at a time $t > t^*$ is at most

$$\mathcal{R}_i(t) = O(\sqrt{t - t^*}) + O(t^*) = O(t^* + \sqrt{t}),$$

where the first term is the regret accumulated after t^* and $O(t^*)$ is an upper bound on the difference between the utility from the best fixed transmission rate (+1) and the worst possible loss (-1).

Thus, following Theorem 3.1, we conclude that at time $t > t^*$, the average strategy profile is an $O(n/\sqrt{t})$-Nash equilibrium, and that the average payoff of each player differs by at most $O(n/\sqrt{t})$ from her payoff in that ϵ-NE. □