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 [BKO11] the following *opinion formation game* was introduced. A weighted directed graph $G(V, E, w)$ is assumed were the vertices stand for the agent and the acres the social influence among them. Each agent $i \in V$ poses 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 incuring 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 [BKO11] they proved that the above game always admits a *Pure Nash Equilibrium* (PNE) $x^* \in [0, 1]^n$ and studied the efficiency of x^*. The 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 the *best response* of $x(t-1)$,

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

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 [GS14] the proved that the $x(t)$ always converges to the PNE x^* and they 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 exits 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 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 is seems quite unatural in today's huge social networks (e.g. Facebook, Twitter e.t.c.) in which each agent may have several hundreds of friends. With this in mind 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 the following variant of the opinion formation game with random payoffs.

Definition 1. *For a given opinion vector $x \in [0, 1]^n$, the 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 $\alpha_i = w_i/(\sum_{j \in N_i} w_{ij} + w_i)$

Our variant has a very natural interpretation: 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 $x \in [0, 1]^n$. As a result, the coefficient α_i measures the reluctancy of agent i to adopt an opinion other than s_i, while the p_{ij} can be seen the *real* influence that j poses on i. Following the common belief that we are influence more by those we interact oftently, the *real* influence p_{ij} can intepreted as measure of how often i meets with j. In our random payoff variant p_{ij} (Definition 1) is exactly the frequency that i meets j.

Equation (3) also helps us to establish the existence of PNE for the our random payoff variant. The notion of *Pure Nash Equilibrium* is properly extended in our case with respect to the expected cost of each agent. Namely, $x^* \in [0,1]$ is PNE if and only if $\mathbf{E}\left[C(x_i, x_{-i})\right] \leq \mathbf{E}\left[C(x_i', x_{-i})\right]$ 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^* with 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 this 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

In this work we study 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 the random cost $C_i(x_i(t), x_{-i}(t))$, where $x(t) \in [0,1]$ is the opinion vector at round t. We are intested in simple and natural update rules that the agents can adopt such as the resulting opinion vector $x(t)$ converges to x^*. Moreover we require that this update rules be a *rational* behavioral assumption for selfish agents.

In Section 2, we study the convergence properties of $x(t)$ if all agents adopt *fictitious play* as their update rule. At round $\tau < t$ each agent i experiences disagreement cost

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

where W_i^τ denotes the agent that i met at round τ. For agent i updating her opinion according to *fictitious play* means that she selects selects her opinion $x_i(t)$ in order to minimize her aggregated cost until round t,

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

Generally speaking if all agents adopt *fictitious play* the resulting strategy vector does not converge to the equilibrium. However, in our case if all agent adopt update rule 4 the resulting opinion vector $x(t)$ converges to x^* 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 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.*

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 $O(|E|)$, where E is the number of edges of graph G. 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 $\log(2\ 10^9)$.

In Section 3 we argue that update rule (4) is a *rational* behavioral assumption for selfish agents in the sense that it ensures *no-regret* for the agents that adopt it. Namely if agent i uses (4) to select $x_i(t)$, then the average disagreement cost that she experiences is similar to the average cost that she would experience by expressing the best fixed opinion. The latter holds no matter the way that the other agents select their opinions and the agents that i meets. The main result of this section is Theorem 2,

Theorem 2. *Consider the fuction $f : [0,1]^2 \mapsto [0,1]$ 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 $x_t = \mathrm{argmin}_{x\in[0,1]} \sum_{\tau=1}^{t-1} f(x,b_\tau)$ then for all t,*

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

The no-regret property of *fictitious play* follows by direct application of Theorem 2 with $f(x,b) = (1-\alpha_i)(x-b)^2 + \alpha(x-s_i)^2$ and $b_t = x_{W_i^t}(t)$, where α_i, s_i are respectively the self confidence coefficient,the internal opinion of i and W_i^t the neighbor that agent i meets at round t.

Even though our update rule (4) 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 achieves convergence rate $O(e^{-\rho t})$ [GS14]. In section 4 we explain this exponential gap. The reason is that the update rule (4) only depends on the opinions of the agents that agent i meets, α_i, and s_i. We call such update rules "opinion dependent". Observe that fictitious play (4) is opinion dependent. In Theorem 3 we show that for any opinion dependent update rule there exists an instance $I = (P, s, \alpha)$ such that $\mathrm{poly}(1/\varepsilon)$ 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.

In order to bypass the lower bound of Theorem 3, the update rules must use more information than simply the opinions of the neighbors, e.g. the indices of the neighbors that agent i meets. Observe that update rules that ensure no regret for the agents must be opinion dependent; Theorem 3 rules out the possibility that they achieve exponential convergence rate. In Section **??** we also argue that update rules that converge exponentially fast, are an unrealistic choice for modelling the true behavior of the agents because of their complicated form. Precisely, we explicitly construct 2 update rules that, while they achieve exponential convergence rate, they are indeed complicated. The first one if a function of the opinions and the indices of the agents that i meets. The second one is a function of the opinions, and the number of neighbors of each agent i. In conclusion, the results of this work indicate that natural dynamics in our limited information exchange setting come at the price of slow convergence rate.

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 Fictitious Play Convergence Rate

In this section we prove that dynamics $x(t) \in [0,1]^n$ produced by the update rule (4) converges to the unique equilibrium $x^* \in [0,1]^n$. For an instance (P,s,α) the opinion vector $x(t) \in [0,1]^n$ is defied as follows:

– Initially all agents adopt their internal opinion, $x_i(0) = s_i$
– At round $t \geq 1$ each agent i
  - selects

fictitious play described as Algorithm **??** 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 **??** 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 $\rho = \min_i a_i$ and C is a universal constant.

At first we present the high level idea of the proof of Theorem 8. According to Algorithm **??**, each agent i at round $t \geq 1$ updates her opinion as $x_i(t) = \mathrm{argmin}_{x\in[0,1]}\sum_{\tau=1}^{t} C_i^\tau(x, x_{W_i^\tau}(\tau-1))$ 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 $\mathbf{E}\left[\|x(t) - x^*\|_\infty\right]$, 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 $|\frac{\sum_{\tau=1}^{t}\mathbf{1}[W_i^\tau=j]}{t} - p_{ij}|$ 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 $\mathbf{P}\left[W_i^\tau = j\right] = p_{ij}$ 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,

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

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

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

We get (7) from the induction step and (8) from inequality (6). Similarly, we can prove that $x_i(t) \geq x_i^* - \delta(t) - (1-\alpha)\frac{\sum_{\tau=1}^{t} e(\tau)}{t}$. As a result $\|x(t) - x^*\|_\infty \leq e(t)$ and the induction is complete. Therefore, we have that with probability at least $1-p$, $\|x(t) - x^*\|_\infty \leq e(t)$ for all $t \geq 1$.

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 $e(t) \leq \sqrt{2.5\ln(D)}\frac{(\ln t)^{3/2}}{t^{\min(\rho,\,1/2)}}$.

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

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 $(b_1, \ldots, b_{t-1})$ 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

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

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 5. *Consider the fuction $f : [0,1]^2 \mapsto [0,1]$ 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 $x_t = \operatorname{argmin}_{x\in[0,1]} \sum_{\tau=1}^{t-1} f(x,b_\tau)$ then for all t,*

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

Theorem 2 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 2 applies and

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

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

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

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

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

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,

$$\left|z^t - y^t\right| \leq \frac{1-\alpha}{t+1}. \tag{9}$$

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 (9) 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

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

4 Lower Bound for Opinion Dependent Dynamics

As we have already discussed, for any instance I with $\rho \ge 1/2$, *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, $\mathbf{E}_I\left[\|x(t) - x^*\|_\infty\right] = O(\mathrm{e}^{-t/2})$. In this section we investigate whether there exists another opinion dependent update rule that the agents can select and ensures a better convergence rate to the equilibrium. We show the following lower bound for the convergence rate. We first give a rigorous definition of update rules that only use the expressed opinions of the neighbors of an agent.

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

$$x_{A,i}(t) = A_t(y_i(1), \ldots, y_i(t-1), a_i, s_i),$$

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

For convenience we restate our lower bound, Theorem 3.

Theorem 6. *Let A be a sequence of functions with corresponding opinion dependent dynamics $x_A(t)$. For any $c > 0$, there exists an instance I such that $\mathbf{E}_I\left[\|x_A(t) - x^*\|_\infty\right] = \Omega(1/t^{1+c})$.*

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. The reduction is formally stated in Lemma 5. Since we prove **??** 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 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 $\theta = \{\theta_t\}_{t=1}^{\infty}$ we define its convergence rate $R_p(\theta_t) = E_p[|\theta_t(X_1, \ldots, X_t) - p|]$, 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 quantity $E_p[|\theta_t(Y_1, \ldots, Y_t) - p|]$ is the expected distance of the estimated value $\hat{\theta}_t$ from the parameter p, when the distribution that generated the samples is $B(p)$. To simplify notation we will also write it as $E_p[|\hat{\theta}_t - p|]$. The error rate $R_p(\theta_t)$ quantifies the rate of convergence of the estimated value $\hat{p} = \theta_t(Y_1, \ldots, Y_t)$ 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 $R_p(\theta_t) \le \frac{1}{2\sqrt{t}}$ for any $p \in [0,1]$ and clearly satisfies the above requirement.

We show now that any opinion dependent algorithm 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 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 $p \in [0,1]$, 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 $\alpha_c = 1/2$ 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 $\theta_A^t = x_A^c(t)$ has error $R_p(\theta_t) = \mathbf{E}_{I_p}\left[\|x_A(t) - x^*\|_\infty\right]$.

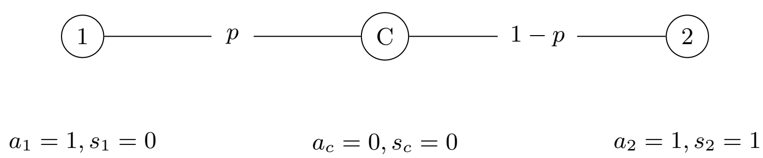

Fig. 1: The Lower Bound Instance

Meaning that if A does not satisfy Theorem **??** 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 4 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

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

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 $|\theta_t - p| \le 1$ (making the natural assumption that estimators do not output values outside $[0,1]$). Simply setting $p = 1/2$ in inequality (10) would give us a p satisfying the requirements of Claim 4. 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 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 $X_1, \ldots, X_t$ 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 and 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

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

[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.

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

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

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,

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

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, as shown in Figure 2 , and since there are at

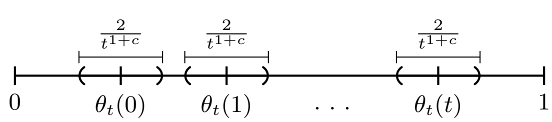

Fig. 2: Estimator output at time t

most $t+1$ such values we have for all $t \geq k$ 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 $\mathbf{P}[p \in B_k] = 0$.

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 2 update rules showing that information about the graph G combined with agents that do not act selfishly, can restore the exponential convergence rate. Our first update rule, depends not only on the expressed opinions of the agents but also on their indices and matrix P.

In [BT97], section 6.3.5, they show a convergence rate guarantee for 1 assuming that there exists a window of B rounds such that all agents get to meet all their neighbors at least once from round t to round $t+B$. In our randomized setting this is not true but we can easily adapt this to hold with high probability. In our problem agent i simply needs to wait to meet the neighbor j with the smallest weight p_{ik}. Therefore, after

Algorithm 1 Tsitsiklis

1: $x_i(0) \leftarrow s_i$.
2: Let d_i be the number of non-negative entries of row i of P.
3: Keep an array A of length d_i.
4: At round t:
  5: Meet neighbor k, get the opinion $x_k(t-1)$ and the index k.
  6: $A_k \leftarrow x_k(t-1)$.
  7: $x_i(t) = \sum_{j=1}^{d_i} p_{ij} A_j$

$\log(1/\delta)/\min_j p_{ij}$ rounds we have that with probability at least $1-\delta$ agent i saw all her neighbors at least once. Since we want this to be true for all agents we shall roughly take $B = 1/\min_{p_{ij}>0} p_{ij}$. Omitting the details of this straightforward adaptation we get that with high probability

$$\|x(t) - x^*\|_\infty = O(\|P\|_\infty^{t/(B+1)}).$$

Precisely, our update rule depends on the expressed opinions, the number of neighbors of each agent, and the number of agents n.

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 2 an agent could just average the opinions that she gets in this window. Unfortunately this would again result in a 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 $S_1, \dots, S_t$ be t i.i.d samples drawn from P, i.e. $(S_1, \dots, S_t) \sim P^t$. The empirical distribution $\hat{P}_t$ is the following estimator of the density of P.

$$\hat{P}_t(A) = \frac{\sum_{i=1}^t \mathbf{1}[S_i \in A]}{t}, \tag{11}$$

where $A \subseteq [n]$. In words, $\hat{P}_t$ simply counts how many times the value i appeared in the samples $S_1, \dots S_t$. 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 $\{1, \dots, n\}$ and let $S_{\mathcal{A}}(t)$ be the Vapnik-Chervonenkis shatter coefficient, defined by*

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

Then

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

Algorithm 2 Graph Aware Update Rule

1: $x_i(0) \leftarrow s_i$.
2: $M_1 = O(\ln(n/\varepsilon))$, $M_2 = O(d^2 \ln(d))$
3: **for** $l = 1, \dots, \ln(1/\varepsilon)$ **do**
4: Keep a set A of tuples $(x, \mathrm{freq}(x))$ of and an array B of length M_1.
5: **for** $j = 1, \dots, M_1$ **do**
6: **for** $k = 1, \dots, M_2$ **do**
7: Get the opinion X_k of a random neighbor.
8: **if** X_j is not in A **then**
9: Insert $(X_k, 1)$ to A.
10: **else**
11: $(X_k, \mathrm{freq}(X_k)) \leftarrow (X_k, \mathrm{freq}(X_k) + 1)$.
12: **end if**
13: **end for**
14: Divide all frequencies of A by M_2.
15: Round all frequencies of ÎŚ to the closest multiple of $1/d_i$.
16: $B(j) \leftarrow \alpha_i \sum_{x \in A} \mathrm{freq}(x) + (1 - \alpha_i) s_i$.
17: **end for**
18: $x_i(t) \leftarrow \mathrm{majority}_j B(j)$.
19: **end for**

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 $\mathbf{E}\left[\|x_t - x^*\|_\infty\right] \le \varepsilon$.*

Proof. According to the update rule 2 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 $\mathcal{A} = \{\{1\}, \{2\}, \dots, \{d_i\}\}$ and use Lemma 7 to get that

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

since $S_{\mathcal{A}} \le n$. 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 14 corresponds to the empirical distribution of equation (11). 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 $\mathbf{E}\left[\|x_T - x^*\|_\infty\right] = (1 - \varepsilon)\varepsilon + \varepsilon \le 2\varepsilon$.

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.

BT97. Dimitri P. Bertsekas and John N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, 1997.

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.

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 8. *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 $e(t) \leq \sqrt{2\ln(D)}\frac{(\ln t)^{3/2}}{t^{\min(\rho,\, 1/2)}}$.

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

$$e(t+1) = e(t)\left(1 - \frac{\rho}{t+1}\right) + \delta(t+1) - \delta(t) + \frac{\delta(t)}{t+1} \tag{12}$$

For $t = 0$ we have that

$$e(1) = (1-\rho)e(0) + \delta(1) = (1-\rho)e(0) + \sqrt{\ln D} \tag{13}$$

Observe that for $D > \mathrm{e}^2$, $\delta(t)$ is decreasing for all $t \geq 1$. Therefore, $\delta(t+1) - \delta(t) + \frac{\delta(t)}{t+1} \leq \frac{\delta(t)}{t+1}$ and from equations (12) and (13) we get that for all $t \geq 0$

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

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

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

We observe that

$$\sum_{\tau=1}^{t}\frac{\sqrt{\ln\tau}}{\tau^{3/2-\rho}} \leq \int_{\tau=1}^{t}\frac{\sqrt{\ln\tau}}{\tau^{3/2-\rho}}\mathrm{d}\tau \tag{14}$$

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 **??** 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,

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

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

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

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

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

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 10. *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 $A = [0, 1]$.