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.

1 Introduction

1.1 Friedkin-Johsen Model and Opinion Formation Games

In [?] the following *opinion formation game* was introduced. A weighted directed graph $G(V, E, w)$ is assumed where and the vertices stand for the n agents ($|V| = n$) and the acres for the social influence among them. Each agent $i \in V$ possess an *internal opinion* $s_i \in [0, 1]$ and a *self confidence coefficient* $w_i > 0$. The strategy of each agent i is the opinion $x_i \in [0, 1]$ that she publicly expresses incurring her cost

$$C_i(x_i, x_{-i}) = \sum_{j \in N_i} w_{ij}(x_i - x_j)^2 + w_i(x_i - s_i)^2 \tag{1}$$

where N_i denotes i's *neighbors* and w_{ij} stands for the social influence j imposes on i. In [?] they proved that the above game always admits a *Pure Nash Equilibrium* (PNE) $x^* \in [0, 1]^n$ and studied its efficiency with respect to the total disagreement cost. They proved that the *Price of Anarchy* is less than 9/8 in case G is bidirectional and $w_{ij} = w_{ji}$.

In the repeated version of the game defined in (1), at each round t each agent i selects an opinion $x_i(t)$ and then suffers cost $C_i(x_i(t), x_{-i}(t))$. If each agent updates her opinion to be the *best response* of $x(t-1)$,

$$x_i(t) = \underset{x \in [0,1]}{\operatorname{argmin}}\, C_i(x, x_{-i}(t-1)) = \frac{\sum_{j \in N_i} w_{ij} x_j(t-1) + w_i s_i}{\sum_{j \in N_i} w_{ij} + w_i} \tag{2}$$

we obtain the Friedkin-Johsen model (FJ-model), which is one of the most influential models in opinion dynamics. The convergence properties of the FJ-model have been extensively studied. In [?] they proved that $x(t)$ always converges to the PNE x^* and provided bounds for the convergence time for various graph topologies. As a result, the above *opinion formation game* has some nice algorithmic properties: It always admits a unique equilibrium point x^* and there exists a simple but most importantly rational update rule for selfish agents that leads the overall system to equilibrium.

1.2 Opinion Formation Games with Random Payoffs

Our work is motivated by the fact that the definition of the cost $C_i(x_i, x_{-i})$ in (1) implies that agent i meets with all of her neighbors. This is more clear in the update rule (2). Each agent, in order to compute her best response, has to learn the opinion of all her neighbors. The latter seems quite unnatural in today's huge social networks (e.g. Facebook, Twitter etc.), in which each user may have several hundreds of friends. Thus, it is far more reasonable to assume that each day an agent meets a small subset of her acquaintances and suffers a cost based on how much she disagrees with them. To capture the above thoughts, we introduce a variant of the opinion formation game in which the disagreement cost of each agent i is a random variable depending on the random meetings of i.

Definition 1. *For a given opinion vector $x \in [0, 1]^n$, the disagreement cost of agent i is the random variable $C_i(x_i, x_{-i})$ defined as follows:*

- *i meets one of her neighbors j with probability $p_{ij} = w_{ij} / \sum_{j \in N_i} w_{ij}$*
- *suffers cost $(1 - a_i)(x_i - x_j)^2 + a_i(x_i - s_i)^2$*

where $\alpha_i = w_i / (\sum_{j \in N_i} w_{ij} + w_i)$

The main difference of the original opinion formation game with our variant is that in the first case an opinion vector $x \in [0, 1]^n$ defines *deterministically* the cost $C_i(x_i, x_{-i})$ of each agent i, whereas in the second case it defines (according to Definition 1) a probability distribution on the cost $C_i(x_i, x_{-i})$ that i suffers. Recent works [?,?] study games with random payoffs. The reason is that the random payoff setting is more suitable to model realistic situations in which randomness naturally occurs because of incomplete information.

The cost $C_i(x_i, x_{-i})$ in (1) can be written equivalently

$$C_i(x_i, x_{-i}) = W_i \left((1 - \alpha_i) \sum_{j \in N_i} p_{ij}(x_i - x_j)^2 + \alpha_i(x_i - s_i)^2 \right) \tag{3}$$

where $W_i = \sum_{j \in N_i} w_{ij} + w_i$ is a positive constant independent of the opinion vector $x \in [0,1]^n$. Thus, the random cost in Definition 1 has a natural interpretation: the coefficient α_i measures the reluctance of agent i to adopt an opinion other than s_i, while p_{ij} can be seen as the *real* influence that j poses on i. In Definition 1, p_{ij} is the frequency that i meets j, meaning that the influence that j poses on i is just a measure on how often they meet. The latter aligns with the common belief that we are influenced more by those we interact more often. Equation (3) also helps to establish the existence of PNE for our random payoff variant. In our case, the notion of PNE extends with respect to the expected cost of each agent. Namely, $x^* \in [0,1]$ is a PNE if and only if $\mathbf{E}\left[C(x_i^*, x_{-i}^*)\right] \le \mathbf{E}\left[C(x_i, x_{-i}^*)\right]$ for each agent i. Since $\mathbf{E}\left[C_i(x_i, x_{-i})\right] = (1-\alpha_i)\sum_{j \in N_i} p_{ij}(x_i - x_j)^2 + \alpha_i(x_i - s_i)^2$, it follows from (3) that the opinion formation game with random payoffs has the same equilibrium x^* as the original opinion formation game.

Instead of denoting an instance of the opinion formation game using a graph G and weights w_{ij}, w_i we adopt the following more convenient notation.

Definition 2. *We denote an instance of the opinion formation game with random payoffs as* (P, s, α)*.*

- P *is a* $n \times n$ *matrix with non-negative elements* p_{ij}*, with* $p_{ii} = 0$ *and* $\sum_{j=1}^n p_{ij}$ *is either* 0 *or* 1*.*
- $s \in [0,1]^n$ *is the internal opinion vector.*
- $\alpha \in [0,1]^n$ *the self confidence coefficient vector.*

We use the matrix P to simplify notation, $p_{ij} = w_{ij}/(\sum_{j \in N_i} w_{ij} + w_i)$ if $j \in N_i$ and 0 otherwise. If $N_i \neq \emptyset$ then $\sum_{j=1}^n p_{ij} = 1$ otherwise it is 0. We remark that in case $N_i = \emptyset$, $\alpha_i = 1$ and agent i suffers cost $(x_i - s_i)^2$. Abusing notation we will sometimes refer to the graph G.

1.3 Our Results

We focus on the repeated version of the game in Definition 1. At round t, each agent i selects an opinion $x_i(t) \in [0,1]$ and then suffers disagreement cost

$$(1-\alpha_i)(x - x_{W_i^\tau}(\tau))^2 + \alpha_i(x - s_i)^2$$

where W_i^t denotes the neighbor that i met at round t. We are interested in simple and natural update rules that the agents can adopt such that the resulting opinion vector $x(t) \in [0,1]^n$ converges to x^*.

In Section 2, we study the convergence properties of $x(t)$ if all agents update their opinion as follows:

$$x_i(t) = \operatorname*{argmin}_{x \in [0,1]} \sum_{\tau=0}^{t-1} (1-\alpha_i)(x - x_{W_i^\tau}(\tau))^2 + \alpha_i(x - s_i)^2 \tag{4}$$

Roughly speaking, the above update says "*play the best according to what you have observed*". According to this principle Brown proposed *fictitious play* [**?**], which is one of the most intuitive and simple models of playing in finite games. Abusing terminology we refer to (4) as fictitious play. We show that in our infinite strategy game, if all agents adopt fictitious play, the resulting opinion vector $x(t)$ converges to x^* with the following rate.

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

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

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

The update rule (4) guarantees convergence while vastly reducing the information exchange between the agents at each round. In (4) each agent i learns the opinion of only one agent at each round whereas in the classical FJ-model (2), agent i must learn the opinions of all her neighbors. In terms of total communication needed to get within distance ε of the equilibrium x^*, the update rule (4) needs $O(n \log n)$ communication while (2) needs $O(|E|)$. Of course for this difference to be significant we need each agent to have at least

$O(\log n)$ friends. A large social network like Facebook has approximately 2 billion users and each user has usually more that 100 friends which is far more than $\log(2\ 10^9)$.

In Section 3 we argue apart that from its simplicity, fictitious play is a *rational game play* for selfish agents in a much stronger sense. At each round t each agent i, selects an opinion $x_i(t) \in [0,1]$ and suffers a cost $(1-\alpha_i)(x_i(t) - x_{W_i^t}(t))^2 + \alpha_i(x_i(t) - s_i)^2$. Since agent i is selfish and only interests in minimizing her cost, it is reasonable to assume that she selects $x_i(t)$ according to an *no-regret algorithm* A for the *online convex optimization problem* where the adversary at round t chooses a function $f_t(x) = (1-\alpha_i)(x-b_t)^2 + \alpha_i(x-s_i)^2$. In Theorem 3 we prove that fictitious play is a no-regret algorithm for the above OCO problem.

Theorem 2. *Consider the function* $f : [0,1]^2 \mapsto [0,1]$ *with* $f(x,b) = (1-\alpha)(x-b)^2 + \alpha(x-s)^2$. *Let* $\{b_t\}_{t=1}^{\infty}$ *be an arbitrary sequence with* $b_t \in [0,1]$. *If* $x_t = \mathrm{argmin}_{x\in[0,1]} \sum_{\tau=0}^{t-1} f(x, b_\tau)$ *then for all* t,

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

Even though the update rule (4) has the above desired properties, the convergence rate of the produced dynamics is outperformed by the convergence rate of the classical FJ-model. For a fixed instance $I = (P, s, \alpha)$, fictitious play converges with rate $\widetilde{O}(1/t^{\min(\rho,1/2)})$ while FJ-model converges with rate $O(e^{-\rho t})$ [?]. As a result the following question arises

Question 1. Can the agent adopt other no-regret algorithms such that the resulting dynamics converges exponential fast?

In section 4 we answer this question in the negative. The reason that fictitious play converges slowly is that the update rule (4) only depends on the opinions of the agents that agent i meets, α_i, and s_i. This is also true for any no-regret algorithm that i uses to select $x_i(t)$ (see Definition **??** in Section 3). We call such update rules "*opinion dependent*". In Theorem 3 we show that for any opinion dependent update rule there exists an instance $I = (P, s, \alpha)$ where $\mathrm{poly}(1/\varepsilon)$ rounds are required to achieve convergence within error ε.

Theorem 3. *Let* A *be an* opinion dependent *update rule, which all agents use to update their opinions. For any* $c > 0$ *there exists an instance* $I = (P, s, a)$ *such that*

$$\mathbf{E}\left[\|x_A(t) - x^*\|_\infty\right] = \Omega(1/t^{1+c}),$$

where $x_A(t)$ *denotes the opinion vector produced by* A.

To prove Theorem 3, we show that opinion dependent rules with "small round complexity" for any instance I, imply the existence of estimators for Bernoulli distributions with "small" sample complexity. Then with a simple argument presented in Lemma 6, we show that such estimators cannot exist. In Section 4 we also briefly discuss two well-known sample complexity lower bounds developed in the field of statistical estimation and explain why they do not work in our case. We mention that Theorem 3 also applies for "opinion dependent rules" that also depend on the agents' indices that i met.

In Section **??**, we present a simple update rule that is not opinion dependent and achieves error rate $e^{-O(\sqrt{t})}$. This update rule is a function of the opinions and the indices of the agents that i met, α_i, s_i and the i-th row the matrix P. Because of Theorem 3 the dependency on the row P_i is inevitable in order to obtain exponential convergence. This update rules also reveals that the slow convergence of *opinion dependent* update rules is not due to the reduced information exchange (learning the opinion of only one agent) between the agents, but due to the fact that the agents are "oblivious" to influence matrix P of the game. The assumption that the agents are aware of the influence matrix P is up to discussion, however our results suggest that exponential convergence is ensured by the knowledge P and has much less to do with the information exchange among the agents.

1.4 Related Work

Our work belongs to the line of work studying the seminal Friedkin-Jonhsen model [?]. Bindel et al. in [?] 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 [?,?,?,?]. In [?] 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 [?], 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 [?,?].

In [?], [?], [?] 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 [?] they proved that no-regret dynamics converge to NE in the case of congestion games. Later in [?] 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 [?] 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 Fictitious Play Convergence Rate

In this section that if all agent select their opinion according to update rule (4) then the dynamics $x(t)$ converges to the unique equilibrium point x^*. For an instance (P, s, α) the opinion vector $x(t) \in [0,1]^n$ produced by (4) is defined as follows:

- Each agent i adopts her internal opinion, $x_i(0) = s_i$
- At round $t \geq 1$, each agent i updates her opinion as follows:

$$x_i(t) = \operatorname*{argmin}_{x \in [0,1]} \sum_{\tau=0}^{t-1} (1-\alpha_i)(x - x_{W_i^\tau}(\tau))^2 + \alpha_i (x - s_i)^2 \tag{4}$$

where W_i^t denotes the agent that i met at round t. The opinion vector $x(t)$ due to the random meeting of the agents. Our convergence metric is $\mathbf{E}\left[\|x(t) - x^*\|_\infty\right]$ where the expectation is taken over the random meeting of the agents. Our convergence result is stated in Theorem 7 and it is the main result of the section.

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

At first we present the high level idea of the proof of Theorem 7. According to Algorithm **??**, each agent i at round $t \geq 1$ updates her opinion as $x_i(t) = \operatorname{argmin}_{x\in[0,1]} \sum_{\tau=1}^{t} C_i^\tau(x, x_{W_i^\tau}(\tau - 1))$ where W_i^τ is the random variable denoting the agent j that i met at round τ ($\mathbf{P}\left[W_i^\tau = j\right] = p_{ij}$). 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 $\mathbf{E}[\|x(t)-x^*\|_\infty]$, 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:

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

Now assume that $|\frac{\sum_{\tau=1}^t \mathbf{1}[W_i^\tau = j]}{t} - p_{ij}|$ 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 $|\frac{\sum_{\tau=1}^t \mathbf{1}[W_i^\tau = j]}{t} - p_{ij}| \neq 0$ and the above analysis does not hold. In Lemma 1 we use the fact that $|\frac{\sum_{\tau=1}^t \mathbf{1}[W_i^\tau = j]}{t} - p_{ij}|$ 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:

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

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

Since W_i^τ are independent random variables with $\mathbf{P}[W_i^\tau = j] = p_{ij}$ 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,

$$\left|\frac{\sum_{\tau=1}^t x^*_{W_i^\tau}}{t} - \sum_{j\neq i} p_{ij}x_j^*\right| \leq \delta(t) \tag{6}$$

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

$$
\begin{aligned}
x_i(t) &= (1-\alpha_i)\frac{\sum_{\tau=1}^{t} x_{W_i^\tau}(\tau-1)}{t} + \alpha_i s_i \\
&\le (1-\alpha_i)\frac{\sum_{\tau=1}^{t} x^*_{W_i^\tau} + \sum_{\tau=1}^{t} e(\tau-1)}{t} + \alpha_i s_i \qquad (7)\\
&\le (1-\alpha_i)\left(\frac{\sum_{\tau=1}^{t} x^*_{W_i^\tau}}{t} + \frac{\sum_{\tau=0}^{t-1} e(\tau)}{t}\right) + \alpha_i s_i \\
&\le (1-\alpha_i)\left(\sum_{j\in N_i} p_{ij}x_j^* + \delta(t) + \frac{\sum_{\tau=0}^{t-1} e(\tau)}{t}\right) + \alpha_i s_i \qquad (8)\\
&\le x_i^* + \delta(t) + (1-\alpha)\left(\frac{\sum_{\tau=0}^{t-1} e(\tau)}{t}\right)
\end{aligned}
$$

We get (7) from the induction step and (8) from inequality (6). 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 7 folows by direct application of Lemma 1 and 2.

3 Fictitious Play is no-regret

In this section we explain why *fictitious play* is a rational behavioral assumption in the repeated version of the opinion formation game defined in 1. Based on this game we consider an appropriate *Online Convex Optimization* problem. This problem can be viewed as the following a game played between an adversary and a player. *At round t,*

1. *the player selects a value $x_t \in [0,1]$.*
2. *the adversary observes the x_t and selects a $b_t \in [0,1]$*
3. *the player receives cost $f(x_t, b_t) = (1-\alpha)(x_t - b_t)^2 + \alpha(x_t - s)^2$.*

where α, s are constants in $[0,1]$. The goal of the player is to pick x_t based on the history $(b_0,\dots,b_{t-1})$ in a way that minimizes the total cost. We emphasize that the agent has to select x_t before seeing b_t, otherwise the problem becomes trivial. We show that a good strategy that the player can follow is

$$
x_t = \operatorname*{argmin}_{x\in[0,1]} \sum_{\tau=0}^{t-1} f(x, b_\tau)
$$

which is known as *fictitious play*. We prove that in our OCO problem, *fictitious play* is a *no-regret* algorithm meaning that it provides guarantees concerining the cost that agent i receives. Informally speaking if an algorithm A is *no-regret* for an OCO problem then for any selection of the adversary, the the total cost that player suffers is less than the cost that she would suffer by selecting any fixed value. The gaurantees that *fictitious play* ensures are presented in Theorem which is the main result of this section.

Theorem 3. *Consider the function $f : [0,1]^2 \mapsto [0,1]$ with $f(x,b) = (1-\alpha)(x-b)^2 + \alpha(x-s)^2$. Let $\{b_t\}_{t=1}^{\infty}$ be an arbitrary sequence with $b_t \in [0,1]$. If $x_t = \mathrm{argmin}_{x\in[0,1]} \sum_{\tau=0}^{t-1} f(x,b_\tau)$ then for all t,*

$$\sum_{\tau=0}^{t} f(x_\tau, b_\tau) \leq \min_{x\in[0,1]} \sum_{\tau=0}^{t} f(x, b_\tau) + O\left(\log t\right)$$

Theorem 3 explains why update rule (4) is a reasonable choice for selfish agents. In our repeated game, agent i selects at round each t an opinion $x_i(t) \in [0,1]$ and then suffers cost $f_i(x_i(t), x_{W_i^t}(t)) = (1-\alpha_i)(x_i(t) - x_{W_i^t}(t))^2 + \alpha_i(x_i(t) - s_i)^2$. If agent i selects $x_i(t)$ according to the update rule 4 then Theorem 3 applies and

$$\frac{1}{t}\sum_{\tau=0}^{t} f_i(x_i(\tau), x_{W_i^\tau}(\tau)) \leq \frac{1}{t}\min_{x\in[0,1]} \sum_{\tau=0}^{t} f_i(x, x_{W_i^\tau}(\tau)) + O\left(\frac{\log t}{t}\right)$$

meaning that the time averaged total disagreement cost that agent i suffers is similar to the time averaged cost of the best fixed opinion.

The rest of the section is dedicated to prove Theorem 3. we first prove that a similar update rule that also takes into account the value b_t admits no-regret. Obviously knowing the value b_t before time selecting x_t is direct contrast with the OCO framework, however proving the no-regret property for this algorithm easily extends to proving the no-regret property for fictitious play.

Lemma 3. *Let $\{b_t\}_{t=0}^{\infty}$ be an arbitrary sequence with $b_t \in [0,1]$. Let $y_t = \mathrm{argmin}_{x\in[0,1]} \sum_{\tau=0}^{t} f(x,b_\tau)$ then for all t,*

$$\sum_{\tau=0}^{t} f(y_\tau, b_\tau) \leq \min_{x\in[0,1]} \sum_{\tau=0}^{t} f(x, b_\tau)$$

Proof. By definition of y_t, $\min_{x\in[0,1]} \sum_{\tau=0}^{t} f(x,b_\tau) = \sum_{\tau=0}^{t} f(y_t, b_\tau)$, so

$$\begin{aligned}
\sum_{\tau=0}^{t} f(y_\tau, b_\tau) - \min_{x\in[0,1]} \sum_{\tau=0}^{t} f(x, b_\tau) &= \sum_{\tau=0}^{t} f(y_\tau, b_\tau) - \sum_{\tau=0}^{t} f(y_t, b_\tau) \\
&= \sum_{\tau=0}^{t-1} f(y_\tau, b_\tau) - \sum_{\tau=0}^{t-1} f(y_t, b_\tau) \\
&\leq \sum_{\tau=0}^{t-1} f(y_\tau, b_\tau) - \sum_{\tau=0}^{t-1} f(y_{t-1}, b_\tau) \\
&= \sum_{\tau=0}^{t-2} f(y_\tau, b_\tau) - \sum_{\tau=0}^{t-2} f(y_{t-1}, b_\tau)
\end{aligned}$$

Continuing in the same way, we get $\sum_{\tau=0}^{t} f(y_\tau, b_\tau) \leq min_{x\in[0,1]} \sum_{\tau=0}^{t} f(x, b_\tau)$.

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 x_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 $f(x_t, b_t)$ that y_t takes as input doesn't change dramatically the minimum point of the sum. Thus, for each t the numbers x_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, $f(x_t, b_t) \leq f(y_t, b_t) + 2\frac{1-\alpha}{t+1} + \frac{(1-\alpha)^2}{(t+1)^2}$.*

Proof. We first prove that for all t,

$$|x_t - y_t| \leq \frac{1-\alpha}{t+1}. \tag{9}$$

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

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

The last inequality follows from the fact that $b_\tau \in [0,1]$. We now use inequality (9) to bound the difference $f(x_t, b_t) - f(y_t, b_t)$.

$$\begin{aligned}f(x_t, b_t) &= \alpha(x_t - s)^2 + (1-\alpha)(x_t - y_t)^2 \\ &\le \alpha(y_t - s)^2 + 2\alpha\,|y_t - s|\,|x_t - y_t| + \alpha\,|x_t - y_t|^2 \\ &\quad + (1-\alpha)(y_t - y_t)^2 + 2(1-\alpha)\,|y_t - y_t|\,|x_t - y_t| + (1-\alpha)\,|x_t - y_t|^2 \\ &\le f(y_t, b_t) + 2\,|x_t - y_t| + |y_t - x_t|^2 \\ &\le f(y_t, b_t) + 2\frac{1-\alpha}{t+1} + \frac{(1-\alpha)^2}{(t+1)^2}\end{aligned}$$

Theorem 3 easily follows since

$$\begin{aligned}\sum_{\tau=0}^{T} f(x_\tau, b_\tau) &\le \sum_{\tau=0}^{T} f(y_\tau, b_\tau) + \sum_{\tau=0}^{T} 2\frac{1-\alpha}{\tau+1} + \sum_{\tau=0}^{T}\frac{(1-\alpha)^2}{(\tau+1)^2} \\ &\le \min_{x\in[0,1]}\sum_{\tau=0}^{T} f(x, y_\tau) + 2(1-\alpha)(\log T + 1) + (1-\alpha)\frac{\pi^2}{6} \\ &\le \min_{x\in[0,1]}\sum_{\tau=0}^{T} f(x, y_\tau) + O(\log T)\end{aligned}$$

4 Lower Bound for Opinion Dependent Dynamics

In this section, we prove that for a certain class of update rules, the resulting dynamics $\{x(t)\}_{t=0}^{\infty}$ always admits slow convergence (Theorem 3). As already mentioned, this class is called *opinion dependent* and means that each agent i updates her opinion according to the opinions that she observed, her s_i and α_i.

Definition 3 (Opinion Dependent Dynamics). *Let $A = (A_t)_{t=1}^{\infty}$ be a sequence of functions, where $A_t : [0,1]^{t+2} \mapsto [0,1]$. The sequence A produces the opinion dependent dynamics $x_A(t) \in [0,1]^n$ defined as*

$$x_i^A(t) = A_t(y_i(0), \ldots, y_i(t-1), a_i, s_i),$$

where $y_i(t)$ is the opinions of the neighbor that agent i meets at round t.

For example the update rule (4) is *opinion dependent* since $x_i(t) = (1-\alpha_i)\sum_{\tau=0}^{t-1} y(\tau)/t + \alpha_i s_i$ meaning that for (4), $A_t(y_0, \ldots, y_{t-1}, \alpha, s) = (1-\alpha)\sum_{\tau=0}^{t-1} y_\tau/t + \alpha s$.

One first thing that seems questionable about the above class is that $x_i(t)$ does not depend on the indices of the neighbors that i met. Another may be the fact that Definition 3 implies that whenever two agents i, j admit the same self confidence coefficient and internal opinion ($\alpha_i = \alpha_j, s_i = s_j$) then they adopt the same update rule. The only reason that these case are excluded is only to simplify notation and our results extend trivially them (see Remark 1).

We are interested in lower bounds about the convergence rate of *opinion dependent dynamics* because these bounds also hold for the convergence rate of the *no-regret dynamics* for our repeated game. Assume that each agent i updates $x_i(t)$ according to a no-regret algorithm A^i for the OCO problem where the adversary selects the functions $(1-\alpha_i)(x-b_t)^2 + \alpha_i(x-s_i)^2$. By Definition **??**, we have that $x_i(t) = A^i_t(y_i(0), \ldots, y_i(t-1))$. Notice that each agent i selects a no-regret algorithm for a different OCO problem defined by her α_i, s_i. Obviously two agents i,j with the same self confidence coefficient and internal opinion ($\alpha_i = \alpha_j, s_i = s_j$) select the algorithms A_i and A_j that admit no-regret for the same OCO problem. If we assume that in this case A_i,A_j are the same then the respective *no-regret dynamics* $\{x(t)\}_{t=1}^{\infty}$ are *opinion dependent* since

$$x_i(t) = A_t(y_i(0), \ldots, y_i(t-1), \alpha_i, s_i)$$

where $A_t : \{0,1\}^{t+2} \mapsto [0,1]$. As already mentioned this assumption is removed at the end of the section.

Theorem 3. *Let A be an* opinion dependent *update rule, which all agents use to update their opinions. For any $c > 0$ there exists an instance $I = (P, s, a)$ such that*

$$\mathbf{E}\left[\|x_A(t) - x^*\|_\infty\right] = \Omega(1/t^{1+c}),$$

where $x_A(t)$ denotes the opinion vector produced by A.

At first we show that any opinion dependent A, achieving the previous convergence rate, can be used as an estimator of the parameter $p \in [0,1]$ of Bernoulli random variable with the same asymptotic error rate. This reduction is formally stated in Lemma 5. Since we prove Theorem 3 using a reduction to an estimation problem we shall first briefly introduce some definitions and notation. For simplicity we will restrict the following definitions of estimators and risk to the case of estimating the mean of Bernoulli random variables. 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 $\theta = (\theta_t)_{t=1}^{\infty}$ 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 $\theta = (\theta_t)_{t=1}^{\infty}$ we define its risk $E_p[|\theta_t(X_1, \ldots, X_t) - p|]$, where*

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

The risk $E_p[|\theta_t(Y_1, \ldots, Y_t) - p|]$ is the expected distance of the estimated value θ_t from the parameter p, when the distribution that generated the samples is $B(p)$. For convenience we also write it as $E_p[|\theta_t - p|]$. The risk quantifies the error rate of the estimated value $\hat{p} = \theta_t(Y_1, \ldots, Y_t)$ to the real parameter p as the number of samples t grows. Since p is unknown, any meaningful estimator $\theta = (\theta_t)_{t=1}^{\infty}$ must guarantee that $\lim_{t\to\infty} E_p[|\theta_t - p|] = 0$ for all p. For example, *sample mean* has error rate $E_p[|\theta_t - p|] \leq \frac{1}{2\sqrt{t}}$.

We show now that any opinion dependent update rule A, achieving the convergence rate of Theorem 3, can be used as an estimator of the parameter $p \in [0,1]$ of a Bernoulli random variable with asymptotically the same error rate. The reduction is formally stated and in Lemma 5.

Lemma 5. *Let A an opinion dependent update rule such that for all instances I, $\lim_{t\to\infty} t^{1+c}\mathbf{E}\left[\|x_A(t) - x^*\|_\infty\right] = 0$. Then there exists an estimator $\theta_A = (\theta^A_t)_{t=1}^{\infty}$ such that for all $p \in [0,1]$,*

$$\lim_{t\to\infty} t^{1+c} E_p[|\theta^A_t - p|] = 0$$

Proof. We sketch here the main idea. For a full proof see Section B of the Appendix. For a given $p \in [0,1]$, 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 directed star graph $K_{1,2}$. The agent corresponding to the center node, c, has $\alpha_c = 1/2$ and whereas the leaf nodes have $a_{1,2} = 1$, $s_1 = 0$, $s_2 = 1$, as shown in Figure 1. It follows that the estimator θ_t with $\theta^A_t = 2x^A_c(t)$ has error $E_p[|\theta^A_t - p|] = \frac{1}{2}\mathbf{E}_{I_p}\left[\|x_A(t) - x^*\|_\infty\right]$. Meaning that $\lim_{t\to\infty} t^{1+c} E_p[|\theta^A_t - p|] = 0$ for all $p \in [0,1]$.

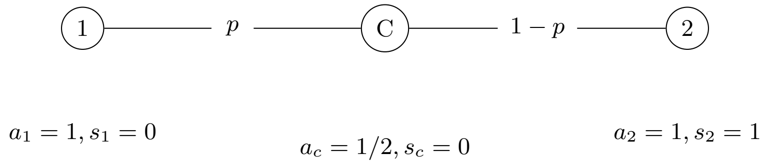

Fig. 1: The Lower Bound Instance

It follows by Lemma 5 that in order to prove Theorem 3 we just need to prove the following claim.

Claim. For any estimator $\theta = (\theta_t)_{t=1}^{\infty}$ 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 above claim states that for any estimator $\theta = (\theta_t)_{t=1}^{\infty}$, we can inspect the functions $\theta_t : \{0, 1\}^t \mapsto [0, 1]$ and then choose a $p \in [0, 1]$ such that the function $E_p[|\theta_t - p|] = \Omega(1/t^{1+c})$. As a result, we have reduced the construction of a lower bound concerning the round complexity of a dynamical process to a lower bound concerning the sample complexity of estimating the parameter p of a Bernoulli distribution.

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 *opinion dependent 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.

Before presenting Theorem **??** we briefly discuss some fundamental results concerning sample complixity lower bounds for statistical estimation. Perhaps the oldest sample complexity lower bound for estimation problems is the well-known Cramer-Rao inequality. Let the function $\theta_t : \{0, 1\}^t \mapsto [0, 1]$ such that $E_p[\theta_t] = p$ for all $p \in [0, 1]$, then

$$\mathbf{E}_p\left[(\theta_t - p)^2\right] \geq \frac{p(1-p)}{t}. \tag{10}$$

Since $\mathbf{E}_p\left[|\theta_t - p|\right]$ can be lower bounded by $\mathbf{E}_p\left[(\theta_t - p)^2\right]$ we can apply the Cramer-Rao inequality and prove Claim 4 for the *unbiased* estimators $\theta = (\theta_t)_{t=1}^{\infty}$. An estimator $\theta = (\theta_t)_{t=1}^{\infty}$ is *unbiased* if $E_p[\theta_t] = p$ for all $p \in [0, 1]$. Obviously, we need to prove the claim for any estimator θ, however this is a first indication that our claim holds.

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 [**?**] as

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

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 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 [**?**,**?**] and the very good lecture notes of Duchi, [**?**]. 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

[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.

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 θ 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 θ with samples from the distribution $B(p)$, then with probability 1 the error rate $\mathbf{E}_p\left[|\theta_t - p|\right] \in \Omega(1/t^{1+c})$. Although we do not show the sharp lower bound $\Omega(1/\sqrt{t})$ we prove that no exponential convergence rate is possible and we remark that our proof is fairly simple, intuitive, and could be of independent interest.

Lemma 6. *Let a Bernoulli estimator $\theta = (\theta_t)_{t=1}^{\infty}$ with error rate $\mathbf{E}_p\left[|\theta_t - p|\right]$. For any $c > 0$, if we select p uniformly at random in $[0,1]$ then*

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

Proof. Let an estimator $\theta = \{\theta_t\}_{t=1}^{\infty}$, 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 θ with error rate $\mathbf{E}_p\left[|\theta_t - p|\right]$ there exists another estimator θ' that satisfies the above property and $\mathbf{E}_p\left[|\theta'_t - p|\right] \le \mathbf{E}_p\left[|\theta_t - p|\right]$ for all $p \in [0,1]$. 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

$$A = \{p \in [0,1] : \lim_{t\to\infty} t^{1+c}\mathbf{E}_p\left[|\theta_t - p|\right] = 0\}$$

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

$$A_k = \{p \in [0,1] : \text{for all } t \ge k,\ t^{1+c}\mathbf{E}_p\left[|\theta_t - p|\right] \le 1\}$$

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

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

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

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

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

$$t^{1+c}\min_{0\le i\le t}|\theta_t(i) - p|\sum_{i=0}^{t}\binom{t}{i}p^i(1-p)^{t-i} \le t^{1+c}\sum_{i=0}^{t}\binom{t}{i}|\theta_t(i) - p|\, p^i(1-p)^{t-i} = t^{1+c}\mathbf{E}_p\left[|\theta_t - p|\right] \le 1/2.$$

Thus, $\mathbf{P}[p \in A_k] \le \mathbf{P}[p \in B_k]$. At first we write the set B_k in the following equivalent form

$$B_k = \cap_{t=k}^{\infty}\{p \in [0,1] : \min_{0 \le i \le t} |\theta_t(i) - p| \le 1/t^{1+c}\}$$

As a result,

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

Each value $\theta_t(i)$ "covers" length $1/t^{1+c}$ from its left and right, as shown in Figure 2, and since there are at

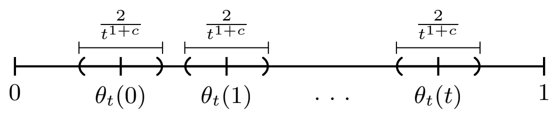

Fig. 2: Estimator output at time t

most $t+1$ such values we have for all $t \ge k$ the set

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

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

Remark 1. The only point that we use that the update rules are *opinion dependent* is in Lemma 5. It is not diffcult to see that the reduction still holds if the update rules also depend on the indices of the neighbors that an agent meets or if agents i, j with $s_i = s_j$ adopt different update rules. As a result, Theorem 3 still applies.

5 Faster Update Rules

We already discussed that the reason that opinion dependent dynamics suffer slow convergence is that the update rule depends only on the expressed opinions. In this section we provide an update rules showing that information about the graph G combined with agents that do not act selfishly, can restore the exponential convergence rate. Our update rule, depends not only on the expressed opinions of the agents but also on their indices and matrix P. Having this knowledge, one could try to come up with an update rule resembling the original update rule of the FJ model. In update rule 1, each agent could store the *most recent* opinions of the random neighbors that she meets in an array and then update her opinion according to their weighted sum (each agent knows row i of P). The problem with this approach is that the opinions of the neighbors that she keeps in her array are *outdated*, i.e. the opinion of neighbor of agent i is different than what she expressed in their last meeting. The good news are that as long as this outdatedness is bounded we can still achieve exponential convergence to the equilibrium. By bounded outdatedness we mean that there exists a number of rounds B such that all agents have met all their neighbors at least once from t to $t+B$.

Remark 2. It is necessary to know the matrix P in order for this update rule to work. We first observe that the lower bound of Section 4 also holds in case the algorithm learns the index of the chosen neighbor, since the reduction involves only two neighbors with different opinions and therefore, they are distinguishable. Therefore, if we tried to learn P by observing the frequencies of the indices of the neighbors and run update rule 1 with the empirical frequencies instead of the p_{ij}, our lower bound ensures that the rate of convergence would not be $O(1/t^{1+c})$ for any $c > 0$. Intuitively, if we know P then the algorithm converges exponentially, since the slow part of the process is learning the probabilities p_{ij} precisely.

In [**?**], they show a convergence rate guarantee for 1 assuming that there exists a such a window B. In the following we briefly summarize their result. For completeness we give here a prove tailored for our purposes. Using a simple induction we get that bounded outdatedness preserves the exponential convergence.

Algorithm 1 Asynchronous Update Rule

```
Initially $x_i(0) = s_i$ for all agent $i$.
Each agent $i$ keeps an array $M_i$ of length $d_i$.
At round $t \geq 1$ each agent $i$:
    $x_i(t) = (1-\alpha_i)\sum_{j=1}^{d_i} p_{ij} M_i[j] + \alpha_i s_i$
    Meets neighbor $W_i^t$ and learns the opinion $x_{W_i^t}(t)$.
    $M_i[W_i^t] \leftarrow x_{W_i^t}(t-1)$.
```

Lemma 7. *Let $\rho = \min_i a_i$, and $\pi_{ij}(t) \in \mathbf{N}$ be the most recent round before round t, that agent i met agent j. If for all $t \geq B$, $t - B \leq \pi_{ij}(t) \leq t-1$ then, for all $t \geq kB$, $\|x(t) - x^*\|_\infty \leq (1-\rho)^k$.*

Proof. To prove our claim we use induction on k. For the induction base $k = 1$,

$$\begin{aligned} |x_i(t) - x_i^*| &= |(1-\alpha_i)\sum_{j \in N_i} p_{ij}(x_j(\pi_{ij}(t)) - x_j^*)| \\ &\leq (1-\alpha_i)\sum_{j \in N_i} p_{ij}|(x_j(\pi_{ij}(t)) - x_j^*)| \leq (1-\rho) \end{aligned}$$

From the induction hypothesis we have for $\pi_{ij}(t) \geq (k-1)B$, that $|x_j(\pi_{ij}(t)) - x_j^*| \leq (1-\rho)^{k-1}$. For $k \geq 2$, we again have that $|x_i(t) - x_i^*| \leq (1-\rho)\sum_{j \in N_i} p_{ij}|(x_j(\pi_{ij}(t)) - x_j^*)|$. Since $t - B \leq \pi_{ij}(t)$ and $t \geq kB$, we have that $\pi_{ij}(t) \geq (k-1)B$ and the induction hypothesis applies.

In our randomized setting there does not exist fixed length window is not true but we can easily adapt this to hold with high probability. To do this observe that agent i simply needs to wait to meet the neighbor j with the smallest weight p_{ij}. Therefore, after $\log(1/\delta)/\min_j p_{ij}$ rounds we have that with probability at least $1-\delta$ agent i met all her neighbors at least once. Since we want this to be true for all agents we shall roughly take $B = 1/\min_{p_{ij}>0} p_{ij}$. In Section C of the Appendix we give the detailed argument that leads to the Lemma 8, showing that the convergence rate of update rule 1 is exponential.

Lemma 8. *Let $x(t)$ be the dynamics corresponding to update rule 1. We have*

$$\mathbf{E}\left[\|x(t) - x^*\|_\infty\right] \leq 2\exp\left(-\rho(1-\rho)\min_{ij} p_{ij}\frac{\sqrt{t}}{4\ln(nt)}\right)$$

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 2. We restate Lemma ?? for completeness.

Lemma 9. *Let $e(t)$ be a function satisfying the recursion*

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

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 $\rho \leq 1/2$ then

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

– If $\rho > 1/2$ then

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

Theorem 7. *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 **??** after t rounds satisfies*

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

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

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

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

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

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

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 $2\pi^2 \geq e^{2.5}$, Lemma 2 applies and $e(t) \leq C\sqrt{\log n} \frac{\log t^{3/2}}{t^{\min(\rho,1/2)}}$ for some universal constant C. Finally,

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

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 8. *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 $p \in [0,1]$,*

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

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 $x_i(t) = A_t(x^2, \dots, x^2)$. 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 $x_i(t) = A_t((1-x)^2, \dots, (1-x)^2)$ 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, $x_i(t) = A_t(\frac{1}{2}x^2 + \frac{1}{2}(x-y_1)^2, \dots, \frac{1}{2}x^2 + \frac{1}{2}(x-y_{t-1})^2) = f_t(y_1, \dots, y_{t-1})$. The estimator $\hat{\theta_A}$ is the following sequences $\{\hat{\theta}_t\}_{t=1}^{\infty}$

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

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 $p \in [0,1]$, we construct an appropriate instance I_p s.t. $R_p(t) = \mathbf{E}_p\left[|\hat{\theta}_t - p|\right] \leq 2\mathbf{E}_{I_p}\left[\|x^t - x^*\|_\infty\right]$. Consider the graph of Figure 1, which has a central node with $a_c = 1/2$ and $s_c = 0$ and two leaf nodes $1,2$ with $a_1 = a_2 = 1$, $s_1 = 1$ and $s_2 = 0$. The weights are $p_{c1} = p$ and $p_{c2} = 1-p$. Obviously, nodes 1 and 2 always have constant opinions, 1 and 0 respectively. Hence, in each round the center node receives either $h_1(1)$ with probability p or $h_2(0)$ with probability $1-p$.

We just need to prove that in I_p, $\mathbf{E}_p\left[|\hat{\theta}_t - p|\right] \leq \frac{1}{2}\mathbf{E}_{I_p}\left[\|x^t - x^*\|_\infty\right]$. 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 agent-1, then $Y_t = 1$ otherwise $Y_t = 0$. As a result, $\mathbf{P}\left[Y_t = 1\right] = p$ and

$$\begin{aligned}
\mathbf{E}_{I_p}\left[\|x^t - x^*\|_\infty\right] &\geq \mathbf{E}_{I_p}\left[|x_c^t - x_c^*|\right] \\
&= \mathbf{E}_{I_p}\left[\left|f_{t+1}(h_{Y_1}(1), \dots, h_{Y_t}(t)) - \frac{p}{2}\right|\right] \\
&= \mathbf{E}_p\left[\left|\frac{\widehat{\theta}_t}{2} - \frac{p}{2}\right|\right] = \frac{1}{2}R_p(t)
\end{aligned}$$

and the result follows.

We next give a rigorous measure-theoretic proof of Theorem 3.

Theorem 9. *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 $R_k(p) = k^2 \ \mathbf{E}_{X\sim P^k}\left[|\theta_k(X) - p|\right]$. 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

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

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

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

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

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

Now observe that

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

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

C Faster Update Rule

We are now going to state and prove a series of Lemmas that culminate in the proof of Lemma 8. We first turn our attention to the problem of calculating the size of window B, such that with high probability all agents have outdateness at most B. We first state a useful fact concerning the coupons collector problem.

Lemma 10. *Suppose that the collector picks coupons with mixed probabilities, where n is the number of distinct coupons. Let w be the minimum of these probabilities. If he selects $\ln n/w + c/w$ coupons, then:*

$$\mathbf{P}\left[\textit{collector hasn't seen all coupons}\right] \leq \frac{1}{e^c}$$

For convenience of reasoning, we will divide the time in "epochs". The lenght of each epoch is B, so times 1 to B belong to the first epoch etc. The next lemma is a calculation of the appropriate size of B.

Lemma 11. *Suppose we run Algorithm 1 for t rounds. If the size of each epoch is*

$$B = \frac{2}{\min_{ij} p_{ij}} \ln \frac{nt}{p}$$

then with probability at least $1 - p$ all agents pich all their neighbours at least once in every epoch.

Proof. In our setting, coupon i corresponds to the selection of neighbour i. Each node is a collector and wants to gather all $n-1$ coupons during each epoch. Suppose $d = \max_i d_i$ is the maximum degree of the graph. Then, if we set $c = \ln(\frac{nt}{p})$, using Lemma 10 we get that a node hasn't seen at least one neigbour after $c/w + \ln d/w$ samples with probability at most $\frac{p}{nt}$. This means that if we set $D = c/w + \ln d/w = \ln \frac{nt}{p}/w + \ln d/w \geq 2/w \ln \frac{nt}{p}$ when p is small enough, then the probability that a specific agent at a specific epoch hasn't collected all neighbouring opinions at least once is at most $\frac{p}{nt}$. By a simple union bound argument, we get that all agents have seen all their neighbours during all epochs with probability at least $1 - p$.

Since all neigbours are picked at least once during each epoch, the outdateness of each agent is at most twice the length of the epoch. Combining this with Lemma **??** we have the following.

Corollary 1. *If we run Algorithm 1 for t rounds, then with probability at least $1 - p$*

$$\|x^t - x^*\|_\infty \leq (1-\rho)^{\frac{t}{2B}} \leq \exp\left(-\frac{\rho t \min_{ij} p_{ij}}{2\ln(\frac{nt}{p})}\right)$$

We now prove Lemma 8 using the previous results.

Proof. Let $u(t) = \|x^t - x^*\|_\infty$ and $w = \min_{ij} p_{ij}$. From Corollary 1 we obtain:

$$\mathbf{P}\left[u(t) > \exp\left(-\frac{\rho w t}{2\ln(\frac{nt}{p})}\right)\right] \leq p$$

for every probability $p \in [0,1]$. Also, since all the parameters of the problem lie in $[0,1]$, we have

$$\mathbf{E}\left[u(t) | u(t) > r\right] \leq 1$$

Now, by the conditional expectations identity, we get:

$$\begin{aligned}\mathbf{E}\left[u(t)\right] &= \mathbf{E}\left[u(t)|u(t) > r\right]\mathbf{P}\left[u(t) > r\right] + \mathbf{E}\left[u(t)|u(t) \leq r\right]\mathbf{P}\left[u(t) \leq r\right] \\ &\leq p + r\end{aligned}$$

where $r = \exp\left(-\frac{\rho w t}{2\ln(\frac{nt}{p})}\right)$. If we set $p = \exp\left(-\frac{\rho w\sqrt{t}}{2\ln nt}\right)$, then:

$$\mathbf{E}\left[u(t)\right] \leq \exp\left(-\frac{\rho w\sqrt{t}}{2\ln nt}\right) + \exp\left(-\frac{\rho w t}{2\ln(\frac{nt}{p})}\right)$$

We now evaluate r for our choice of probability p:

$$\begin{aligned} r &= \exp\left(-\frac{\rho w t}{2\ln\left(\frac{nt}{p}\right)}\right) \\ &= \exp\left(-\frac{\rho w t}{2\ln\left(\frac{nt}{\exp\left(-\frac{\rho w\sqrt{t}}{2\ln nt}\right)}\right)}\right) \\ &= \exp\left(-\frac{\rho w t}{2\ln nt + 2\frac{\rho w\sqrt{t}}{2\ln nt}}\right) \\ &\leq \exp\left(-\frac{\rho w t}{4\ln(nt)\sqrt{t}}\right) \\ &= \exp\left(-\frac{\rho w\sqrt{t}}{4\ln(nt)}\right)\end{aligned}$$

Using the previous calculation, we obtain:

$$\begin{aligned}\mathbf{E}\left[u(t)\right] &\leq \exp\left(-\frac{\rho w\sqrt{t}}{2\ln(nt)}\right) + \exp\left(-\frac{\rho w\sqrt{t}}{4\ln(nt)}\right) \\ &\leq 2\exp\left(-\frac{\rho w\sqrt{t}}{4\ln(nt)}\right) \\ &= 2\exp\left(-\rho\min_{ij} p_{ij}\frac{\sqrt{t}}{4\ln(nt)}\right)\end{aligned}$$