Fictitious Play in Opinion Formation Games with Random Payoffs

Dimitris Fotakis[1,2], Vardis Kandiros[2], Vasilis Kontonis[2], and Stratis Skoulakis[2]

[1] Yahoo Research, New York, NY, USA.
fotakis@oath.com
[2] National Technical University of Athens, Greece.
fotakis@cs.ntua.gr, vkonton@gmail.com, vkandiros@yahoo.gr, sskoul@gmail.com

Abstract. We study opinion formation games based on the famous model proposed by Friedkin and Johsen. In today's huge social networks the assumption that in each round all agents update their opinions by taking into account the opinions of *all* their friends could be unrealistic. Therefore, we assume that in each round each agent gets to meet with only one random friend of hers. Since it is more likely to meet some friends than others we assume that agent i meets agent j with probability p_{ij}. Specifically, we define an opinion formation game, where at round t, agent i with intrinsic opinion $s_i \in [0,1]$ and expressed opinion $x_i(t) \in [0,1]$ meets with probability p_{ij} neighbor j with opinion $x_j(t)$ and suffers cost that is a convex combination of $(x_i(t) - s_i)^2$ and $(x_i(t) - x_j(t))^2$.
For a dynamics in the above setting to be considered as natural it must be simple, converge to the equilibrium x^*, and perhaps most importantly, it must be a reasonable choice for selfish agents. In this work we show that *fictitious play*, is a natural dynamics for the above game. We prove that, after $O(1/\varepsilon^2)$ rounds, the opinion vector is within error ε of the equilibrium. Moreover, we show that fictitious play admits no-regret and thus, is a reasonable algorithm for agents to implement.
The classical Friedkin-Johsen dynamics converges to the equilibrium within error ε after only $O(\log(1/\varepsilon))$ rounds whereas in our setting fictitious play needs $\widetilde{O}(1/\varepsilon^2)$ rounds. A natural question is whether there exists a simple dynamics for our problem with better rate of convergence. We answer this question in the negative by showing that no-regret algorithms cannot converge with less than $\mathrm{poly}(1/\varepsilon)$ rounds. Interestingly, we show that when agents meet a neighbor *uniformly* at random, there exists a dynamics that only needs $O(\log(1/\varepsilon)^2)$ rounds, resembling the convergence rate of the original Friedkin-Johsen model.

Keywords: Opinion Dynamics · No Regret · Fictitious Play

1 Introduction

1.1 Friedkin-Johsen Model and Opinion Formation Games

In the Friedkin-Johsen model (FJ-model), there exist a set of n agents. Each agent i poses an internal opinion $s_i \in [0,1]$ and a self confidence coefficient $\alpha_i \in (0,1]$. At each round $t \geq 1$, agent i *publicly* expresses the opinion $x_i(t)$ that is defined as follows:

$$x_i(t) = (1-\alpha_i)\sum_{j\neq i} p_{ij}x_j(t-1) + \alpha_i s_i \tag{1}$$

where $p_{ij} \in [0,1]$ is the influence that j has imposes on i and it is assumed that $\sum_{j\neq i} p_{ij} = 1$. The simplicity of the update rule (1) makes the FJ-model plausible, because in real social networks it is very unlikely that agents change their opinions according to complex rules.

Based on the FJ-model, in [BKO11], they propose an *opinion formation game* in which the strategy that each agent i adopts, is the opinion $x_i \in [0,1]$ that she publicly expresses incurring her a cost

$$C_i(x_i, x_{-i}) = (1-\alpha_i)\sum_{j\neq i} p_{ij}(x_i - x_j)^2 + \alpha_i(x_i - s_i)^2. \tag{2}$$

In [**?**] they prove that it always admits a unique Nash Equilibrium x^* and they provide bound for the Price of Anarchy. It is easy to see that if each agent i at round $t \geq 1$ updates her opinion so as to minimize her individual cost defined in (2), the resulting dynamics (*best response dynamics*) is the dynamics defined by the FJ-model. We denote an instance of the above game as (P, s, α) where $P_{n\times n}$ is a stochastic matrix with $P_{ij} = p_{ij}$ and zero diagonal entries, s is the initial opinion vector and α the self confidence vector. In [GS14] it is proved that for any instance $I = (P, s, a)$ of the FJ-model, the opinion vector $x(t) = (x_1(t), \ldots, x_n(t))$ converges to the unique equilibrium point x^*. Meaning that the agents can adopt a simple but more importantly a *rational* update rule, in the sense that they minimize their individual cost, that ensures fast convergence to the equilibrium point.

1.2 Opinion Formation Games with Random Payoffs

In the game defined by (2), agent's i cost $C_i(x_i, x_{-i})$ is a deterministic function of the opinion vector x. Many recent works (see e.g. [ZLZ17]) study games with random payoffs, that is agent's i cost $(C_i(x_i, x_{-i}))$ is a random variable. The random payoff setting can be much more realistic, since randomness may naturally occur because of incomplete information, noise or other stochastic factors. Motivated by this line of research we introduce a random payoff variant of the opinion formation game (2).

Definition 1. *Let $I = (P, s, a)$ an instance of the opinion formation game and x the opinion vector. Each agent i,*

- *picks another agent j with probability p_{ij}*
- *suffers cost $C_i(x_i, x_{-i}) = (1-a_i)\sum_{j\in N_i}(x_i - x_j)^2 + a_i(x_i - s_i)^2$*

This game is more compatible to a realistic setting since in real world networks (e.g. Facebook, Twitter e.t.c.), each agent may have several hundreds of friends. As a result it is far more reasonable to assume that every day, each agent meets a random small subset of her acquaintances and suffers a cost based on how much she disagrees with them. In the opinion formation game with random payoffs we can define the Pure Nash Equilibrium with respect to the expected cost of each agent. More precisely, an opinion vector x^* is a PNE if and only if for each agent i,

$$\mathbf{E}\left[C(x_i, x_{-i})\right] \leq \mathbf{E}\left[C(x_i', x_{-i})\right], \text{ for all } x_i' \in [0,1]$$

Since for a given opinion vector x, the expected cost of agent i is $\mathbf{E}\left[C(x_i, x_{-i})\right] = (1-a_i)\sum_{j\in N_i}(x_i - x_j)^2 + a_i(x_i - s_i)^2$ the results in [BKO11] apply, meaning that a unique PNE x^* always exists. However, the *best response dynamics* in the random payoff game does not converge to the equilibrium. In this work, we investigate whether there exists a natural dynamical process that leads the system to the unique equilibrium point x^*.

1.3 Our Results and Techniques

In this work we propose Algorithm 1 as a dynamics for the random payoff game of Definition 1. Note that the dynamics described in Algorithm 1 is the *fictitious play* in the game defined by the instance $I = (P, s, a)$. Generally speaking *fictitious play* does not guarantee convergence to the equilibrium. In Section 3 we show

Algorithm 1 Fictitious Play

1: Initially, each agent i adopts the opinion s_i.
2: At round $t \geq 1$, each agent i meets another agent j with probability p_{ij}.
3: Suffers cost $C_i^t(x_i(t-1), x_j(t-1)) = (1-a_i)(x_i(t-1) - x_j(t-1))^2 + a_i(x_i(t-1) - s_i)^2$
4: Updates opinion as follows:

$$x_i(t) = \operatorname*{argmin}_{x \in [0,1]} \sum_{\tau=1}^{t} C_i^\tau(x, x_j(\tau-1)) \tag{3}$$

that, for our game with random payoffs, fictitious play converges to equilibrium, with the following rate.

Theorem 1. *Let $I = (P, s, \alpha)$ be any instance of the opinion formation game of Definition 1 with equilibrium $x^* \in [0,1]^n$. The opinion vector $x(t)$ produced by Algorithm 1 after t rounds satisfies*

$$\mathbf{E}\left[\|x(t) - x^*\|_\infty\right] \leq C\sqrt{\log n}\frac{(\log t)^2}{t^{\min(1/2,\rho)}},$$

where $\rho = \min_{i \in V} a_i$ and C is a universal constant.

The update rule (3) guarantees convergence while vastly reducing the information exchange between the agents at each round. Namely, each agent i learns the opinion of only one agent whereas in the classical FJ-model (1), agent i must learn the opinions of all the agents j with with $p_{ij} > 0$. In terms of total communication needed to get within distance ε of the equilibrium, the update rule (3) needs $O(n \log n)$ communication while (1) needs $O(|E|)$, where E is the set of positive entries in matrix P. Of course for this difference to be significant we need each agent to have at least $O(\log n)$ friend (agents j with $p_{ij} > 0$). A large social network like Facebook has approximately 2 billion users and each user has usually more that 100 friends which far more than $\log(2\ 10^9)$.

Apart from converging to the equilibrium, our update rule (3) is also a rational behavioral assumption since it ensures *no-regret* for the agents. Having no-regret means that the average cost for each agent i after T rounds is close to the average cost that she would suffer by expressing any fixed opinion. This is a very important feature of our update rule because even players that selfishly will to minimize their incurred cost, could choose to play according to it. In Section 4 we show the following theorem.

Theorem 2. *For every instance $I = (P, s, a)$, for every agent i*

$$\sum_{\tau=1}^{t} C_i^\tau(x_i(\tau)) \leq min_{x \in [0,1]} \sum_{\tau=1}^{t} C_i^\tau(x) + O(\log t)$$

Even though our update rule (3) has the above desired properties, for a fixed instance $I = (P, s, \alpha)$ it only achieves convergence rate of $\widetilde{O}(1/t^{\min(\rho,1/2)})$, while the original FJ-model outperforms our update rule since it achieves convergence rate $O(e^{-\rho t})$ [GS14]. We investigate whether this gap is due to an inherent characteristic of the random payoff setting (i.e. learning the opinion of just one random neighbor vs learning all of them) or we could find another natural dynamics that converges exponentially fast. Perhaps suprisingly both of the above questions admit a negative answer. In section 5, we prove that exponential convergence cannot be acheived if agents want to ensure no-regret. An informal statement of the theorem is the following.

Theorem 3. *If each agent updates her opinion according to a no-regret algorithm A then for any $c > 0$ there exists an instance $I = (P, s, a)$ such that $\mathbf{E}\left[\|x_t - x^*\|_\infty\right] = \Omega(1/t^{1+c})$.*

To prove this we first show that the existence of a no-regret algorithm (that achieves this convergence rate to the equilibrium) implies the existence of an algorithm that uses i.i.d samples from a Bernoulli random variable $B(p)$ to estimate its success probability p with the same asymptotic error rate for all $p \in [0, 1]$. Then we prove that such an estimator does not exists(Theorem ??). However, this lower bound is not due to the reduced information exchange of the random payoff setting. Update rules that ensure no-regret to the agents admit *slow* convergence because they are totally ignorant of the influence matrix P. Update rules that may use some knowledge on the influence matrix P does not necessarily suffer this the forementioned lower bound. In Section 6, we present an update rule that using some information of the matrix P can acheive exponential convergence.

1.4 Related Work

Our work belongs to the line of work studying the seminal Friedkin-Jonhsen model [FJ90]. Bindel et al. in [BKO11] defined an opinion formation game based on the FJ-model and bounded the inefficiency of its equilibrium point with respect to the total disagreement cost. Subsequent work bounded the inefficiency of its equilibrium in variants of the latter game [BGM13,**?**,**?**,**?**]. In [GS14] they show that the convergence time depends on the spectral radius of the adjacency matrix of the graph G and provided bounds in special graph topologies. In [BFM16], a variant of the opinion formation game in which social relations depend on the expressed opinions, is studied. They prove that, the discretized version of the above game admits a potential function and thus best-response converges to the Nash equilibrium. Convergence results in other discretized variants of the FJ-model can be found in [YOA$^+$13,**?**].

In [FV97], [FS99], [SA] they prove that in a finite if each agent updated her mixed strategy according to a no-regret algorithm the resulting time-averaged distribution converges to Coarse Correlated Equilibrium. In the same spirit in [BEDL06] they proved that no-regret dynamics converge to NE in the case of congestion games. Later in [EMN09] they studied no regret dynamics in games with infinite strategy space. They proved that for a large class of games with concave utility function (socially concave games), the time-averaged strategy vector converges to the pure Nash equilibrium. More recent work investigate a stronger notion of convergence of no-regret dynamics. In [CHM17] they show that, in n-person finite generic games that admit unique Nash equilibrium, the strategy vector converges *locally* and exponentially fast to the PNE. They also provide conditions for *global* convergence. Our results fit in this line of research since we show that for a game with *infinite* strategy space, the strategy vector (not the time-averaged) converges to the unique Nash equilibrium.

2 Preliminaries

2.1 Notation

Let $G(V, E)$ be a graph. We denote by V the set of agents, and $N_i \subset V$ the set of neighbors of agent i. Let $n = |V|$. We denote by $x(t) \in [0, 1]^n$ the vector of opinions of the agents at round t and $x_i(t)$ its i-th coordinate. We denote by $Q([0, 1])$ the set of rationals in $[0, 1]$.

2.2 Online Convex Optimization and No Regret Algorithms

The *Online Convex Optimization* (OCO) framework can be seen as a game played between a player and an adversary. Let $K \subseteq \mathbf{R}^n$ be a convex set and a set of functions $\mathcal{F}$ defined over K. *At round* t,

1. *the agent chooses* $x_t \in K$.
2. *the adversary observes the* x_t *and selects a function* $f_t(x) \in \mathcal{F}$.
3. *the player receives cost* $f_t(x_t)$.

The goal of the player is to pick x_t based on the history $(f_1, \ldots, f_{t-1})$ in a way that minimizes the total cost. We emphasize that the agent has to choose x_t before seeing f_t, otherwise the problem becomes trivial.

Definition 2. *An OCO algorithm A, at round t selects a vector $x^t \in K$ according to the history, $x^t = A_t(f_1, \ldots, f_t)$.*

Definition 3. *Let the OCO algorithm A and the sequence of functions $\{f_1, \dots, f_T\} \in F^T$. The regret of A is defined by*

$$R_A(T) = \sum_{t=1}^{T} f_t(x_t) - min_{x \in K} \sum_{t=1}^{T} f_t(x)$$

If $R_A(T) = o(T)$ for any sequence $\{f_1, \dots, f_T\}$ then A is no-regret algorithm

According to the feasibility set K and the set of functions F, different no-regret algorithms with differnt regret bounds can be derived. Some seminal examples are Zinkevic's algorithm that achieves regret $R_A(T) = O\left(\sqrt{T}\right)$, when K is convex, closed and bounded and F is the set of convex functions with bounded first derivative. Hazan et al [] proposed an algorithm with $R_A(T) = O(\log T)$, when F is the set of twice diffentiable strongly convex functions.

No regret dynamics and repeated games: Agent i with s_i, a_i is the player. At round t, she adopts an opinion $x_i(t) \in [0,1]$ and then the adversary selects a function $C_i^t(x) = (1-a_i)(x-y_t)^2 + a_i(x-s_i)^2$, where $y^t \in [0,1]$. Let A a no-regret in the above OCO setting. Clearly, each agent i is willing to adopt as $x_i(t)$ the suggestion of A, ensuring that average cost is similar to the average cost of the best fixed opinion. As a result if all agents adopt the no-regret algorithm A, for a fixed instance I of the game the following dynamical process is defined.

Algorithm 2 no-regret dynamics

Let an instance $I = (G, s, a)$ of the opinion formation game and $x_1(0), \dots, x_n(0)$ the initial opinions.
At round $t \geq 1$, each agent i:
1: Adopts an opinion $x_i(t) = A_t(C_i^1, \dots, C_i^{t-1})$
2: Meets uniformly at random one of her friends $j \in N_i$
3: Suffers cost $C_i^t(x_i(t))$, where $C_i^t(x) = (1-a_i)(x - x_j(t))^2 + a_i(x-s_i)^2$

We denote as $x_A(t)$ the opinion vector when the algorithm A is selected. In case $A_t(C_i^1, \dots, C_i^{t-1}) = \text{argmin}_{x \in [0,1]} \sum_{\tau=1}^{t-1} C_i^t(x)$ we obtain the *fictitious play* defined in 1 and for simplicity we denote it as $x(t)$. As we have already menioned, in Section 3 we bound the convergence rate of the *fictitious play*, $\mathbf{E}[\|x(t) - x^*\|_\infty]$ and in Section 4, we prove that it is no-regret algorithm for the specific OCO setting. The latter does not hold in more general OCO settings. In section 4, we investigate whether there exists an algorithm A with significantly better asymptotic rate of convergence.

3 Fictitious Play Convergence Rate

In this section we prove that fictitious play described as Algorithm 1 converges to the unique equilibrium x^*. The main result of the section is Theorem 8.

Theorem 4. *Let $I = (P, s, \alpha)$ be any instance of the opinion formation game of Definition 1 with equilibrium $x^* \in [0,1]^n$. The opinion vector $x(t)$ produced by Algorithm 1 after t rounds satisfies*

$$\mathbf{E}[\|x(t) - x^*\|_\infty] \leq C\sqrt{\log n}\frac{(\log t)^{3/2}}{t^{\min(1/2, \rho)}},$$

where $\rho = \min_i a_i$ and C is a universal constant.

At first we present the high level idea of the proof of Theorem 8. According to Algorithm 1, each agent i at round $t \geq 1$ updates her opinion as $x_i(t) = \text{argmin}_{x \in [0,1]} \sum_{\tau=1}^{t} C_i^\tau(x, x_{W_i^\tau}(\tau - 1))$ where W_i^τ is the random variable denoting the agent j that i met at round τ ($\mathbf{P}[W_i^\tau = j] = p_{ij}$). The above update rule can be written equivalently as:

$$x_i(t) = (1-\alpha_i)\frac{\sum_{\tau=1}^{t} x_{W_i^\tau}(\tau - 1)}{t} + \alpha_i s_i$$

Since we are interested in bounding the $\mathbf{E}\left[\|x(t)-x^*\|_\infty\right]$, we can use the fact $x_i^* = (1-\alpha_i)\sum_{j\neq i} x_j^* + \alpha_i s_i$ to bound $|x_i(t)-x_i^*|$ as follows:

$$\begin{aligned}
|x_i(t)-x_i^*| &= (1-\alpha_i)\left|\frac{\sum_{\tau=1}^{t} x_{W_i^\tau}(\tau-1)}{t} - \sum_{j\neq i} p_{ij}x_j^*\right| \\
&= (1-\alpha_i)\left|\sum_{j\neq i}\frac{\sum_{\tau=1}^{t}\mathbf{1}[W_i^\tau=j]x_j(\tau-1)}{t} - \sum_{j\neq i} p_{ij}x_j^*\right| \\
&\leq (1-\alpha_i)\sum_{j\neq i}\left|\frac{\sum_{\tau=1}^{t}\mathbf{1}[W_i^\tau=j]x_j(\tau-1)}{t} - p_{ij}x_j^*\right|
\end{aligned}$$

Now assume that $\left|\frac{\sum_{\tau=1}^{t}\mathbf{1}[W_i^\tau=j]}{t} - p_{ij}\right|$ were 0 for all $t\geq 1$, then with simple algebraic manipulations we can prove that $\|x(t)-x^*\|_\infty \leq e(t)$ where $e(t)$ satisfies the recursive equation $e(t) = (1-\rho)\frac{\sum_{\tau=0}^{t-1} e(\tau)}{t}$. It follows that $\|x(t)-x^*\|_\infty \leq 1/t^\rho$ meaning that $x(t)$ converges to x^*. Obviously $\left|\frac{\sum_{\tau=1}^{t}\mathbf{1}[W_i^\tau=j]}{t} - p_{ij}\right| \neq 0$ and the above analysis does not hold. In Lemma 1 we use the fact that $\left|\frac{\sum_{\tau=1}^{t}\mathbf{1}[W_i^\tau=j]}{t} - p_{ij}\right|$ tends to 0 with probability 1 (W_i^τ are independent random variables) to obtain a similar recursive relation for $e(t)$. Then in Lemma 2 we upper bound the solution of this recursive equation.

Lemma 1. *Let $e(t)$ the solution of the following recursion,*

$$e(t) = \delta(t) + (1-\rho)\frac{\sum_{\tau=0}^{t-1} e(\tau)}{t}$$

where $e(0) = \|x(0)-x^*\|_\infty$ and $\delta(t) = \sqrt{\frac{\ln(\pi^2 n t^2/6p)}{t}}$. Then,

$$\mathbf{P}\left[\text{for all } t\geq 1,\ \|x(t)-x^*\|_\infty \leq e(t)\right] \geq 1-p$$

Proof. We remind that W_i^τ denotes the agent j that i met at round τ and that this happens with probability p_{ij} and x^* is the unique equilibrium point of the instance $I=(P,s,\alpha)$. At first we prove that with probability at least $1-p$, for all $t\geq 1$ and all agents i:

$$\left|\frac{\sum_{\tau=1}^{t} x_{W_i^\tau}^*}{t} - \sum_{j\neq i} p_{ij}x_j^*\right| \leq \delta(t) \tag{4}$$

where $\delta(t) = \sqrt{\frac{\log(\pi^2 n t^2/(6p))}{t}}$.

Since W_i^τ are independent random variables with $\mathbf{P}\left[W_i^\tau = j\right] = p_{ij}$ and $\mathbf{E}\left[x_{W_i^\tau}^*\right] = \sum_{j\neq i} p_{ij}x_j^*$. By the Hoeffding's inequality we get

$$\mathbf{P}\left[\left|\frac{\sum_{\tau=1}^{t} x_{W_i^\tau}^*}{t} - \sum_{j\neq i} p_{ij}x_j^*\right| > \delta(t)\right] < 6p/(\pi^2 n t^2).$$

To bound the probability of error for all rounds $t=1$ to ∞ and all agents i, we apply the union bound

$$\sum_{t=1}^{\infty}\mathbf{P}\left[\max_i\left|\frac{\sum_{\tau=1}^{t} x_{W_i^\tau}^*}{t} - \sum_{j\neq i} p_{ij}x_j^*\right| > \delta(t)\right] \leq \sum_{t=1}^{\infty}\frac{6}{\pi^2}\frac{1}{t^2}\sum_{i=1}^{n}\frac{p}{n} = p$$

As a result with probability $1-p$ we have that for all $t\geq 1$ and all agents i,

$$\left|\frac{\sum_{\tau=1}^{t} x_{W_i^\tau}^*}{t} - \sum_{j\neq i} p_{ij}x_j^*\right| \leq \delta(t) \tag{5}$$

Now we can prove our claim by induction. Assume that $\|x^\tau - x^*\|_\infty \le e(\tau)$ for all $\tau \le t-1$. Then

$$
\begin{aligned}
x_i(t) &= (1-\alpha_i)\frac{\sum_{\tau=1}^t x_{W_i^\tau}(\tau-1)}{t} + \alpha_i s_i \\
&\le (1-\alpha_i)\frac{\sum_{\tau=1}^t x^*_{W_i^\tau} + \sum_{\tau=1}^t e(\tau-1)}{t} + \alpha_i s_i \qquad (6)\\
&\le (1-\alpha_i)\left(\frac{\sum_{\tau=1}^t x^*_{W_i^\tau}}{t} + \frac{\sum_{\tau=0}^{t-1} e(\tau)}{t}\right) + \alpha_i s_i \\
&\le (1-\alpha_i)\left(\sum_{j\in N_i} p_{ij}x_j^* + \delta(t) + \frac{\sum_{\tau=0}^{t-1} e(\tau)}{t}\right) + \alpha_i s_i \qquad (7)\\
&\le x_i^* + \delta(t) + (1-\alpha)\left(\frac{\sum_{\tau=0}^{t-1} e(\tau)}{t}\right)
\end{aligned}
$$

We get (6) from the induction step and (7) from inequality (5). Similarly, we can prove that $x_i(t) \ge x_i^* - \delta(t) - (1-\alpha)\frac{\sum_{\tau=1}^t e(\tau)}{t}$. As a result $\|x(t) - x^*\|_\infty \le e(t)$ and the induction is complete. Therefore, we have that with probability at least $1-p$, $\|x(t)-x^*\|_\infty \le e(t)$ for all $t \ge 1$.

Lemma 2. *Let $e(t)$ be a function satisfying the recursion*

$$e(t) = \delta(t) + (1-\alpha)\frac{\sum_{\tau=0}^{t-1} e(\tau)}{t} \text{ and } e(0) = \|x(0) - x^*\|_\infty,$$

where $\delta(t) = \sqrt{\frac{\ln(Dt^{2.5})}{t}}$, $\delta(0) = 0$, and $D > \mathrm{e}^{2.5}$ is a positive constant. Then $e(t) \le \sqrt{2.5\ln(D)}\frac{(\ln t)^{3/2}}{t^{\min(\rho,\,1/2)}}$.

The proof of Theorem 8 folows by direct application of Lemma 1 and 2.

4 Fictitious Play is no-regret

In this section we consider the Online Convex Optimization problem that we defined in Subsection 2.2. For every agent i with a_i, s_i, the following OCO problem is defined. For $b \in [0,1]$ we define $C_b(z) = \alpha_i(z-s_i)^2 + (1-\alpha_i)(z-b)^2$. The feasibility set is $K = [0,1]$ and the set of functions that the adversary chooses from is $\mathcal{F}_i = \{C_b(z) : b \in [0,1]\}$. Since the functions of $\mathcal{F}_i$ are uniquely determined by b the adversary simply chooses a sequence b^t. We now show that fictitious play admits no-regret for this OCO problem. To simplify notation, since we have fixed an agent i, we drop the subscript i from the following, i.e. we denote α_i by α. Let z^t be the choice that A makes at time t.

Theorem 5. *Let $z^t = \operatorname{argmin}_{z\in[0,1]} \sum_{\tau=0}^{t-1} C^\tau(z)$ then*

$$\sum_{t=0}^T C^t(z^t) \le \min_{z\in[0,1]} \sum_{t=0}^T C^t(z) + O(\log T).$$

In order to prove this claim, we will first define a similar rule y^t that takes into account the function C^t for the "prediction" at time t. Intuitively, this guarantees that the rule admits no regret.

Lemma 3. *Let $y^t = \operatorname{argmin}_{z\in[0,1]} \sum_{\tau=0}^{t} C^\tau(z)$ then*

$$\sum_{t=0}^T C^t(y^t) \le \min_{z\in[0,1]} \sum_{t=0}^T C^t(z)$$

Proof. By definition of y^t, $\min_{z\in[0,1]} \sum_{t=0}^{T} C^t(z) = \sum_{t=0}^{T} C^t(y^T)$, so

$$\begin{aligned}
\sum_{t=0}^{T} C^t(y^t) - \min_{z\in[0,1]} \sum_{t=0}^{T} C^t(z) &= \sum_{t=0}^{T} C^t(y^t) - \sum_{t=0}^{T} C^t(y^T) \\
&= \sum_{t=0}^{T-1} C^t(y^t) - \sum_{t=0}^{T-1} C^t(y^T) \\
&\leq \sum_{t=0}^{T-1} C^t(y^t) - \sum_{t=0}^{T-1} C^t(y^{T-1}) \\
&= \sum_{t=0}^{T-2} C^t(y^t) - \sum_{t=0}^{T-2} C^t(y^{T-1})
\end{aligned}$$

Continuing in the same way, we get $\sum_{t=0}^{T} C^t(y^t) \leq min_{z\in[0,1]} \sum_{t=0}^{T} C^t(z)$.

Now we can derive some intuition for the reason that *fictitious play* admits no regret. Since the cost incurred by the sequence y^t is at most that of the best fixed strategy, we can compare the cost incurred by z^t with that of y^t. However, the functions in $\mathcal{F}_i$ are Lipschitz-continuous and more specifically quadratic. These functions are all "similar" to each other, so the extra function C^t that y^t takes as input doesn't change dramatically the minimum point of the sum. Thus, for each t the numbers z^t and y^t are quite close and as a result the difference in their cost must be quite small.

Lemma 4. *For all t, $C^t(z^t) \leq C^t(y^t) + 2\frac{1-\alpha}{t+1} + \frac{(1-\alpha)^2}{(t+1)^2}$.*

Proof. We first prove that for all t,

$$\left|z^t - y^t\right| \leq \frac{1-\alpha}{t+1}. \tag{8}$$

By definition $z^t = \alpha s + (1-\alpha)\frac{\sum_{\tau=0}^{t-1} b_\tau}{t}$ and $y^t = \alpha s + (1-\alpha)\frac{\sum_{\tau=0}^{t} b_\tau}{t+1}$.

$$\begin{aligned}
\left|z^t - y^t\right| &= (1-\alpha)\left|\frac{\sum_{\tau=0}^{t-1} b_\tau}{t} - \frac{\sum_{\tau=0}^{t} b_\tau}{t+1}\right| \\
&= (1-\alpha)\left|\frac{\sum_{\tau=0}^{t-1} b_\tau - tb^t}{t(t+1)}\right| \\
&\leq \frac{1-\alpha}{t+1}
\end{aligned}$$

The last inequality follows from the fact that $b_\tau \in [0,1]$. We now use inequality (8) to bound the difference $C^t(z^t) - C^t(y^t)$.

$$\begin{aligned}
C^t(z^t) &= \alpha(z^t - s)^2 + (1-\alpha)(z^t - y_t)^2 \\
&\leq \alpha(y^t - s)^2 + 2\alpha\left|y^t - s\right|\left|z^t - y^t\right| + \alpha\left|z^t - y^t\right|^2 \\
&\quad + (1-\alpha)(y_t - y_t)^2 + 2(1-\alpha)\left|y^t - y_t\right|\left|z^t - y^t\right| + (1-\alpha)\left|z^t - y^t\right|^2 \\
&\leq C^t(y^t) + 2\left|z^t - y^t\right| + \left|y^t - z^t\right|^2 \\
&\leq C^t(y^t) + 2\frac{1-\alpha}{t+1} + \frac{(1-\alpha)^2}{(t+1)^2}
\end{aligned}$$

Theorem 5 easily follows since

$$\begin{aligned}\sum_{t=0}^{T} C^t(z^t) &\leq \sum_{t=0}^{T} C^t(y^t) + \sum_{t=0}^{T} 2\frac{1-\alpha}{t+1} + \sum_{t=0}^{T} \frac{(1-\alpha)^2}{(t+1)^2} \\ &\leq \min_{z\in[0,1]} \sum_{t=0}^{T} C^t(z) + 2(1-\alpha)(\log T + 1) + (1-\alpha)\frac{\pi^2}{6} \\ &\leq \min_{z\in[0,1]} \sum_{t=0}^{T} C^t(z) + O(\log T)\end{aligned}$$

5 Lower Bound for no-regret Dynamics

As we have already discussed for any fixed instance I with $\rho \geq 1/2$, *fictitious play* achieves convergence rate $\mathbf{E}_I\left[\|x_A(t) - x^*\|_\infty\right] = O\left(1/\sqrt{t}\right)$, this rate is outperformed by the rate of the original *FJ model* convergence rate $\mathbf{E}_I\left[\|x(t) - x^*\|_\infty\right] = O\left(1/2^t\right)$. In this section we investigate whether there exists another no-regret algorithm A that the agents can select and ensures a better convergence rate to the equilibrium. We show the following

Theorem 6. *Let A a no-regret algorithm and let $x_A(t)$ the opinion vector defined in* **??**. *For any $c > 0$, there exists an instance I_A such that $\mathbf{E}_{I_A}\left[\|x_A(t) - x^*\|_\infty\right] = \Omega(1/t^{1+c})$.*

The above theorem states that rationality in selfish agents comes with the price of slow convergence to the equilibrium point.

Since we prove 6 using a reduction to an estimation problem we shall first briefly introduce some definitions and notation. Given t independent samples from a Bernoulli random variable $B(p)$ an estimator is an algorithm that takes these samples as inputs and outputs an answer in $[0, 1]$.

Definition 4. *An estimator sequence $(\theta_t)_{t=1}^{\infty}$ is a sequence of functions, $\theta_t : \{0,1\}^t \mapsto [0,1]$.*

Perhaps the first estimator that comes to one's mind is the *sample mean*, that is $\theta_t = (1/t)\sum_{i=1}^{t} X_i$. Of course for an estimator to be efficient we would like its answer to be close to the mean p of the Bernoulli that generated the samples. To measure the efficiency of an estimator we define the *risk* which corresponds to the expected loss of an estimator.

Definition 5. *For an estimator $\hat{\theta} = \{\hat{\theta}_t\}_{t=1}^{\infty}$ we define its convergence rate $R_p(\theta_t) = E_p[|\hat{\theta}_t(X_1,\dots,X_t) - p|]$, where*

$$E_p[|\hat{\theta}_t(X_1,\dots,X_t) - p|] = \sum_{(y_1,\dots,y_t)\in\{0,1\}^t} |\hat{\theta}_t(y_1,\dots,y_t) - p| \; p^{\sum_{i=1}^t y_i} \; (1-p)^{t-\sum_{i=1}^t y_i}$$

The quantity $E_p[|\hat{\theta}_t(Y_1,\dots,Y_t) - p|]$ is the expected distance of the estimated value $\hat{\theta}_t$ from the parameter p, when the distribution of the samples is $B(p)$. To simplify notation we will also denote it as $E_p[|\hat{\theta}_t - p|]$. The error rate $R_p(\theta_t)$ quantifies the rate of convergence of the estimated value $\hat{p} = \theta_t(Y_1,\dots,Y_t)$ to the real parameter p. Since p is unknown, any meaningful estimator $\hat{p}$ must guarantee that for all $p \in [0,1]$, $\lim_{t\to\infty} R_p(t) = 0$. For example, *sample mean* has error rate $R_p(t) \leq \frac{1}{2\sqrt{t}}$ for ant $p \in [0,1]$ and clearly satisfies the above requirement.

We show now that any no-regret algorithm A, achieving the convergence rate of Theorem **??**, can be used as an estimator of the parameter $p \in [0,1]$ of a Bernoulli random variable with the same asymptotic error rate. The reduction formally stated and in Lemma 5.

Lemma 5. *Let A be a no-regret algorithm A such that for all instances I, $\lim_{t\to\infty} t^{1+c}\mathbf{E}_I\left[\|x_A(t) - x^*\|_\infty\right] = 0$. Then there exists an estimator $\hat{\theta_A}$ such that for all $p \in [0,1]$,*

$$\lim_{t\to\infty} t^{1+c} R_p(\theta_t) = 0$$

We sketch here the main idea. For a full proof see Section B of the Appendix. For a given a $p \in [0,1]$, we construct an instance I_p such that $x_c^* = p$ for an agent c. Moreover, agent c must receive only values 1 or 0 with probability p and $1-p$ respectively. This can be easily done using the star graph $K_{1,2}$. The agent corresponding to the center node, c, has $\alpha_c = 0$ and whereas the leaf nodes have $a_{1,2} = 1$, $s_1 = 0$, $s_2 = 1$. It follows that the estimator θ_t with $\theta_A^t = x_A^c(t)$ has error $R_p(\theta_t) = \mathbf{E}_{I_p}\left[\|x_A(t) - x^*\|_\infty\right]$. Meaning that if A

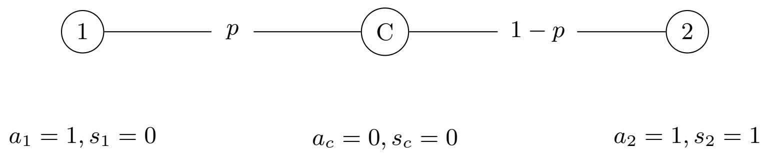

Fig. 1: The Lower Bound Instance

does not satisfy Theorem 6 then $\lim_{t\to\infty} t^{1+c} R_p(\theta_t) = 0$ for all $p \in [0,1]$. Thus we want to prove the following claim

Claim. For all sequences of estimators θ_t Then, there exists a fixed $p \in [0,1]$ such that

$$\lim_{t\to\infty} t^{1+c} \mathbf{E}_p\left[|\theta_t - p|\right] > 0.$$

The crucial point of Claim 5 is the fact that in order to construct the hard instance for the estimator we first inspect the sequence of estimators and then choose a $p \in [0,1]$ so that *all* estimators θ_t of the sequence have error rate $\omega(1/t^{1+c})$.

At this point we should mention that it is known that $\Omega(1/\varepsilon^2)$ samples are needed to estimate the parameter p of a Bernoulli random variable within additive error ε. Another well-known result is that taking the average of the samples is the *best* way to estimate the mean of a Bernoulli random variable. These results would indicate that the best possible rate of convergence for a no-regret dynamics would be $O(1/\sqrt{t})$. However, there is some fine print in these results which does not allow us to use them. In order to explain the various limitations of these methods and results we will briefly discuss some of them.

Perhaps the oldest sample complexity lower bound for estimation problems is the well-known Cramer-Rao inequality. Assuming that θ_t is a sequence of unbiased estimators, that is $\mathbf{E}\left[\theta_t\right] = p$ for all t, the Cramer-Rao lower bound for estimating the mean p of a Bernoulli random variable states that

$$\mathbf{E}_p\left[(\theta_t - p)^2\right] \geq \frac{p(1-p)}{t}. \tag{9}$$

In our setting, we can lower bound $\mathbf{E}_p\left[|\theta_t - p|\right]$ by $\mathbf{E}_p\left[(\theta_t - p)^2\right]$ since $|\theta_t - p| \leq 1$ (making the natural assumption that estimators do not output values outside $[0,1]$). Simply setting $p = 1/2$ in inequality (9) would give us a p satisfying the requirements of Claim 5. The problem with this lower bound is that the assumption that the estimator is unbiased is considered rather restrictive and unrealistic even in the statistics literature in the sense that many efficient practical estimators are not unbiased. Thus, we would like to get a lower bound with minimal assumptions about the estimator.

To the best of our knowledge, sample complexity lower bounds without assumptions about the estimator are given as lower bounds for the *minimax risk*, which was defined [3] by Wald in [Wal39] as

$$\inf_{\theta_t} \sup_{p\in[0,1]} \mathbf{E}_p\left[|\theta_t - p|\right].$$

Minimax risk captures the idea that after we pick the best possible algorithm, an adversary inspects it and picks the worst possible $p \in [0,1]$ to generate the samples that our algorithm will get as input. The methods

[3] Although the minimax risk is defined for any estimation problem and loss function, for simplicity, we write the minimax risk for estimating the mean of a Bernoulli random variable.

of Le'Cam, Fano, and Assouad are well-known information-theoretic methods to establish lower bounds for the minimax risk. For more on these methods see [Yu97,**?**] and the very good lecture notes of Duchi, [Duc]. As we stated before, it is well known that the minimax risk for the case of estimating the mean of a Bernoulli is lower bounded by $\Omega(1/\sqrt{t})$ and this lower bound can be established by Le Cam's method. In order to show why such arguments do no work for our purposes we shall sketch how one would apply Le Cam's method to get this lower bound. To apply Le Cam's method, one typically chooses two Bernoulli distributions whose means are far but their total variation distance is small. Le Cam showed that when two distributions are close in total variation then given a sequence of samples $X_1, \dots, X_t$ it is hard to tell whether these samples were produced by P_1 or P_2. The hardness of this *testing* problem implies the hardness of *estimating* the parameters of a family of distribution. For our problem the two distributions would be $B(1/2 - 1/\sqrt{t})$ and $B(1/2 + 1/\sqrt{t})$. It is not hard to see that their total variation distance is at most $O(1/t)$, which implies a lower bound $\Omega(1/\sqrt{t})$ for the minimax risk. The problem here is that the parameters of the two distributions depend on the number of samples t. The more samples the algorithm gets to see, the closer the adversary takes the 2 distributions to be. For our problem we would like to *fix* an instance and then argue about the rate of convergence of any algorithm on this instance. Namely, having an instance that depends on t does not work for us.

Trying to get a lower bound without assumptions about the estimators while respecting our need for a fixed (independent of t) p we prove Lemma 6. In fact, we show something stronger: for *almost all* $p \in [0, 1]$, any estimator $\hat{\theta}$ cannot achieve rate $o(1/t^{1+c})$. More precisely, suppose we select a p uniformly at random in $[0, 1]$ and run the estimator $\hat{\theta}$ with samples from the distribution $B(p)$, then with probability 1 the error rate $R_p(\theta_t) \in \Omega(1/t^{1+c})$. Although we do not show the sharp lower bound $\Omega(1/\sqrt{t})$ we prove that no exponential convergence rate is possible. Finally, while we do not preclude the possibility of similar results to exist in the statistics literature, we remark that our proof is fairly simple, intuitive, and could be of independent interest.

Lemma 6. *Let $\hat{\theta}$ an estimator for the parameter p of a Bernoulli random variable with error rate $R_p(\theta_t)$. If we select p uniformly at random in $[0, 1]$ then*

$$\mathbf{P}\left[\lim_{t\to\infty} t^{1+c} R_p(\theta_t) \ > 0\right] = 1,$$

for any $c > 0$.

Proof. Let an estimator $\hat{\theta} = \{\theta_t\}_{t=1}^{\infty}$, where $\theta_t : \{0, 1\}^t \mapsto [0, 1]$. The function θ_t can have at most 2^t different values. Without loss of generality we assume that θ_t takes the same value $\theta_t(x)$ for all $x \in \{0, 1\}^t$ with the same number of 1's. For example, $\theta_3(\{1, 0, 0\}) = \theta_3(\{0, 1, 0\}) = \theta_3(\{0, 0, 1\})$. This is due to the fact that for any $p \in [0, 1]$,

$$\sum_{0\le i\le t} \sum_{\|x\|_1 = i} |\theta_t(x) - p|\, p^i (1-p)^{t-i} \ge \sum_{0\le i\le t} \binom{t}{i} \left| \frac{\sum_{\|x\|_1=i} \theta_t(x)}{\binom{t}{i}} - p \right| p^i (1-p)^{t-i}.$$

For any estimator with error $R_p(\theta_t)$ there exists another estimator with $R'_p(t)$ that satisfies the above property and $R'_p(t) \le R_p(\theta_t)$. Thus we can assume that θ_t takes at most $t + 1$ different values. Let A denote the set of p for which the estimator has error rate $o(1/t^{1+c})$, that is

$$A = \{p \in [0, 1] : \lim_{t\to\infty} t^{1+c} R_p(\theta_t) = 0\}$$

We show that if we select p uniformly at random in $[0, 1]$ then $\mathbf{P}[p \in A] = 0$. We also define the set

$$A_k = \{p \in [0, 1] : \text{for all } t \ge k,\ t^{1+c} R_p(\theta_t) \le 1/2\}$$

Observe that if $p \in A$ then there exists t_p such that $p \in A_{t_p}$, meaning that $A \subseteq \cap_{k=1}^{\infty} A_k$. As a result,

$$\mathbf{P}[p \in A] \le \mathbf{P}\left[p \in \bigcup_{k=1}^{\infty} A_k\right] \le \sum_{k=1}^{\infty} \mathbf{P}[p \in A_k]$$

To complete the proof we show that $\mathbf{P}[p \in A_k] = 0$ for all k. Notice that $p \in A_k$ implies that for $t \geq k$, the estimator $\hat{\theta}$ must always have a value $\theta_t(i)$ close to p. Using this intuition we define the set

$$B_k = \{p \in [0,1] : \text{for all } t \geq k,\ t^{1+c} \min_{0 \leq i \leq t} |\theta_t(i) - p| < 1\}$$

We now show that $A_k \subseteq B_k$. Since $p \in A_k$ we have that for all $t \geq k$

$$t^{1+c} \min_{0 \leq i \leq t} |\theta_t(i) - p| \sum_{i=0}^{t} \binom{t}{i} p^i (1-p)^{t-i} \leq t^{1+c} \sum_{i=0}^{t} \binom{t}{i} |\theta_t(i) - p|\, p^i (1-p)^{t-i} = t^{1+c} R_p(\theta_t) \leq 1/2.$$

Thus, $\mathbf{P}[p \in A_k] \leq \mathbf{P}[p \in B_k]$. At first we write the set B_k in the following equivalent form $B_k = \cap_{t=k}^{\infty}\{p \in [0,1] :\ \min_{0 \leq i \leq t} |\theta_t(i) - p| \leq 1/t^{1+c}\}$. As a result,

$$\mathbf{P}[p \in B_k] \leq \mathbf{P}\left[\min_{0 \leq i \leq t} |\theta_t(i) - p| \leq 1/t^{1+c}\right], \text{for all } t \geq k$$

Each value $\theta_t(i)$ "covers" length $1/t^{1+c}$ from its left and right, and since there are at most $t+1$ such values we have for all $t \geq k$ the set

$$\{p \in [0,1]\ :\ \min_{0 \leq i \leq t} |\theta_t(i) - p| \leq 1/t^{1+c}\} = \bigcup_{i=0}^{t} \left(\theta_t(i) - \frac{1}{t^{1+c}},\ \theta_t(i) + \frac{1}{t^{1+c}}\right).$$

For each interval in the above union we have that $\mathbf{P}\left[|\theta_t(i) - p| \leq 1/t^{1+c}\right] \leq 2/t^{1+c}$ and by the union bound we get $\mathbf{P}[p \in B_k] \leq 2(t+1)/t^{1+c}$, for all $t \geq k$. We conclude that $\mathbf{P}[p \in B_k] = 0$.

6 A Graph Aware Update Rule

In Section 5 we saw that in our *imperfect information* for any no-regret algorithm $\mathcal{A}$ there exists a graph G such that the corresponding dynamics needs $\text{poly}(1/\varepsilon)$ rounds to achieve error ε. However, in the *perfect information* FJ-model there exists a simple update rule that requires only $\log(1/\varepsilon)$ rounds. Notice that our lower bound crucially depends on the fact that $\mathcal{A}$ is no-regret. At this point, a natural question is whether, this exponential gap is a generic restriction of our imperfect information model. We answer this question in the negative. More precisely the reason that no-regret dynamics suffer slow convergence is that the update rule depends only on the expressed opinions. In this section we exhibit an update rule that depends on the graph G and achieves exponentially fast convergence. Precisely, our update rule depends on the expressed opinions, the number of neighbors of each agent, and the number of agents n. The latter is interesting because information about the graph G combined with agents that do not act selfishly, can restore the exponential convergence rate.

The main idea of the protocol is straightforward: to counterbalance the imperfect information, the agents can spend some rounds to simulate one round of the original FJ-model. To do this, they agree to stop updating their expressed opinion for a large enough window of rounds so that everybody learns, with high probability, *exactly* the average of the opinions of their neighbors. Following the ideas of Section 3 an agent could just average the opinions that she gets in this window. Unfortunately this would again result in a $\text{poly}(1/\varepsilon)$-round protocol. However this can be fixed by using the additional knowledge of the number of agents and the number of neighbors. Precisely, each agent i keeps an array with the frequencies of the different opinions that she observes. The catch is that at the end of the window, she rounds each frequency to the closest multiple of $1/d_i$, where d_i is the number of neighbors of agent i. This rounding step is crucial to ensure the exponential convergence rate. To see that this works, first notice that if all agents stop updating their opinions for a number of rounds, each agents just needs to specify exactly how many of her neighbors share a specific observed opinion. If the length of the window is large the frequency of a specific opinion at the end of the window will be sufficiently close to the true frequency. Since the frequencies of the opinions that agent i observes can only be multiples of $1/d_i$, we can round the estimated frequencies to the closest multiple of $1/d_i$ to recover the true frequencies and use them to get the exact average of the opinions of the neighbors. To bound the length of the window we use a VC-dimension argument and show that with $n^2 \log n$ rounds each agent knows the frequencies within error smaller than $1/n$ with constant probability, which then can be trivially amplified by repeating the procedure.

We next state a version of the standard VC-Inequality that we will use in our argument. Let P be a discrete distribution over $[n]$, and let $S_1, \ldots, S_t$ be t i.i.d samples drawn from P, i.e. $(S_1, \ldots, S_t) \sim P^t$. The empirical distribution $\hat{P}_t$ is the following estimator of the density of P.

$$\hat{P}_t(A) = \frac{\sum_{i=1}^{t} \mathbf{1}[S_i \in A]}{t}, \tag{10}$$

where $A \subseteq [n]$. In words, $\hat{P}_t$ simply counts how many times the value i appeared in the samples $S_1, \ldots S_t$. We will use the following version of the classical result of Vapnik and Chervonenkis.

Lemma 7. *Let $\mathcal{A}$ be a collection of subsets of $\{1, \ldots, n\}$ and let $S_{\mathcal{A}}(t)$ be the Vapnik-Chervonenkis shatter coefficient, defined by*

$$S_{\mathcal{A}(t)} = \max_{x_1, \ldots, x_t \in [n]} \left| \{ \{x_1, \ldots, x_t\} \cap A : A \in \mathcal{A} \} \right|.$$

Then

$$\mathbf{E}_{P^t} \left[\max_{A \in \mathcal{A}} \left| \hat{P}_t(A) - P(A) \right| \right] \leq 2 \sqrt{\frac{\log 2 S_{\mathcal{A}}(t)}{t}}$$

Algorithm 3 Graph Aware Update Rule

$x_i(0) \leftarrow s_i$.
$M_1 = O(\ln(n/\varepsilon))$, $M_2 = O(d^2 \ln(d))$
for $l = 1, \ldots, \ln(1/\varepsilon)$ **do**
  Keep a set A of tuples $(x, \mathrm{freq}(x))$ of and an array B of length M_1.
  for $j = 1, \ldots, M_1$ **do**
    for $k = 1, \ldots, M_2$ **do**
      Get the opinion X_k of a random neighbor.
      if X_j is not in A **then**
        Insert $(X_k, 1)$ to A.
      else
        $(X_k, \mathrm{freq}(X_k)) \leftarrow (X_k, \mathrm{freq}(X_k) + 1)$.
    Divide all frequencies of A by M_2.
    Round all frequencies of ÎŚ to the closest multiple of $1/d_i$.
    $B(j) \leftarrow \alpha_i \sum_{x \in A} \mathrm{freq}(x) + (1 - \alpha_i) s_i$.
  $x_i(t) \leftarrow \mathrm{majority}_j B(j)$.

Theorem 7. *Let $I = (G(V, E), s, a)$ be an instance of the opinion formation game of Definition 1 with $a > 1/2$. Let d be the maximum degree of the graph G and $n = |V|$. There exists an update rule after $O(d^2 \log^2 n \log^2(1/\varepsilon))$ rounds achieves expected error $\mathbf{E}\left[\|x_t - x^*\|_\infty\right] \leq \varepsilon$.*

Proof. According to the update rule 3 all agents fix their opinions $x_i(t)$ for $M_1 \times M_2$ rounds. To estimate the sum of the opinions each agent estimates the frequencies k_j/d_i. Since the neighbors have at most d_i different opinions we can map the opinions to natural numbers in $[d_i]$. At each round the agent gets the opinion of a random neighbor and therefore the samples X_i that she observes are drawn from a discrete distribution P supported on $[d_i]$. If k_j be the (absolute) frequency of the opinion j namely the number of neighbors that express j as their opinion, then the probability $P(j)$ of opinion j is k_j/d_i. To learn the probabilities $P(j)$ using samples from P, we let $\mathcal{A} = \{\{1\}, \{2\}, \ldots, \{d_i\}\}$ and use Lemma 7 to get that

$$\mathbf{E}_{P^m} \left[\max_{j \in [d_i]} \left| \hat{P}_m(j) - P(j) \right| \right] \leq 2 \sqrt{\frac{\log 2 d_i}{m}},$$

since $S_{\mathcal{A}} \leq n$. Therefore, an agent can draw $m = 100n^2 \log(2n)$ to learn the frequencies k_j/d_i within expected error $1/(5d)$. Notice now that the array A after line 12 corresponds to the empirical distribution of equation (10). Notice that if the agents have estimations of the frequencies k_j/d_i with error smaller than $1/d$, then by rounding them to the closest multiple of $1/d_i$ they learn the frequencies exactly. By Markov's inequality we have that with probability at least $4/5$ the rounded frequencies are exactly correct. By standard Chernoff bounds we have that if the agents repeat the above procedure $\ln(1/\delta)$ times and keep the most frequent of the answers $B(j)$, then they will obtain the correct answer with probability at least $1-\delta$. We know that, having computed the *exact* average of the opinions of the neighbors Î§$(\log(1/\varepsilon))$ rounds are enough to achieve error ε. Since we need all nodes to succeed at computing the exact averages for Î§$(\log(1/\varepsilon))$ rounds, from the union bound we get that for $\delta < \frac{\varepsilon}{n \ln(1/\varepsilon)}$, with probability at least $1-\varepsilon$ the error is at most ε. Finally, from the law of total expectation, after $T = O(d^2 \log d \log(\varepsilon)(\log(n/\varepsilon) + \log\log(1/\varepsilon))$ rounds the expected error is $\mathbf{E}\left[\|x_T - x^*\|_\infty\right] = (1-\varepsilon)\varepsilon + \varepsilon \leq 2\varepsilon$.

References

BEDL06. 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.

BFM16. Vittorio Bilò, Angelo Fanelli, and Luca Moscardelli. Opinion formation games with dynamic social influences. In Yang Cai and Adrian Vetta, editors, *Web and Internet Economics*, pages 444–458, Berlin, Heidelberg, 2016. Springer Berlin Heidelberg.

BGM13. Kshipra Bhawalkar, Sreenivas Gollapudi, and Kamesh Munagala. Coevolutionary opinion formation games. In *Symposium on Theory of Computing Conference, STOC'13, Palo Alto, CA, USA, June 1-4, 2013*, pages 41–50, 2013.

BKO11. David Bindel, Jon M. Kleinberg, and Sigal Oren. How bad is forming your own opinion? In *IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011, Palm Springs, CA, USA, October 22-25, 2011*, pages 57–66, 2011.

CHM17. Johanne Cohen, Amélie Héliou, and Panayotis Mertikopoulos. Hedging under uncertainty: Regret minimization meets exponentially fast convergence. In *Algorithmic Game Theory - 10th International Symposium, SAGT 2017, L'Aquila, Italy, September 12-14, 2017, Proceedings*, pages 252–263, 2017.

Duc. John Duchi. Stats311, Lecture Notes.

EMN09. Eyal Even-Dar, Yishay Mansour, and Uri Nadav. On the convergence of regret minimization dynamics in concave games. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 523–532, 2009.

FJ90. Noah E. Friedkin and Eugene C. Johnsen. Social influence and opinions. *The Journal of Mathematical Sociology*, 15(3-4):193–206, 1990.

FS99. Yoav Freund and Robert E. Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(1):79 – 103, 1999.

FV97. Dean P. Foster and Rakesh V. Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21(1):40 – 55, 1997.

GS14. Javad Ghaderi and R. Srikant. Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate. *Automatica*, 50(12):3209–3215, 2014.

SA. Hart Sergiu and MasâĂŘColell Andreu. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150.

Wal39. Abraham Wald. Contributions to the Theory of Statistical Estimation and Testing Hypotheses. *The Annals of Mathematical Statistics*, 10(4):299–326, December 1939.

YOA+13. Mehmet Ercan Yildiz, Asuman E. Ozdaglar, Daron Acemoglu, Amin Saberi, and Anna Scaglione. Binary opinion dynamics with stubborn agents. *ACM Trans. Economics and Comput.*, 1(4):19:1–19:30, 2013.

Yu97. Bin Yu. Assouad, Fano, and Le Cam. In *Festschrift for Lucien Le Cam*, pages 423–435. Springer, New York, NY, 1997.

ZLZ17. Yichi Zhou, Jialian Li, and Jun Zhu. Identify the nash equilibrium in static games with random payoffs. In *Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017*, pages 4160–4169, 2017.

A Fictitious Play Convergence Rate

We give here the proof of the following technical lemma that we used to derive an upper bound on the rate of convergence of the fictitious play dynamics in Section 3. We restate Lemma **??** for completeness.

Lemma 8. *Let $e(t)$ be a function satisfying the recursion*

$$e(t) = \delta(t) + (1-\alpha)\frac{\sum_{\tau=0}^{t-1} e(\tau)}{t} \text{ and } e(0) = \|x^0 - x^*\|_\infty,$$

where $\delta(t) = \sqrt{\frac{\ln(Dt^2)}{t}}$, $\delta(0) = 0$, and $D > \mathrm{e}^2$ is a positive constant. Then $e(t) \le \sqrt{2\ln(D)}\frac{(\ln t)^{3/2}}{t^{\min(\rho,\,1/2)}}$.

Proof. Observe that for all $t \ge 0$ the function $e(t)$ the following recursive relation

$$e(t+1) = e(t)\left(1 - \frac{\rho}{t+1}\right) + \delta(t+1) - \delta(t) + \frac{\delta(t)}{t+1} \tag{11}$$

For $t = 0$ we have that

$$e(1) = (1-\rho)e(0) + \delta(1) = (1-\rho)e(0) + \sqrt{\ln D} \tag{12}$$

Observe that for $D > \mathrm{e}^2$, $\delta(t)$ is decreasing for all $t \ge 1$. Therefore, $\delta(t+1) - \delta(t) + \frac{\delta(t)}{t+1} \le \frac{\delta(t)}{t+1}$ and from equations (11) and (12) we get that for all $t \ge 0$

$$e(t+1) \le e(t)\left(1 - \frac{\rho}{t+1}\right) + \frac{\sqrt{\ln(D(t+1)^2)}}{(t+1)^{3/2}} \le e(t)\left(1 - \frac{\rho}{t+1}\right) + \frac{\sqrt{2\ln(D(t+1))}}{(t+1)^{3/2}}$$

Now let $g(t) = \frac{\sqrt{2\ln(Dt)}}{t^{3/2}}$ to obtain for all $t \ge 1$

$$\begin{aligned}
e(t) &\le (1 - \frac{\rho}{t})e(t-1) + g(t)\\
&\le (1 - \frac{\rho}{t})(1 - \frac{\rho}{t-1})e(t-2) + (1 - \frac{\rho}{t})g(t-1) + g(t)\\
&\le (1 - \frac{\rho}{t})\cdots(1-\rho)e(0) + \sum_{\tau=1}^{t} g(\tau)\prod_{i=\tau+1}^{t}(1 - \frac{\rho}{i})\\
&\le \frac{e(0)}{t^\rho} + \sum_{\tau=1}^{t} g(\tau)e^{-\rho\sum_{i=\tau+1}^{t}\frac{1}{i}}\\
&\le \frac{e(0)}{t^\rho} + \sum_{\tau=1}^{t} g(\tau)e^{-\rho(H_t - H_\tau)}\\
&\le \frac{e(0)}{t^\rho} + e^{-\rho H_t}\sum_{\tau=1}^{t} g(\tau)e^{\rho H_\tau}\\
&\le \frac{e(0)}{t^\rho} + \frac{\sqrt{2}}{t^\rho}\sum_{\tau=1}^{t}\tau^\rho\frac{\sqrt{\ln(D\tau)}}{\tau^{3/2}}\\
&\le \frac{e(0)}{t^\rho} + \frac{\sqrt{2\ln D}}{t^\rho}\sum_{\tau=1}^{t}\frac{\sqrt{\ln\tau}}{\tau^{3/2-\rho}}
\end{aligned}$$

We observe that

$$\sum_{\tau=1}^{t}\frac{\sqrt{\ln\tau}}{\tau^{3/2-\rho}} \le \int_{\tau=1}^{t}\frac{\sqrt{\ln\tau}}{\tau^{3/2-\rho}}\mathrm{d}\tau \tag{13}$$

since, $\tau \mapsto \frac{\sqrt{\ln\tau}}{\tau^{3/2-\rho}}$ is a decreasing function of τ for all $\rho \in [0,1]$.

– If $\rho \le 1/2$ then

$$\int_{\tau=1}^{t} \tau^{\rho} \frac{\sqrt{\ln \tau}}{\tau^{3/2}} \mathrm{d}\tau \le \sqrt{\ln t} \int_{\tau=1}^{t} \frac{1}{\tau} \mathrm{d}\tau = (\ln t)^{3/2}$$

– If $\rho > 1/2$ then

$$\begin{aligned}
\int_{\tau=1}^{t} \tau^{\rho} \frac{\sqrt{\ln \tau}}{\tau^{3/2}} \mathrm{d}\tau &= \int_{\tau=1}^{t} \tau^{\rho-1/2} \frac{\sqrt{\ln \tau}}{\tau} d\tau \\
&= \frac{2}{3} \int_{\tau=1}^{t} \tau^{\rho-1/2} ((\ln \tau)^{3/2})' \mathrm{d}\tau \\
&= \frac{2}{3} (\ln t)^{3/2} - (\rho - 1/2) \frac{2}{3} \int_{\tau=1}^{t} \tau^{\rho-3/2} (\ln \tau)^{3/2} \mathrm{d}\tau \\
&\le \frac{2}{3} (\ln t)^{3/2}
\end{aligned}$$

Theorem 8. *Let $I = (P, s, \alpha)$ be any instance of the opinion formation game of Definition 1 with equilibrium $x^* \in [0,1]^n$. The opinion vector $x(t)$ produced by Algorithm 1 after t rounds satisfies*

$$\mathbf{E}\left[\|x(t) - x^*\|_\infty\right] \le C \sqrt{\log n} \frac{(\log t)^{3/2}}{t^{\min(1/2, \rho)}},$$

where $\rho = \min_{i \in V} a_i$ and C is a universal constant.

Proof. By Lemma 1 we have that for all $t \ge 1$ and $p \in [0,1]$,

$$\mathbf{P}\left[\|x(t) - x^*\|_\infty \le e_p(t)\right] \ge 1 - p$$

where $e_p(t)$ is the solution of the recursion, $e_p(t) = \delta(t) + (1-\rho) \frac{\sum_{\tau=0}^{t-1} e_p(\tau)}{t}$ with $\delta(t) = \sqrt{\frac{\log(\pi^2 n t^2/(6p))}{t}}$. Setting $p = \frac{1}{12\sqrt{t}}$ we have that

$$\mathbf{P}\left[\|x(t) - x^*\|_\infty \le e(t)\right] \ge 1 - \frac{1}{12\sqrt{t}}$$

where $e(t)$ is the solution of the recursion $e(t) = \delta(t) + (1-\rho) \frac{\sum_{\tau=0}^{t-1} e_p(\tau)}{t}$ with $\delta(t) = \sqrt{\frac{\log(2\pi^2 n t^{2.5})}{t}}$. Since $2\pi^2 \ge e^{2.5}$, Lemma 2 applies and $e(t) \le C \sqrt{\log n} \frac{\log t^{3/2}}{t^{\min(\rho, 1/2)}}$ for some universal constant C. Finally,

$$\mathbf{E}\left[\|x(t) - x^*\|_\infty\right] \le \frac{1}{12\sqrt{t}} + (1 - \frac{1}{12\sqrt{t}}) C \sqrt{\log n} \frac{(\log t)^{3/2}}{t^{\min(\rho, 1/2)}} \le (C + \frac{1}{12}) \sqrt{\log n} \frac{(\log t)^{3/2}}{t^{\min(\rho, 1/2)}}$$

B Lower bound for no-regret Dynamics

In the following Lemma we show how we can use algorithm A to construct an estimator $\hat{\theta_A}$ for Bernoulli distributions. We restate **??** for completeness.

Theorem 9. *Let the no-regret algorithm A such that for all instances I, $\lim_{t \to \infty} t^{1+c} \mathbf{E}_I\left[\|x_A(t) - x^*\|_\infty\right] = 0$. Then there exists an estimator $\hat{\theta_A}$ such that for all $q \in Q[0,1]$,*

$$\lim_{t \to \infty} t^{1+c} R_q(t) = 0$$

Proof. At first we remind that an estimator $\hat{\theta}$ is a sequence of functions $\{\hat{\theta}_t\}_{t=1}^{\infty}$, where $\theta_t : \{0,1\}^t \mapsto [0,1]$. We construct such a sequence using the algorithm A. We also remind that when an agent i runs algorithm A, she selects $x_i(t)$ according to the cost functions $\{C_i^\tau\}$ that she has already received

$$x_i(t) = A_t(C_i^1, \ldots, C_i^{t-1})$$

Consider an agent i with $a_i = 1$ and $s_i = 0$ that runs A. Then $C_i^t(x) = x^2$ for all t and $x_i(t) = A_t(x^2, \ldots, x^2)$. The latter means that $x_i(t)$ only depends on t, $x_i(t) = h_0(t)$. Equivalently, if $a_i = 1$ and $s_i = 1$ then $x_i(t) = A_t((1-x)^2, \ldots, (1-x)^2)$ and $x_i(t) = h_1(t)$. Finally, consider an agent i with $a_i = 1/2$ and $s_i = 0$. In this case $C_i^t = \frac{1}{2}x^2 + \frac{1}{2}(x - y_t)^2$, where $y_t \in [0,1]$ is the opinion of the neighbor $j \in N_i$ that i met at round t. As a result, $x_i(t) = A_t(\frac{1}{2}x^2 + \frac{1}{2}(x-y_1)^2, \ldots, \frac{1}{2}x^2 + \frac{1}{2}(x - y_{t-1})^2) = f_t(y_1, \ldots, y_{t-1})$. The estimator $\hat{\theta_A}$ is the following sequences $\{\hat{\theta}_t\}_{t=1}^{\infty}$

$$\hat{\theta}_t(Y_1, \ldots, Y_t) = \frac{1}{2} f_{t+1}(h_{Y_1}(1), \ldots, h_{Y_t}(t))$$

Observe that $\hat{\theta}_t : \{0,1\}^t \mapsto [0,1]$ meaning that $\hat{\theta_A}$ is a valid estimator for Bernoulli distributions. Now for any $p \in Q([0,1])$, we construct an appropriate instance I_p s.t. $R_p(t) = \mathbf{E}_p\left[|\hat{\theta}_t - p|\right] \leq 2\mathbf{E}_{I_p}\left[\|x^t - x^*\|_\infty\right]$. For $p = \frac{k}{n}$ consider the following instance I_p with $n+1$ agents:

- A central agent with $s_c = 0$ and $a_c = 1/2$.
- Directed edges from the central agent to all the other agents.
- k agents with $s_i = 0$ and $a_i = 1$
- $n-k$ agents with $s_i = 1$ and $a_i = 1$

We just need to prove that in I_p, $\mathbf{E}_p\left[|\hat{\theta}_t - p|\right] \leq 2\mathbf{E}_{I_p}\left[\|x^t - x^*\|_\infty\right]$. Notice that $x_c^* = \frac{p}{2}$ and $x_i^* = s_i$ if $i \neq c$. .
At round t, if the oracle returns to the center agent the value $h_1(t)$ of a 1-agent, then $Y_t = 1$ otherwise $Y_t = 0$. As a result, $\mathbf{P}\left[Y_t = 1\right] = p$ and

$$\begin{aligned}
\mathbf{E}_{I_p}\left[\|x^t - x^*\|_\infty\right] &\geq \mathbf{E}_{I_p}\left[||x_c^t - x_c^*||\right] \\
&= \mathbf{E}_p\left[|\frac{\widehat{\theta_t}}{2} - \frac{p}{2}\right]| = R_p(t)
\end{aligned}$$

Theorem 10. *Let $g(p) = \lim_{t\to\infty} t^{1+c} R_p(t)$ be a continuous function. Then there exists q_0 such that*

$$\lim_{t\to\infty} t^{1+c} R_{q_0}(t) > 0$$

Proof. laksdokasodkaok

We next give a rigorous measure-theoretic proof of Theorem **??**.

Theorem 11. *Let $\theta_t : \{0,1\}^t \to [0,1]$ be a sequence of estimators for the success probability p of a Bernoulli random variable with distribution P. There exists $p \in [0,1]$ such that*

$$\lim_{t\to\infty} t^2\ \mathbf{E}_{X\sim P^t}\left[|\theta_t(X) - p|\right] > 0.$$

Proof. Observe that $\theta_t(\{0,1\}^t)$ has cardinality at most 2^t. Since

$$\sum_{0\leq i\leq t} \sum_{\|x\|_1 = i} |\theta_t(x) - p|\, p^i (1-p)^{t-i} \geq \sum_{0\leq i\leq t} \binom{t}{i} \left| \frac{\sum_{\|x\|_1 = i} \theta_t(x)}{\binom{t}{i}} - p \right| p^i (1-p)^{t-i}.$$

Thus, without loss of generality we assume that $\theta_t(\{0,1\}^t)$ contains at most $t+1$ discrete points.

In the following, we work in the measure space $(\mathbf{R}, \mathcal{M}, \mu)$, where μ is the Lebesgue measure, and $\mathcal{M}$ is the σ-algebra of the Lebesgue measurable sets. Suppose that there exists no such $p \in [0,1]$. Let

$$A = \{p \in [0,1]\ :\ \lim_{t\to\infty} t^2\ \mathbf{E}_{X\sim P^t}\left[|\theta_t(X) - p|\right] = 0\}.$$

Then $A = [0,1]$ and A is measurable as an interval. Notice that,

$$A \subseteq \bigcup_{t=1}^{\infty} \bigcap_{k=t}^{\infty} A_k,$$

where $A_k = \{p \in [0,1] \ : \ R_k(p) < 1/2\}$, and $R_k(p) = k^2 \ \mathbf{E}_{X \sim P^k} [|\theta_k(X) - p|]$. We have that $R_k : [0,1] \to [0, +\infty)$ is polynomial of degree t in p and therefore it is a measurable function. Thus, A_k is measurable. We now show that

$$A_k \subseteq B_k \coloneqq \{p \in [0,1] \ : \ k^2 \min_{0 \le i \le k} |\theta_k(i) - p| < 1\}.$$

We prove this by contradiction. Suppose that $p \in A_k$ but $p \notin B_k$. Since $p \in A_k$ we have that

$$R_k(p) = k^2 \sum_{i=0}^{k} \binom{k}{i} |\theta_k(i) - p| \, p^i (1-p)^{k-i} \ge k^2 \min_{0 \le i \le k} |\theta_k(i) - p| \sum_{i=0}^{k} \binom{k}{i} p^i (1-p)^{k-i} \ge 1.$$

Since the functions $p \mapsto k^2 \, |\theta_k(i) - p|$ are measurable, their pointwise minimum is measurable and therefore the sets B_k are also measurable. We next proceed to bound $\mu(B_k)$. Since θ_k can only take k different values we have that there exist $k+1$ intervals (a_{k_i}, b_{k_i}) of length at most $2/k^2$ such that $B_k = \bigcup_{i=0}^{k} (a_{k_i}, b_{k_i})$. Since μ is subadditive we have

$$\mu(B_k) \le \sum_{i=0}^{k} \frac{2}{k^2} = \frac{2(k+1)}{k^2}.$$

Now observe that

$$\mu(A) \le \mu\left(\bigcup_{t=1}^{\infty} \bigcap_{k=t}^{\infty} A_k\right) \le \mu\left(\bigcup_{t=1}^{\infty} \bigcap_{k=t}^{\infty} B_k\right) \le \sum_{t=1}^{\infty} \mu\left(\bigcap_{k=t}^{\infty} B_k\right) \le \sum_{t=1}^{\infty} \lim_{k \to \infty} \mu(B_k) = 0$$

Which is a contradiction since we assumed $A = [0,1]$.