Contribution Title*

First Author[1[0000–1111–2222–3333]], Second Author[2,3[1111–2222–3333–4444]], and Third Author[3[2222––3333–4444–5555]]

[1] Princeton University, Princeton NJ 08544, USA
[2] Springer Heidelberg, Tiergartenstr. 17, 69121 Heidelberg, Germany `lncs@springer.com`
http://www.springer.com/gp/computer-science/lncs
[3] ABC Institute, Rupert-Karls-University Heidelberg, Heidelberg, Germany
`{abc,lncs}@uni-heidelberg.de`

Abstract. The best paper.

Keywords: First keyword · Second keyword · Another keyword.

* Supported by MOP

1 Preliminaries

In the Friedkin-Johsen model an undirected graph $G(V,E)$ is assumed, where V denotes the agents and E the social relations between them. Each agent i poses an internal opinion $s_i \in [0,1]$ and a self confidence coefficient $\alpha_i \in [0,1]$. At each time step $t \geq 1$, agent i updates her opinion as follows:

$$x_i(t) = (1-\alpha_i)\frac{\sum_{j\in N_i} x_j(t-1)}{d_i} + \alpha_i s_i$$

d_i denotes the degree of node i and N_i the set of her neighbors. The above process have been studied and it is known to converge to a unique equilibrium point $x^* \in [0,1]^n$, where $n = |V|$.

2 No Regret

We consider the following online convex optimization problem. At each time step t, the player i selects a real number x^t and a fucntion $f^t(x)$ arrives. The player then suffers $f^t(x^t)$ cost. The functions $f^t(x)$ have the following form:

$$f^t(x) = \alpha(x-s_i)^2 + (1-\alpha)(x-a_t)^2$$

where $s_i, \alpha \in [0,1]$ and are independent of t and $a_t \in [0,1]$. In other words, the function f^t is uniquely determined by the number a_t.

We show that for this class of function $fictitiousplay$ admits no regret.

Theorem 1. *Let* $x^t = \arg min_{x\in[0,1]} \sum_{\tau=1}^{t-1} f^\tau(x)$ *then*

$$\sum_{t=1}^{T} f^t(x^t) \leq min_{x\in[0,1]} \sum_{t=1}^{T} f^t(x) + O(logT)$$

At first we prove that the sequence $y^t = argmin_{x\in[0,1]} \sum_{\tau=1}^{t} f^\tau(x)$ admits no regret.

Lemma 1. *Let* $y^t = argmin_{x\in[0,1]} \sum_{\tau=1}^{t} f^\tau(x)$ *then*

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

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

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

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

Now we can derive some intuition for the reason that $fictitiousplay$ admits no regret. Since the cost incured by the sequence y^t is at most that of the best fixed strategy, we can compare the cost incured by x^t with that of y^t. However, 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 2. *For all t, $\left|x^t - y^t\right| \leq \frac{1-\alpha}{t}$*

Proof. By definition $x^t = \alpha s_i + (1-\alpha)\frac{\sum_{\tau=1}^{t-1} a_\tau}{t-1}$ and $y^t = \alpha s_i + (1-\alpha)\frac{\sum_{\tau=1}^{t} a_\tau}{t}$.

$$\begin{aligned}\left|x^t - y^t\right| &= (1-\alpha)\left|\frac{\sum_{\tau=1}^{t-1} a_\tau}{t-1} - \frac{\sum_{\tau=1}^{t} a_\tau}{t}\right| \\ &= (1-\alpha)\left|\frac{\sum_{\tau=1}^{t-1} a_\tau - (t-1)a_t}{t(t-1)}\right| \\ &\leq \frac{1-\alpha}{t}\end{aligned}$$

The last inequality follows from the fact that $a_\tau \in [0,1]$.

Lemma 3. *For all t, $f^t(x^t) \leq f^t(y^t) + 2\frac{1-\alpha}{t} + \frac{(1-a)^2}{t^2}$.*

Proof.

$$\begin{aligned}f^t(x^t) &= \alpha(x^t - s_i)^2 + (1-\alpha)(x^t - a_t)^2 \\ &\leq \alpha(y^t - s_i)^2 + 2\alpha\left|y^t - s_i\right|\left|x^t - y^t\right| + \alpha\left|x^t - y^t\right|^2 \\ + (1-\alpha)(y_t - a_t)^2 + 2(1-\alpha)\left|y^t - a_t\right|\left|x^t - y^t\right| + (1-\alpha)\left|x^t - y^t\right|^2 & \\ &\leq f^t(y^t) + 2\left|x^t - y^t\right| + \left|y^t - x^t\right|^2 \\ &\leq f^t(y^t) + 2\frac{1-\alpha}{t} + \frac{(1-\alpha)^2}{t^2}\end{aligned}$$

Now Theorem 1 easily follows since:

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

3 Full memory with limited information exchange

In each round, every agent learns a neighbour's opinion and updates the appropriate cell in memory. Because an agent learns only one opinion in each round, he uses outdated information about his other neighbours in order to compute his opinion. We first introduce some convinient notation.

Definition 1. *$\pi_{ij}(t)$ is the last time that i learned j's opinion until time t.*

Obviously, if at time t an agent i learned k's opinion, then $\pi_{ik}(t) = t$. The rule, according to which agent i updates his opinion at time t is the following(for simplicity, assume that $\alpha_i = \frac{1}{d+1}$ and that the graph is d-regular):

$$x_i(t+1) = (1-\alpha_i)\frac{\sum_{j\in N_i} x_j(\pi_{ij}(t))}{d_i} + \alpha_i s_i$$

Notice that in contrast with the traditional model, instead of $x_j(t)$ we have $x_j(\pi_{ij}(t))$. That is because the information about j is probably outdated. Denote with x^* the solution vector of the system. We would like to prove that such a process converges to the solution of the linear system involving the laplacian matrix of the graph. In order to analyse the algorithm, we will divide the time into "epochs". Each epoch has a length $D \geq d$, so the first epoch is from time 1 to time D, the second from $D+1$ to $2D$ e.t.c. Time 0 does not belong at any epoch. The numbering of epochs begins from 1. We will also assume that during each epoch, each agent picks every one of his neighbours for update at least once. Later we are going to find a suitable epoch length D, such that this holds with high probability. In other words, for a specific time in epoch i every agent has information which has "bounded" outdateness, since every coordinate was updated at least once during the $i-1$ epoch. We are going to use this fact to prove the following lemma.

Lemma 4. *Denote with x^* the solution vector of the system. For every time t, which belongs to epoch T, it holds:*

$$\|x(t) - x^*\|_\infty \leq (1-\alpha_i)^T \|x(0) - x^*\|_\infty$$

Proof. We are going to use induction in time. The base case is for time $t = 1$. We are going to prove that:

$$\|x(1) - x^*\|_\infty \leq (1-\alpha_i)\|x(0) - x^*\|_\infty$$

We assume that everybody has the same initial vector $x(0)$ stored in memory, so for each i holds:

$$x_i(1) = (1-\alpha_i)\frac{\sum_{j \in N_i} x_j(0)}{d} + \alpha_i s_i(1)$$

Since x^* is the solution of the system, we have:

$$x_i^* = (1-\alpha_i)\frac{\sum_{j \in N_i} x_j^*}{d} + \alpha_i s_i(2)$$

Subtracting (2) from (1) we get:

$$\begin{aligned}
|x_i(1) - x_i^*| &= |(1-\alpha_i)\frac{\sum_{j \in N_i}(x_j(0) - x_j^*)}{d}| \\
&\leq (1-\alpha_i)\frac{\sum_{j \in N_i}|x_j(0) - x_j^*|}{d} \\
&\leq (1-\alpha_i)\|x(0) - x^*\|_\infty \\
\Rightarrow \|x(1) - x^*\|_\infty &\leq (1-\alpha_i)\|x(0) - x^*\|_\infty
\end{aligned}$$

So the base case is verified. Let the inductive hypothesis hold for all times until t. Suppose that time $t+1$ belongs to epoch T. We fix an agent i. Then, if $\pi_{ij}(t)$ belongs to the previous epoch, by the induction hypothesis it holds:

$$|x_j(\pi_{ij}(t)) - x_j^*| \leq (1-\alpha_i)^{T-1} \|x(0) - x^*\|_\infty$$

On the other hand, if $\pi_{ij}(t)$ belongs to the current epoch T, by the induction hypothesis we have:

$$|x_j(\pi_{ij}) - x_j^*| \leq (1-\alpha_i)^T \|x(0) - x^*\|_\infty \leq (1-\alpha_i)^{T-1} \|x(0) - x^*\|_\infty$$

So, for every neighbour of agent i, the value that i stores for this neighbour is "close" to the optimal. Now, using again equation (1), we have:

$$\begin{aligned}
|x_i(t+1) - x_i^*| &= |(1-\alpha_i)\frac{\sum_{j \in N_i}(x_j(\pi_{ij}(t)) - x_j^*)}{d}| \\
&\leq (1-\alpha_i)\frac{\sum_{j \in N_i}|x_j(\pi_{ij}(t)) - x_j^*|}{d} \\
&\leq (1-\alpha_i)\frac{\sum_{j \in N_i}\|x(\pi_{ij}(t)) - x^*\|_\infty}{d} \\
&\leq \frac{\sum_{j \in N_i}(1-\alpha_i)^{T-1}\|x(0) - x^*\|_\infty}{d+1} \\
&= (1-\alpha_i)(1-\alpha_i)^{T-1}\|x(0) - x^*\|_\infty \\
&= (1-\alpha_i)^T \|x(0) - x^*\|_\infty
\end{aligned}$$

Corollary 1. *If the algorithm runs for $O(D\log\frac{1}{\epsilon})$ iterations, it gets ϵ-close to the solution x^*.*

We now turn our attention to the problem of computing the appropriate length D of epochs. A useful fact concerning the coupons collector problem is the following.

Lemma 5. *Suppose that the collector picks $n\ln n + cn$ coupons, where n is the number of distinct coupons. Then:*

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

Theorem 2. *After t rounds, with probability at least $1-p$: $\|x^t - x^*\|_\infty \leq (1-a)^{\frac{t}{2d\ln(\frac{dt}{p})}}$*

Proof. In our setting, coupon i corresponds to the selection of neighbour i. Each node is a collector and wants to gather all d_i coupons during each epoch. Suppose $d = \max_i d_i$ is the maximum degree of the graph. Then, if we set $n = d$ and $c = \ln(\frac{dt}{p})$, using the previous lemma we get that a node hasn't seen at least one neigbour after $cd + d\ln d$ samples with probability at most $\frac{p}{dt}$. This means that if we set $D = cd + d\ln d = d\ln\frac{dt}{p} + d\ln d \geq 2d\ln\frac{dt}{p}$ 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 $\frac{p}{dt}$. By a simple union bound argument, we get that all agents have seen all their neighbours during all epochs with probability at least $1-p$. We observe that at time t approximately $\frac{t}{D}$ epochs have passed. Therefore, using the previous result about the convergence rate, we get that with probability at least $1-p$:

$$\|x^t - x^*\|_\infty \leq (1-a)^{\frac{t}{D}} \leq (1-a)^{\frac{t}{2d\ln(\frac{dt}{p})}}$$

We are now going to translate this result to one that involves the expected value of the error. First of all, we set $Err(t) = \|x^t - x^*\|_\infty$ the error after t rounds and using the conditional expectations identity, we get:

$$\begin{aligned}&\mathbb{E}[Err(t)] = \mathbb{E}[Err(t)|Err(t) > r]\mathbb{P}[Err(t) > r] + \mathbb{E}[Err(t)|Err(t) \leq r]\mathbb{P}[Err(t) \leq r]\\ &\leq p + r\end{aligned}$$

where $p = (1-a)^{\frac{\sqrt{t}}{2d\ln dt}}$ and $r = (1-a)^{\frac{t}{2d\ln(\frac{dt}{p})}}$. We now evaluate r:

$$\begin{aligned} r &= (1-a)^{\frac{t}{2d\ln\left(\frac{dt}{p}\right)}}\\ &= (1-a)^{\frac{t}{2d\ln\left(\frac{dt}{(1-a)^{\frac{\sqrt{t}}{2d\ln dt}}}\right)}}\\ &= (1-a)^{\frac{t}{2d\ln dt + 2d\frac{\sqrt{t}}{2d\ln dt}\ln\left(\frac{1}{(1-a)}\right)}}\\ &\leq (1-a)^{\frac{t}{4d\ln(dt)\ln\left(\frac{1}{1-a}\right)\sqrt{t}}}\\ &= (1-a)^{\frac{\sqrt{t}}{4d\ln(dt)\ln\left(\frac{1}{1-a}\right)}}\end{aligned}$$

Using the previous calculation, we obtain:

$$\mathbb{E}[Err(t)] \leq (1-a)^{\frac{\sqrt{t}}{2d\ln dt}} + (1-a)^{\frac{\sqrt{t}}{4d\ln(dt)\ln\left(\frac{1}{1-a}\right)}} \leq 2(1-a)^{\frac{\sqrt{t}}{4d\ln(dt)\ln\left(\frac{1}{1-a}\right)}}$$

This means that the error decreases subexponentially with the number of rounds in expectation.

4 Constant Memory with full window

In this work we investigate an variant of the above process. We denote by V the set of agents, and $N_i \subset V$ the set of neighbors of agent i. Let $x(t) \in [0,1]^n$ be the opinion vector at round t. We denote by $x_i(t)$ the opinion of agent i at round t. At round $t+1$, $x(t+1)$ is constructed in the following way. At first, each agent i gets to see the opinion $x_j(t)$, where $j \in N_i$ is picked uniformly at random from the set of her neighbors. Let W_i^t be the random variable corresponding to the selected neighbor. Now each agent $i \in V$ updates her opinion as follows

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

We remark that each agent i get to see *only* the opinion $x_{W_i^t}$ and not the label W_i^t of her neighbor.

Let $x^* \in [0,1]^n$ be the unique equilibrium point of the given instance I. We prove that the above stochastic process has the following convergence rate to x^*.

Theorem 3.

$$\mathbf{E}\left[\|x^t - x^*\|_\infty\right] \leq \sqrt{\frac{\log(n)}{t}}$$

We start by stating the standard Hoeffding bound

Lemma 6 (Hoeffding's Inequality). *Let $X = (X_1 + \cdots + X_t)/t$, $X_i \in [0,1]$. Then,*

$$\mathbb{P}\left[|X - \mathbf{E}[X]| > \lambda\right] < 2e^{-2t\lambda^2}.$$

Lemma 7. *With probability at least $1-p$, $\|x^t - x^*\|_\infty \leq e(t)$, where $e(t)$ satisfies the following recursive relation*

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

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

Corollary 2. *The function $e(t)$ satisfies the following recursive relation*

$$e(t+1) - e(t) + \alpha\frac{e(t)}{t+1} = \delta(t+1) - \delta(t) + \frac{\delta(t)}{t+1}$$

Proof. As we have already mentioned for any instance I there exists a unique equilibrium vector x^*. Since $W_i^\tau \sim \mathrm{U}(N_i)$ we have that $\mathbf{E}\left[x^*_{W_i^\tau}\right] = \frac{\sum_{j\in N_i} x_j^*}{|N_i|}$. Since W_i^τ are independent random variables, we can use Hoeffding's inequality (Lemma 3) to get

$$\mathbb{P}\left[\left|\frac{\sum_{\tau=1}^{t} x^*_{W_i^\tau}}{t} - \frac{\sum_{j\in N_i} x_j^*}{|N_i|}\right| > \delta(t)\right] < \frac{6}{\pi^2}\frac{p}{nt^2},$$

where $\delta(t) = \sqrt{\frac{\log(\pi^2 n t^2/(6p))}{t}}$. Therefore, by the union bound

$$\mathbb{P}\left[\text{for all } t \geq 1 : \max_{i\in V}\left|\frac{\sum_{\tau=1}^{t} x^*_{W_i^\tau}}{t} - \frac{\sum_{j\in N_i} x_j^*}{|N_i|}\right| > \delta(t)\right] \leq$$
$$\sum_{t=1}^{\infty}\mathbb{P}\left[\max_{i\in V}\left|\frac{\sum_{\tau=1}^{t} x^*_{W_i^\tau}}{t} - \frac{\sum_{j\in N_i} x_j^*}{|N_i|}\right| > \delta(t)\right] \leq$$
$$\sum_{i=1}^{\infty}\frac{6}{\pi^2}\frac{1}{t^2}\sum_{i=1}^{n}\frac{p}{n} = p$$

As a result with probability $1-p$ we have that for all t

$$\left|\frac{\sum_{\tau=1}^{t} x^*_{W_i^\tau}}{t} - \frac{\sum_{j\in N_i} x^*_j}{|N_i|}\right| \le \delta(t) \tag{1}$$

We will our claim by induction. We assume that $\|x^\tau - x^*\|_\infty \le e(\tau)$ for all $\tau \le t-1$. Then

$$\begin{aligned}
x_i(t) &= (1-\alpha_i)\frac{\sum_{\tau=1}^{t} x_{W_i^\tau}(\tau-1)}{t} + \alpha_i s_i \\
&\le (1-\alpha_i)\frac{\sum_{\tau=1}^{t} x^*_{W_i^\tau} + \sum_{\tau=1}^{t} e(\tau-1)}{t} + \alpha_i s_i \qquad (2)\\
&\le (1-\alpha_i)\left(\frac{\sum_{\tau=1}^{t} x^*_{W_i^\tau}}{t} + \frac{\sum_{\tau=0}^{t-1} e(\tau)}{t}\right) + \alpha_i s_i \\
&\le (1-\alpha_i)\left(\frac{\sum_{j\in N_i} x^*_j}{|N_i|} + \delta(t) + \frac{\sum_{\tau=0}^{t-1} e(\tau)}{t}\right) + \alpha_i s_i \qquad (3)\\
&\le x^*_i + \delta(t) + (1-\alpha)\left(\frac{\sum_{\tau=0}^{t-1} e(\tau)}{t}\right)
\end{aligned}$$

We get (2) from the induction step and (3) from inequality (1). 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 $\|x_i(t) - x^*\|_\infty \le e(t)$.

In order to bound the convergence time of the system, we just need to bound the convergence rate of the function $e(t)$. The following lemma provides us with a simple upper bound for the convergence rate of our process.

Lemma 8.

$$e(t) \le \frac{e(0)}{t^\alpha} + \frac{O\left(\sqrt{\log(\frac{nd}{p})}\right)}{t^\alpha}\sum_{\tau=1}^{t}\tau^\alpha\frac{\sqrt{\log\tau}}{\tau^{3/2}}$$

Proof. At first, since $\frac{\delta(t)}{t}$ is a decreasing function for $p \le 1/4$, we have that $e(t) \le (1-\frac{\alpha}{t}) + g(t)$, where $g(t) = \frac{\delta(t)}{t}$.

$$\begin{aligned}
e(t) &\le (1-\frac{\alpha}{t})e(t-1) + g(t)\\
&\le (1-\frac{\alpha}{t})(1-\frac{\alpha}{t-1})e(t-2) + (1-\frac{\alpha}{t})g(t-1) + g(t)\\
&\le (1-\frac{\alpha}{t})\cdots(1-\alpha)e(0) + \sum_{\tau=1}^{t} g(\tau)\prod_{i=\tau+1}^{t}(1-\frac{\alpha}{i})\\
&\le \frac{e(0)}{t^\alpha} + \sum_{\tau=1}^{t} g(\tau)e^{-\alpha\sum_{i=\tau+1}^{t}\frac{1}{i}}\\
&\le \frac{e(0)}{t^\alpha} + \sum_{\tau=1}^{t} g(\tau)e^{-\alpha(H_t-H_\tau)}\\
&\le \frac{e(0)}{t^\alpha} + e^{-\alpha H_t}\sum_{\tau=1}^{t} g(\tau)e^{\alpha H_\tau}\\
&\le \frac{e(0)}{t^\alpha} + \frac{O\left(\sqrt{\log(\frac{n}{p})}\right)}{t^\alpha}\sum_{\tau=1}^{t}\tau^\alpha\frac{\sqrt{\log\tau}}{\tau^{3/2}}
\end{aligned}$$

Lemma 9.

$$e(t) = \begin{cases} O\left(\sqrt{\log(\frac{n}{p})}\frac{(\log t)^{3/2}}{t^{\alpha}}\right) & \text{if } \alpha \leq 1/2 \\ O\left(\sqrt{\log(\frac{n}{p})}\frac{(\log t)^{3/2}}{t^{1/2}}\right) & \text{if } \alpha > 1/2 \end{cases}$$

Proof. We observe that

$$\sum_{\tau=1}^{t} \tau^{\alpha}\frac{\sqrt{\log\tau}}{\tau^{3/2}} \leq \int_{\tau=1}^{t} \tau^{\alpha}\frac{\sqrt{\log\tau}}{\tau^{3/2}}d\tau,$$

since $\tau^{\alpha}\frac{\sqrt{\log\tau}}{\tau^{3/2}}$

- If $\alpha \leq 1/2$ then

$$\int_{\tau=1}^{t} \tau^{\alpha}\frac{\sqrt{\log\tau}}{\tau^{3/2}}d\tau \leq \int_{\tau=1}^{t} \tau^{1/2}\frac{\sqrt{\log\tau}}{\tau^{3/2}}d\tau = O((\log t)^{3/2}))$$

- If $\alpha > 1/2$ then

$$\begin{aligned} \int_{\tau=1}^{t} \tau^{\alpha}\frac{\sqrt{\log\tau}}{\tau^{3/2}}d\tau &= \int_{\tau=1}^{t} \tau^{\alpha-1/2}\frac{\sqrt{\log\tau}}{\tau}d\tau \\ &= \frac{2}{3}\int_{\tau=1}^{t} \tau^{\alpha-1/2}((\log\tau)^{3/2})'d\tau \\ &= \frac{2}{3}(\log t)^{3/2} - (\alpha-1/2)\frac{2}{3}\int_{\tau=1}^{t} \tau^{\alpha-3/2}(\log\tau)^{3/2}d\tau \\ &= O\left((\log t)^{3/2}\right) \end{aligned}$$

5 Lower bound for unbiased estimators

We first state a well known result from point estimation:

Lemma 10 (Cramer-Rao bound). *Let $\hat{\theta}$ be an estimator for the parameter θ of a distribution P_θ, where θ is a continuous parameter. Suppose that the estimator is unbiased, that is:$E[\hat{\theta}] = 0$, for all distributions P_θ. Under suitable regularity conditions, which are met by bernoulli distributions, it holds:*

$$Var[\hat{\theta}] \geq \frac{1}{nI_\theta}$$

, where I_θ is the Fischer information of distribution P_θ.

In the case of Bernoulli random variables, we can easily show that for a Bernoulli with probability θ, $I_\theta = \frac{1}{\theta(1-\theta)}$. Applying the above result we obtain

$$var[\hat{\theta}] \geq \frac{\theta(1-\theta)}{n} \geq \frac{1}{4n}$$

for every unbiased estimator $\hat{\theta}$. We can use this fact to show that a specific class of distributed protocols that solve our problem has "large" variance. Let P be a protocol that when restricted to the star topology acts like an unbiased estimator for the weighted mean value of the neighbours, which can be an arbitrary real number in $[0,1]$. We will show that for all star topologies the variance of the solution is $\geq \frac{(1-a)^2}{4n}$ after n samples. Suppose that for a specific star topology the preceding claim doesn't hold. Then, we notice that if $\hat{\theta}$ is the output of the protocol is this topology, then $\frac{(\hat{\theta}-as)}{1-a}$ is an unbiased estimator of the mean value of the neighbours, and this holds for every possible value of the mean. By our hypothesis, we have:

$$Var[\frac{\hat{\theta}-as}{1-a}] = \frac{1}{(1-a)^2}Var[\hat{\theta}] < \frac{1}{4n}$$

, a contradiction, since our constructed estimator is unbiased.