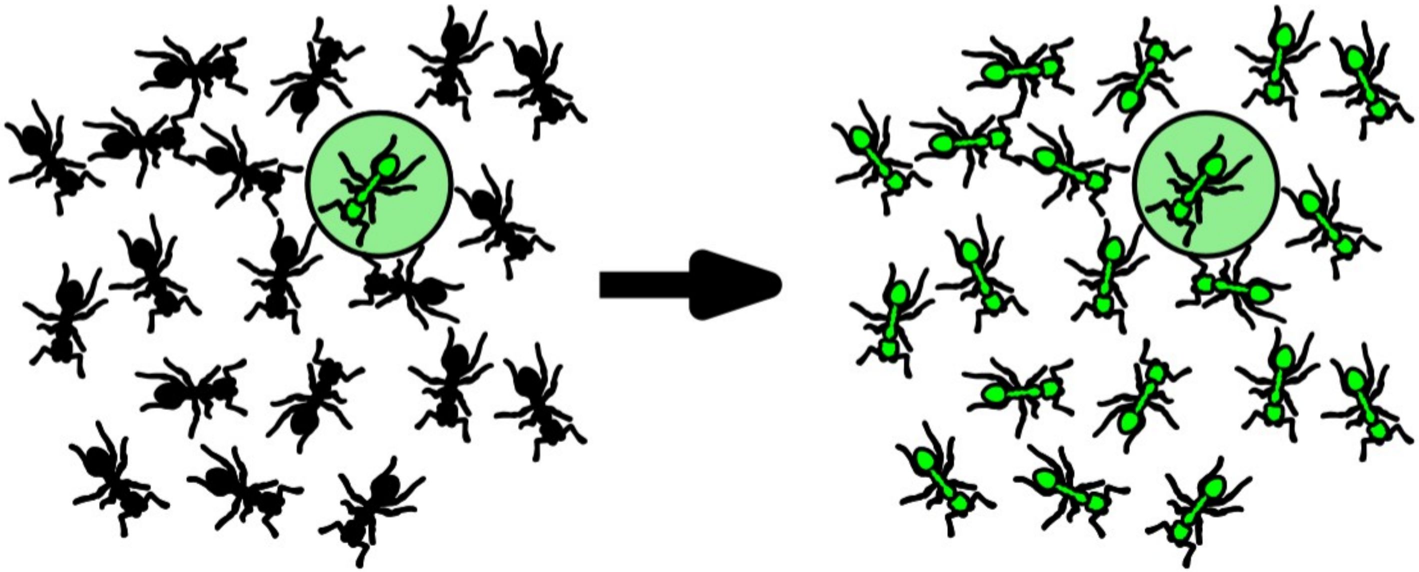
Natural Distributed Algorithms

- Lecture 2 -
Rumor Spreading in the Noisy PUSH Model

Emanuele Natale
CNRS - UCA

CdL in Informatica
Università degli Studi di Roma “Tor Vergata”

Rumor Spreading Problem

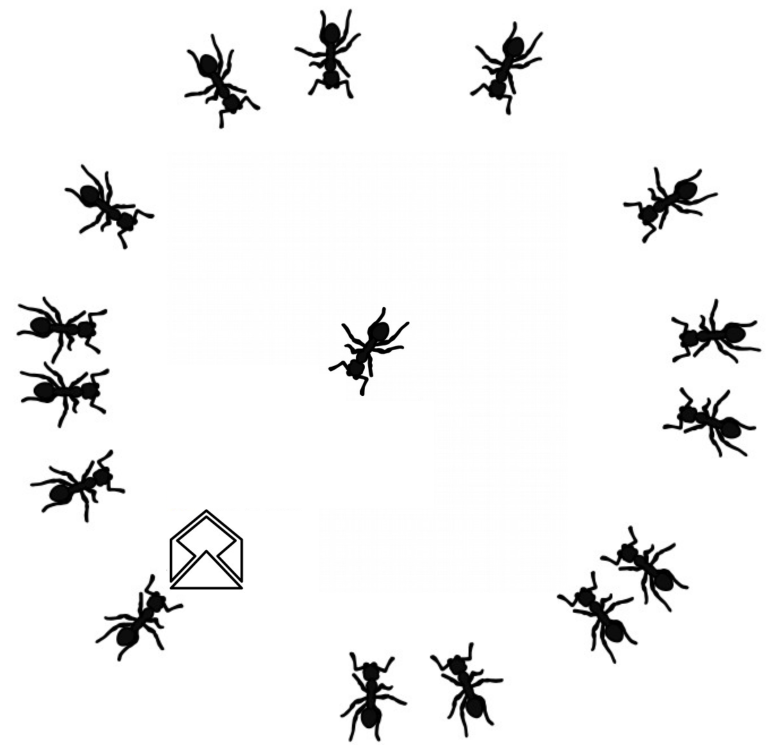


- One **source** node in a **special state**
- **Goal configuration**: all agents in the **special state**

Stochastic Interactions: PUSH Model

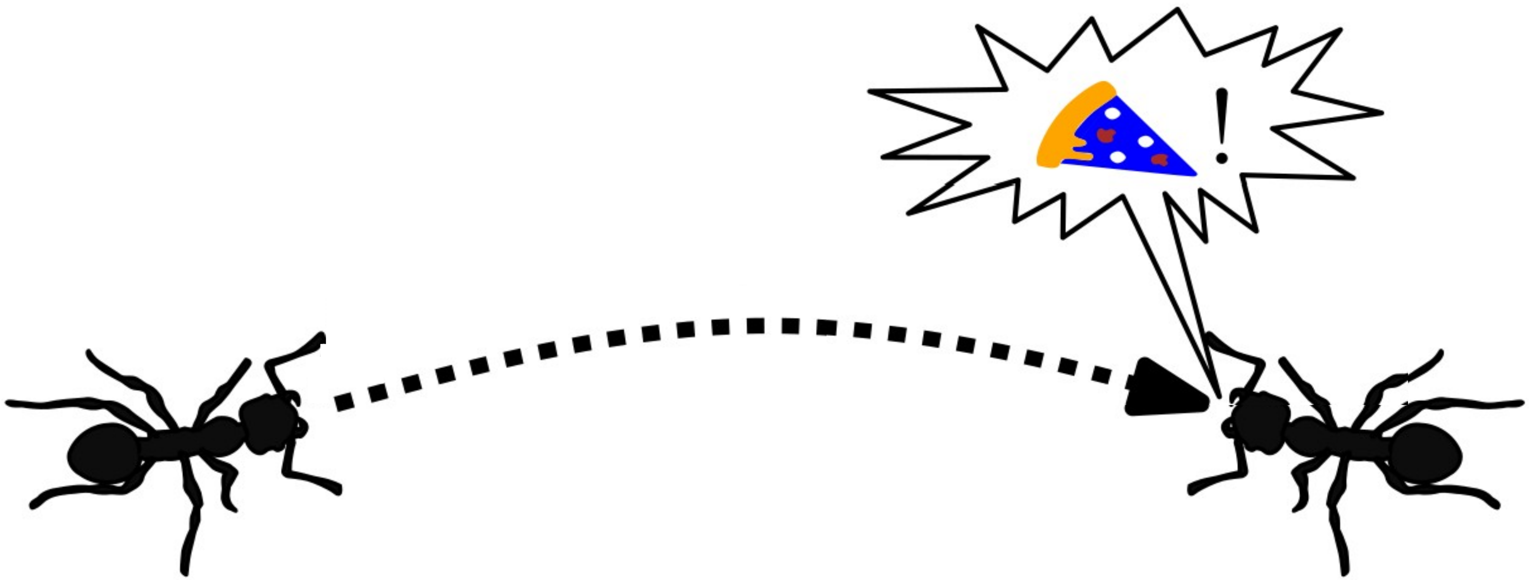
Desired features
- Stochastic
- Parsimonious (Anonymous)
- Active (uni-directional)

(Uniform) PUSH model
[Demers '88]
(single binary message)
- Discrete parallel time
- At each round each agent **can send one bit** message
- Each message is received by **one agent** chosen independently and uniformly at random



Noisy Communication

Before being received,
each bit is flipped with probability $\frac{1}{2} - \epsilon$



1 bit per message,
noise flips it with prob $\frac{1}{2} + \epsilon$

Rumor Spreading in the **Noisy** PUSH Model: "Chinese-Whispers" **Does Not Work**

RECALL

If each agent sends the **rumor** as soon as it receives it…

If $M^{(t)}$ is the received message, $p_t := \Pr\left(M^{(t)} = 1\right)$ and $q_t := 1 - p_t = \Pr\left(M^{(t)} = 0\right)$

$$\begin{pmatrix} p_t \\ q_t \end{pmatrix} = \begin{pmatrix} \frac{1}{2}+\epsilon & \frac{1}{2}-\epsilon \\ \frac{1}{2}-\epsilon & \frac{1}{2}+\epsilon \end{pmatrix} \begin{pmatrix} p_{t-1} \\ q_{t-1} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}+\epsilon & \frac{1}{2}-\epsilon \\ \frac{1}{2}-\epsilon & \frac{1}{2}+\epsilon \end{pmatrix}^t \begin{pmatrix} p_0 \\ q_0 \end{pmatrix}$$

$$\approx \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Breathe-Before-Speaking Algorithm

Feinerman, Ofer, Bernhard Haeupler, and Amos Korman. 2017. "Breathe before Speaking: Efficient Information Dissemination despite Noisy, Limited and Anonymous Communication." Distributed Computing 30 (5): 339–55. https://doi.org/10.1007/s00446-015-0249-4.

Idea. Reduce number of "message hops" from the source.

Algorithm "Breathe-Before-Speaking" (informal description)

<u>Stage 1</u>:

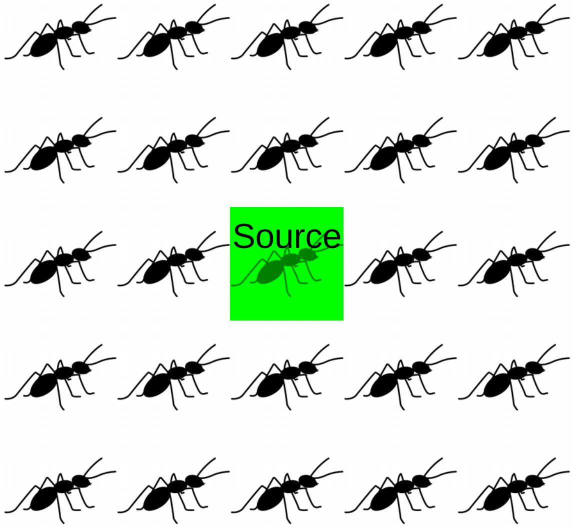
For $O(\log n)$ **phases**:

At each round of each phase:

- each informed agent **sends** its **opinion**
- each uninformed agent **only listens** to incoming messages

At the end of the phase each uninformed that has received a message becomes informed with the opinion that it has received

<u>Stage 2</u>: ……



Example with phases of 3 steps

Breathe-Before-Speaking Algorithm: Core Idea of Stage 1

If a phase lasts $k \geq 2$ rounds (each agent sends k messages)

$$\mathbb{E}\left[I^{(t)} \mid I^{(t-1)} = i\right] = i + (n-i)\left(1 - \left(1 - \frac{1}{n}\right)^{ik}\right)$$

using $1 - x \leq e^{-x}$

$$\geq i\left(1 + \left(\frac{n}{i} - 1\right)\left(1 - e^{-k\frac{i}{n}}\right)\right)$$

using $e^{-x} \geq 1 - \frac{x}{2}$ and assuming $i \leq \frac{n}{k}$

$$\geq i\left(1 + \left(\frac{n}{i} - 1\right)\frac{k}{2}\frac{i}{n}\right)$$

$$\geq i\left(1 + \left(1 - \frac{i}{n}\right)\frac{k}{2}\right)$$

$$\geq i\left(1 + \frac{k}{4}\right)$$

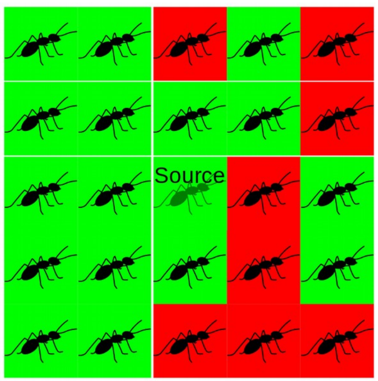
How large should *k* be?

From Stage 1 to Stage 2: How Many Rounds in a Phase?

RECALL

If $M^{(t)}$ is the received message,
$p_t := \Pr\left(M^{(t)} = 1\right)$ and
$q_t := 1 - p_t = \Pr\left(M^{(t)} = 0\right)$

$$\begin{pmatrix} p_t \\ q_t \end{pmatrix} = \begin{pmatrix} \frac{1}{2} + \epsilon & \frac{1}{2} - \epsilon \\ \frac{1}{2} - \epsilon & \frac{1}{2} + \epsilon \end{pmatrix} \begin{pmatrix} p_{t-1} \\ q_{t-1} \end{pmatrix}$$



We want the *bias* $\delta_t = |p_t - q_t|$ to be as large as possible

What bias is enough? $\rightarrow$ Typically $\Omega(\sqrt{n \log n})$

How are we going to *use* the bias?

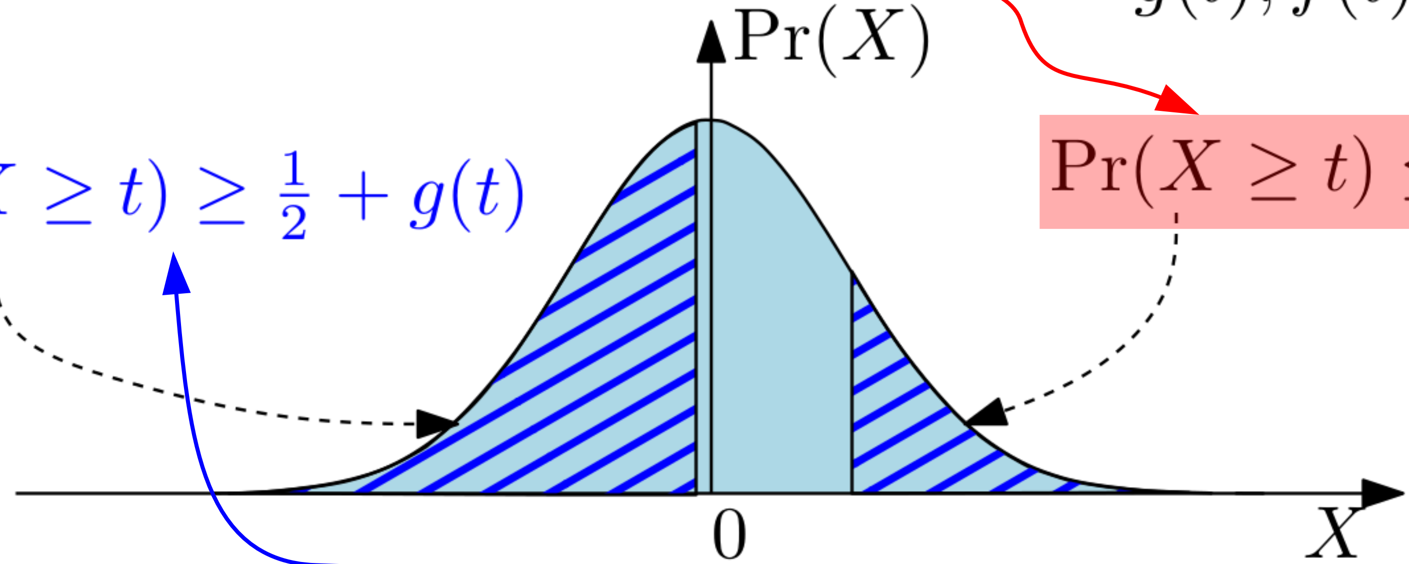
Algorithm "Breathe-Before-Speaking" (informal description)
Stage 1: …… (all agents got an opinion)
Stage 2: For $O(\log n)$ **phases**, at each round each agent **sends** its **opinion** and **counts** the opinions of received messages. At the end of the phase each agent takes **the most frequent** opinion.

Mathematical Challenge: Small Deviations Theory

There is a well-established **large deviations theory** concerned with showing that the probability that a r.v. exceeds a certain bound is very small

$$g(t), f(t) << 1$$

$$\Pr(X \geq t) \geq \frac{1}{2} + g(t)$$



$$\Pr(X \geq t) \leq f(t)$$

Warning: In Multi-Agent Systems (MAS), ensuring that a node doesn’t make an error can be computationally expensive (e.g. needs many messages).

Since in MAS we have many agents, if we can easily solve the **majority consensus problem**, it is enough that the **majority is right**.

Majority Amplification of Small Deviations

Idea of Stage 2. Use majority rule to amplify bias (reduce noise).

optimal rule

If $M^{(t)}$ is the received message,
$p_t := \Pr\left(M^{(t)} = 1\right)$ and
$q_t := 1 - p_t = \Pr\left(M^{(t)} = 0\right)$

$$\delta_t = |p_t - q_t|$$
$$= \left|\left(\tfrac{1}{2} + \epsilon\right) p_{t-1} + \left(\tfrac{1}{2} - \epsilon\right) q_{t-1} - \left(\left(\tfrac{1}{2} - \epsilon\right) p_{t-1} + \left(\tfrac{1}{2} + \epsilon\right) q_{t-1}\right)\right|$$

(Recall: bias decreases exponentially if we forward directly!)

$$= 2\epsilon\delta_{t-1}$$

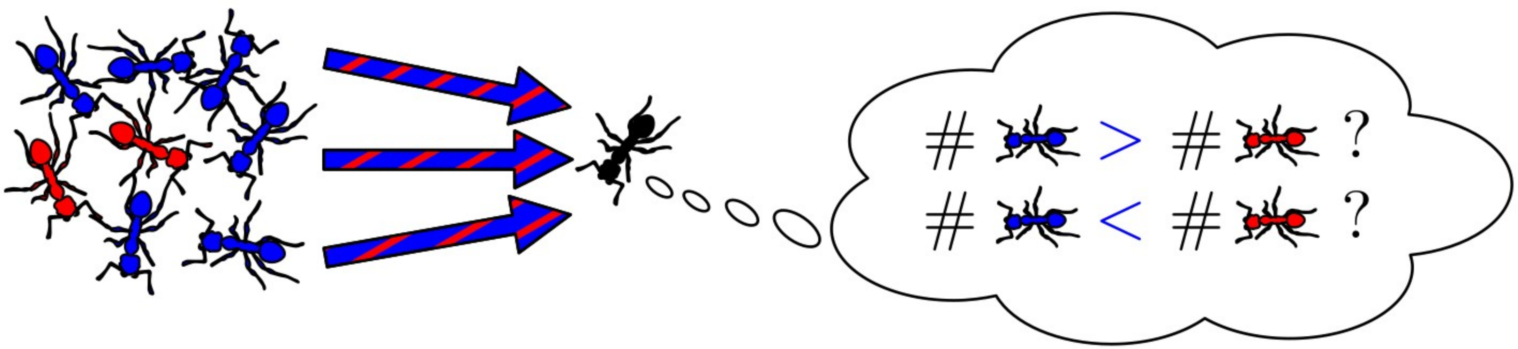
effective bias

How much does bias increases if we wait that we receive *k* messages and then send the majority? (*Breathe before speaking*!)

Attention! If we wait *k* rounds we can only say that the majority of nodes will receive at least *k/2* messages (Why? Chernoff bound...)

Majority Amplification Lemma (1/2)

Stage 2: For $O(\log n)$ **phases**, at each round each agent **sends** its **opinion** and **counts** the opinions of received messages.
At the end of the phase each agent takes **the most frequent** opinion.



Lemma

If $X_1, ..., X_n$ are i.i.d. binary r.v.s with bias $\lambda = |\Pr(X_i = 1) - \Pr(X_i = 0)|$ then the majority value m has bias

$$|\Pr(m = 1) - \Pr(m = 0)| \geq \min\left\{\sqrt{n}\lambda, \frac{1}{4}\right\}.$$

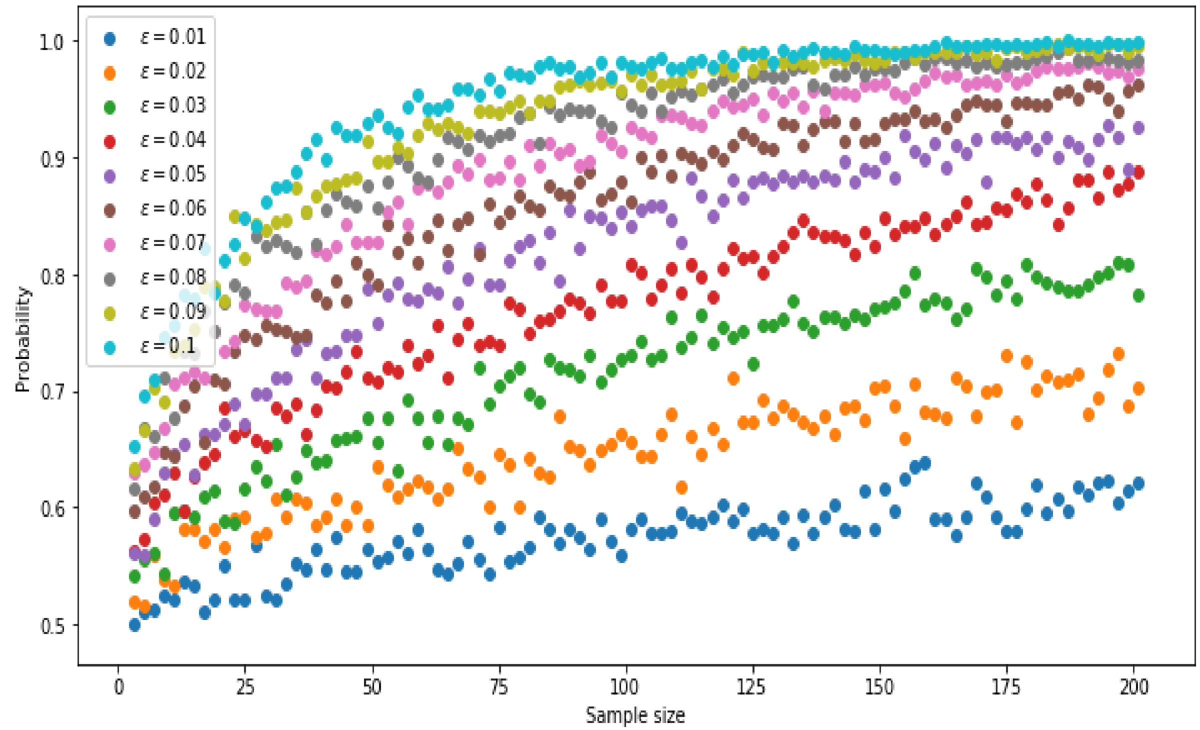
How is the Lemma used?

Application of Majority Amplification Lemma (2/2)

$$|\Pr(m=1) - \Pr(m=0)| \geq \min\left\{\sqrt{n}\lambda, \tfrac{1}{4}\right\}.$$

Let $X_1 = M^{(t)}, \ldots, X_n = M^{(t+n)}$ be the noisy messages received during a phase of Stage 2

The effective bias is $\lambda = 2\epsilon\delta$, where δ is the difference between the fraction of nodes with correct and wrong opinion



$$|\Pr(m=1) - \Pr(m=0)| \geq \min\left\{\sqrt{\text{phase length}} \cdot 2\epsilon\delta, \tfrac{1}{4}\right\}.$$

If phase length $> \frac{4}{\epsilon^2}$ then δ grows by a factor 2 at each phase.
Once the effective bias is $> \frac{1}{4}$, one phase of $O(\log n)$ is enough.
(Chernoff bounds)

Breathe-Before-Speaking Algorithm: Summary

Stage 1:

For $O(\log n)$ **phases**: $(t_i = O\left(\frac{1}{\epsilon^2}\right)$ except first and last phase)

At each of the t_i rounds of each phase:

- each informed agent **sends** its **opinion**
- each uninformed agent **only listens** to incoming messages

At the end of the phase each uninformed that has received a message becomes informed with the opinion that it has received

Stage 2:

For $O(\log n)$ **phases**: $(t_i = O\left(\frac{1}{\epsilon^2}\right)$ except first and last phase)

At each of the t_i rounds of each phase:

- each agent **sends** its **opinion**
- each agent **counts** the opinions of received messages.

At the end of the phase each agent takes **the most frequent** opinion.

(Half of the) Proof of Majority Amplification

If $X_1, \dots, X_n$ are i.i.d. binary r.v.s with n odd and bias $\lambda = |\Pr(X_i = 1) - \Pr(X_i = 0)|$ then the majority value m has bias

$$|\Pr(m=1) - \Pr(m=0)| \geq \min\left\{\sqrt{n}\lambda, \frac{1}{4}\right\}.$$

(Half-)Proof

Two cases: $n > \frac{1}{\lambda^2}$ and $n < \frac{1}{\lambda^2}$

First case: We use Chebyshev's inequality

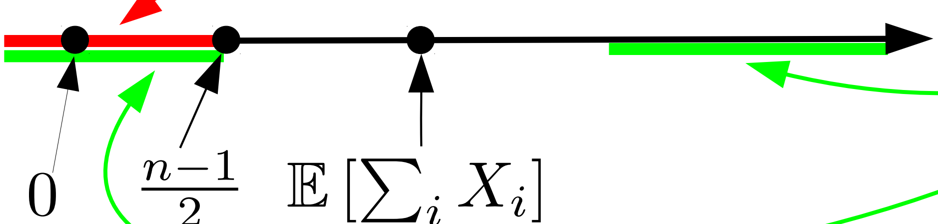
$$P(|Y - \mathbb{E}[Y]| \geq x) \leq \frac{\mathrm{Var[Y]}}{x^2}$$

$\mathbb{E}\left[\sum_i X_i\right] = n\left(\frac{1}{2} + \lambda\right)$ $\qquad$ $\mathrm{Var}\left[\sum_i X_i\right] = n\left(\frac{1}{4} - \lambda^2\right)$

(Correct opinion is w.l.o.g. 1)

$$P(m=0) = P\left(\sum_i X_i \leq \frac{n-1}{2}\right) \leq P\left(\left|\sum_i X_i - \mathbb{E}\left[\sum_i X_i\right]\right| > n\lambda - \frac{1}{2}\right)$$

$\leq \frac{1}{3}$



x is computed such that this is true

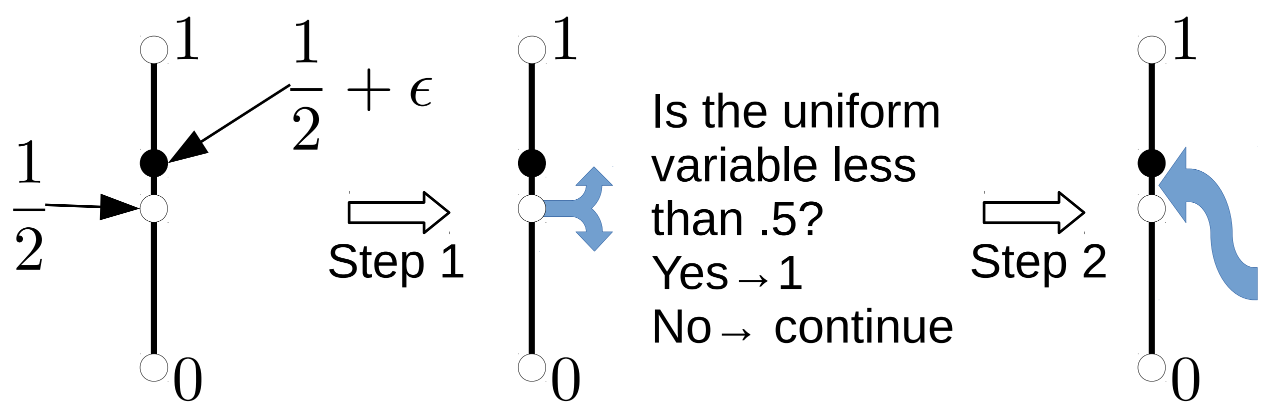
The $n < \frac{1}{\lambda^2}$ Case: Intuition

Observation:
We can simulate a coin C_p with probability *p* using a uniform random variable

```
[1]: from random import uniform

[11]: def coin(p):
          return uniform(0,1)<p

[12]: [coin(1/3) for i in range(10)]

[12]: [False, False, False, False, False,
      True, False, True, False, False]
```

(Correct opinion is w.l.o.g. 1)

We can simulate $C_{\frac{1}{2}+\epsilon}$ as follows:



Given that the uniform variable is more than .5, is it ***here***?
Yes → 1
No → 0

Step 1: Flip $C_{\frac{1}{2}}$

Step 2: Flip coin with prob.
$P(U \in (\frac{1}{2}, \frac{1}{2}+\epsilon) \,|\, U \geq \frac{1}{2})$
$= \frac{P(U \in (\frac{1}{2}, \frac{1}{2}+\epsilon))}{P(U \geq \frac{1}{2})} = 2\epsilon$

$$C_{\frac{1}{2}+\epsilon} = C_{\frac{1}{2}} + (1 - C_{\frac{1}{2}})C_{2\epsilon}$$

Throw n variables $C_{\frac{1}{2}}$ and n corresponding variables $C_{2\epsilon}$. If the majority of $C_{\frac{1}{2}}$s is 1, ok, otherwise, if we miss z 1s for the majority to be 1, we look at the prob that at least z $C_{2\epsilon}$s are 1.

...Project Idea

As for the case $n < \frac{1}{\lambda^2}$,
there are two proofs, both accessible but with lots of calculations:

- Feinerman, Ofer, Bernhard Haeupler, and Amos Korman. 2017. “Breathe before Speaking: Efficient Information Dissemination despite Noisy, Limited and Anonymous Communication.” Distributed Computing 30 (5): 339–55. https://doi.org/10.1007/s00446-015-0249-4.
- Fraigniaud, Pierre, and Emanuele Natale. 2019. “Noisy Rumor Spreading and Plurality Consensus.” Distributed Computing 32 (4): 257–76. https://doi.org/10.1007/s00446-018-0335-5.

Project: (First, understand and then) write a detailed exposition of either one of the proofs.

More Project Ideas

- Write the analysis **in expectation** of Phase 1 of the Breathe-Before-Speaking algorithm.
  The reasoning is similar to what we have done for the “Chinese-whispers protocol” in order to solve the non-noisy rumor spreading problem. You can also check the analysis w.h.p. (much, much harder than in expectation!), in

  - Feinerman, Ofer, Bernhard Haeupler, and Amos Korman. 2017. “Breathe before Speaking: Efficient Information Dissemination despite Noisy, Limited and Anonymous Communication.” Distributed Computing 30 (5): 339–55. https://doi.org/10.1007/s00446-015-0249-4.
  - Fraigniaud, Pierre, and Emanuele Natale. 2019. “Noisy Rumor Spreading and Plurality Consensus.” Distributed Computing 32 (4): 257–76. https://doi.org/10.1007/s00446-018-0335-5.

- Simulate the Breathe-Before-Speaking algorithm on Erdős-Rényi graphs (random graphs where each edge is added with probability p), varying p, the noise parameter ϵ and the duration of the time-windows.
  Simulations should be performed using open-source software with some effort to make them efficient (e.g. coded in Python using Numpy), and the source code should be made publicly available (e.g. on Gitlab) and GPL licensed.

Likelihood Ratio Test (Neyman-Pearson Lemma)

$X_1, ..., X_n$ i.i.d. r.v.s, all with distribution either P_0 or P_1.

$\psi(\mathbf{x}) \in \{0,1\}$ with $\mathbf{x} = (x_1, ..., x_n)$ is any function that guesses whether the distribution is P_0 (when $\psi = 0$) or P_1 (when $\psi = 1$).

$$P_0(\psi = 1) + P_1(\psi = 0)$$

$$= P_0(\psi = 1) + 1 - P_1(\psi = 1) = 1 + \sum_{\substack{\mathbf{x} \text{ s.t.} \\ \psi(\mathbf{x})=1}} \left(P_0(\mathbf{x}) - P_1(\mathbf{x})\right)$$

$$= 1 + \sum_{\substack{\mathbf{x} \text{ s.t.} \\ \psi(\mathbf{x})=1 \\ P_0(\mathbf{x}) \geq P_1(\mathbf{x})}} \left(P_0(\mathbf{x}) - P_1(\mathbf{x})\right) + \sum_{\substack{\mathbf{x} \text{ s.t.} \\ \psi(\mathbf{x})=1 \\ P_0(\mathbf{x}) < P_1(\mathbf{x})}} \left(P_0(\mathbf{x}) - P_1(\mathbf{x})\right)$$

If we define $\psi = \mathbf{1}\left[P_0(\mathbf{x}) < P_1(\mathbf{x})\right]$, error is minimized

Back to Majority Amplification

Chebyshev's Inequality

Recall **Markov's inequality**: Given positive random variable X

$$\mathbb{E}\left[X\right] = \sum_{i=1}^{\infty} i \Pr\left(X = i\right) \ \geq \sum_{i=t}^{\infty} i \Pr\left(X = i\right)$$

$$\geq \sum_{i=t}^{\infty} t \Pr\left(X = i\right) \ = t \Pr\left(X \geq t\right)$$

To prove **Chebyshev's inequality**, given a r.v. Y, set

$$X = (Y - \mathbf{E}[Y])^2$$

...and apply Markov's inequality

$$P\left(|Y - \mathbb{E}\left[Y\right]| \geq t\right) =$$

$$P\left(\left(Y - \mathbb{E}\left[Y\right]\right)^2 \geq t^2\right) \leq \frac{\mathbb{E}\left[\left(Y - \mathbb{E}\left[Y\right]\right)^2\right]}{t^2} = \frac{\mathrm{Var}\left[\mathrm{Y}\right]}{t^2}$$

Back to Proof of Lemma

Experiments on Majority Amplification

```
In [1]: import numpy as np
```

```
In [2]: x = [2*i+1 for i in range(1,101)]
```

```
In [3]: eps = [0.01*i for i in range(1,11)]
```

```
In [4]: bias = [2*epsilon for epsilon in eps]
```

```
In [5]: trials = 1000
        result = np.empty([len(eps), len(x)])
        for i in range(len(eps)):
            prob = 1/2+eps[i]
            for j in range(len(x)):
                experiments = np.random.binomial(x[j], prob, trials)
                successes = list(map(lambda y: y>x[j]/2, experiments))
                result[i][j] = sum(successes)/trials
```

```
In [6]: import matplotlib.pylab as plt
```

```
In [7]: plt.figure(figsize=(12,6))
        for i in range(len(eps)):
            plt.scatter(x, result[i][:], label='$\epsilon=$'+str(eps[i]))
        plt.xlabel('Sample size')
        plt.ylabel('Probability')
        plt.legend()
```

Jupyter-lab notebook (included in the Anaconda Python distribution)

Back to the Lemma Application

Computing t

$$\mathbb{E}\left[\sum_i X_i\right] = (2\gamma+1)\left(\frac{1}{2}+\delta\right) = \gamma\left(1+\delta+\frac{1}{\gamma}\left(\frac{1}{2}+\delta\right)\right)$$

$$\begin{aligned} t &= \gamma\left(1+\delta+\frac{1}{\gamma}\left(\frac{1}{2}+\delta\right)\right) - \gamma \\ &= \gamma\left(\delta+\frac{1}{2\gamma}(1+2\delta)\right) < 2\gamma\delta(1+\delta) \end{aligned}$$

$$\begin{aligned} P\left(\left|\sum_i X_i - \mathbb{E}\left[\sum_i X_i\right]\right| > 2\gamma\delta(1+\delta)\right) &< \frac{\gamma(1+\delta)}{2\left(2\gamma\delta(1+\delta)\right)^2} \\ &= \frac{1}{8\gamma}\frac{1}{\delta^2(1+\delta)} < \frac{1}{8\gamma\delta^2} < \frac{1}{3} \end{aligned}$$

Back to the Proof Intuition

Requirement for Concentration via CB

Let X_i be independent r.v.s with $|X_i| \leq M$, then

$$\Pr\left(\left|\sum_i X_i - \mathbb{E}\left[\sum_i X\right]\right| \geq \Delta\right) \leq 2e^{-\frac{\Delta^2}{2\left(\sum_i \mathbb{E}[X_i^2] + \frac{M}{3}\Delta\right)}}$$

Since we want "w.h.p." we need that on the r.h.s. $e^{\text{quantity}} \leq \frac{1}{n}$, that is

$$\text{quantity} = -\frac{\Delta^2}{2\mathbb{E}\left[\sum_i X_i\right] + \frac{2}{3}\Delta} \leq \log\frac{1}{n} = -\log n$$

that is $3\Delta^2 - 2\log n\Delta - 6\log n\mathbb{E}\left[\sum_i X_i\right] \geq 0$, which requires that

$$\Delta \geq \frac{1}{3}\log n\left(1 + \sqrt{1 + 18\frac{\mathbb{E}\left[\sum_i X_i\right]}{\log n}}\right)$$

Note that we need at least $\Delta \leq \mathbb{E}[\sum_i X_i]$.