Consensus

or

approximate majority
quantile summaries
selection problem
…

Milan Vojnovic
Microsoft Research

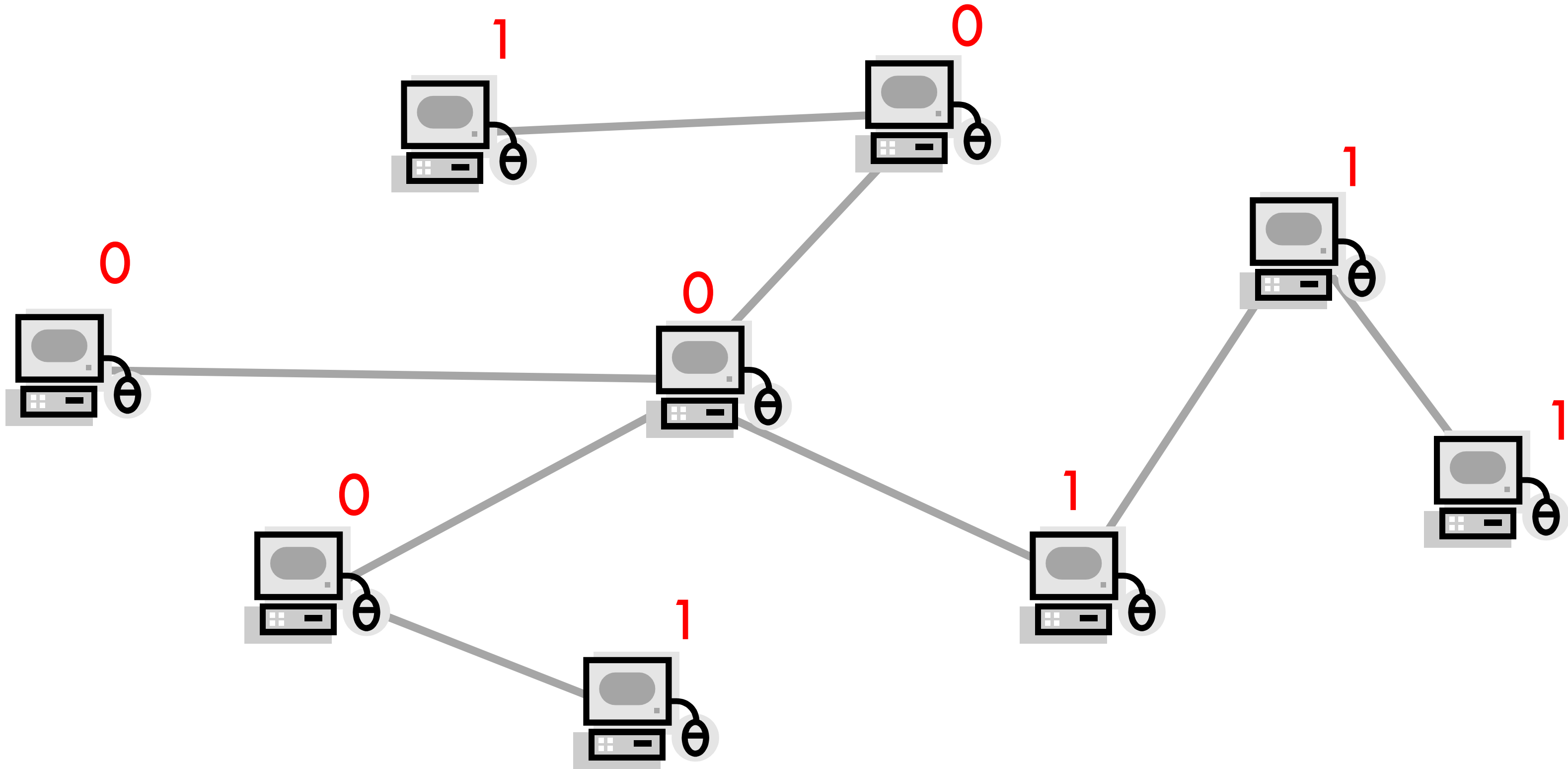
Workshop on Performance and Control of Large-Scale Networks
Eindhoven, Netherlands, June 30-July 2, 2014

A retrospective talk …



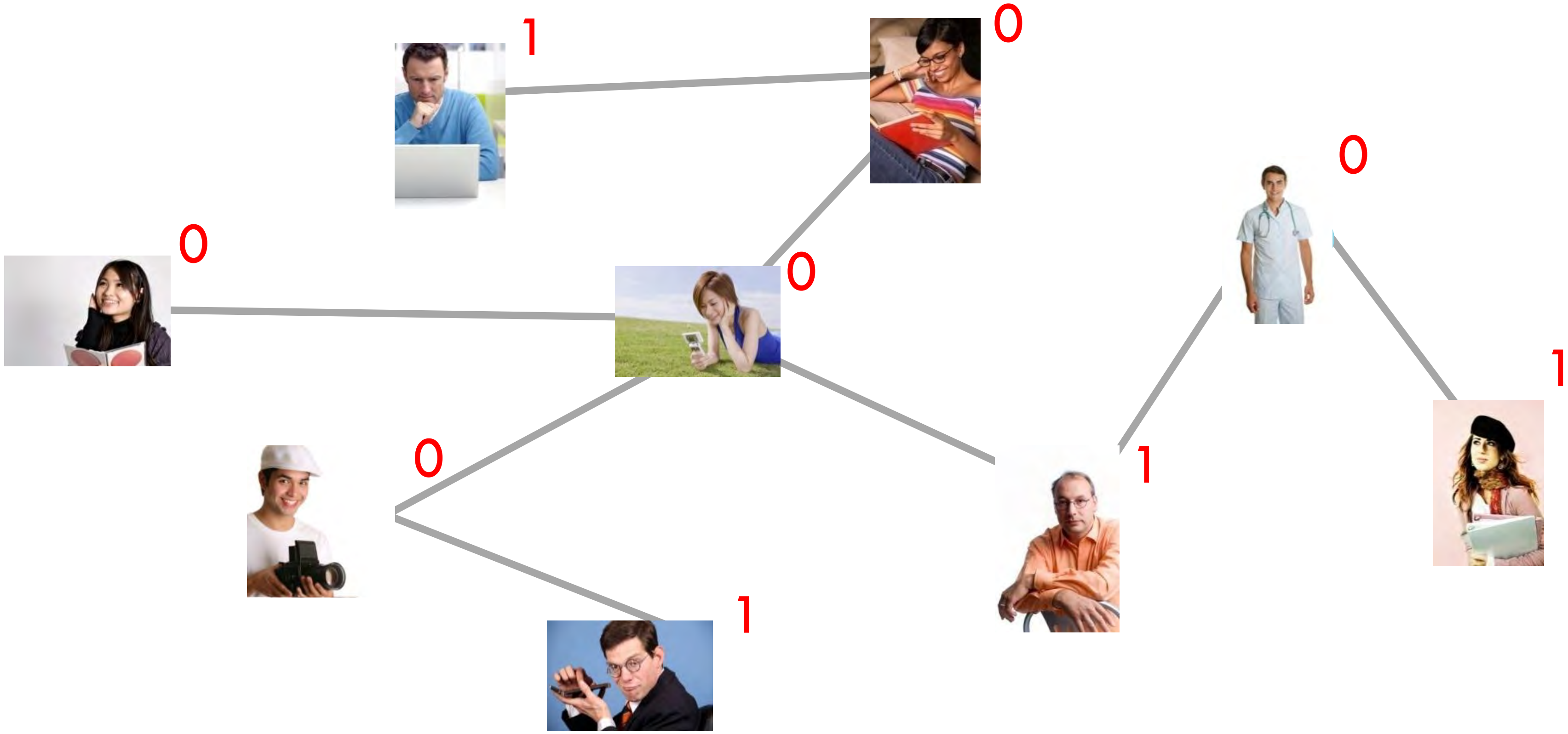
…

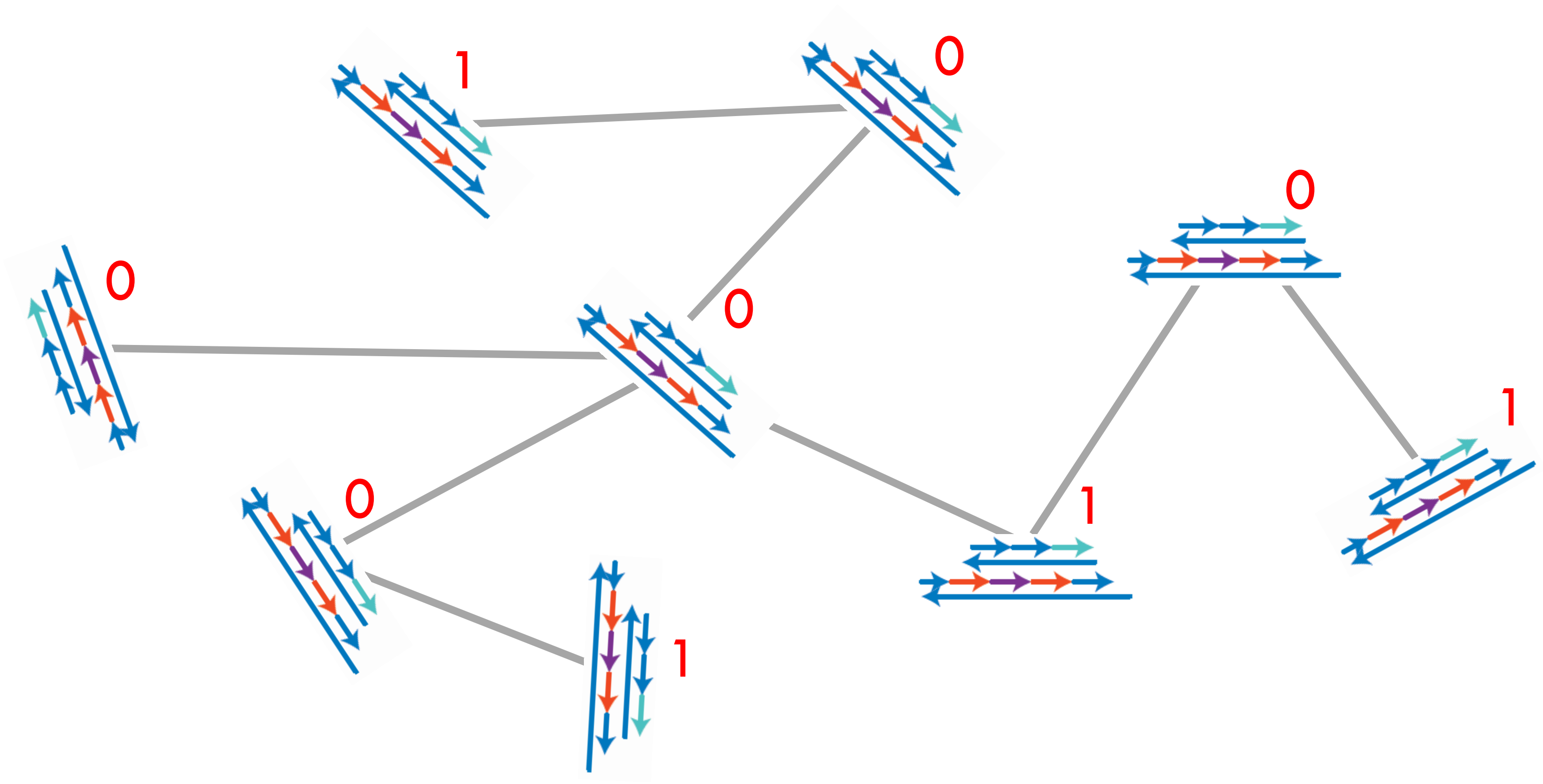
Approximate majority



Input: each node holds a binary value, either 0 or 1
Output: each node to report the majority vote (with high probability)

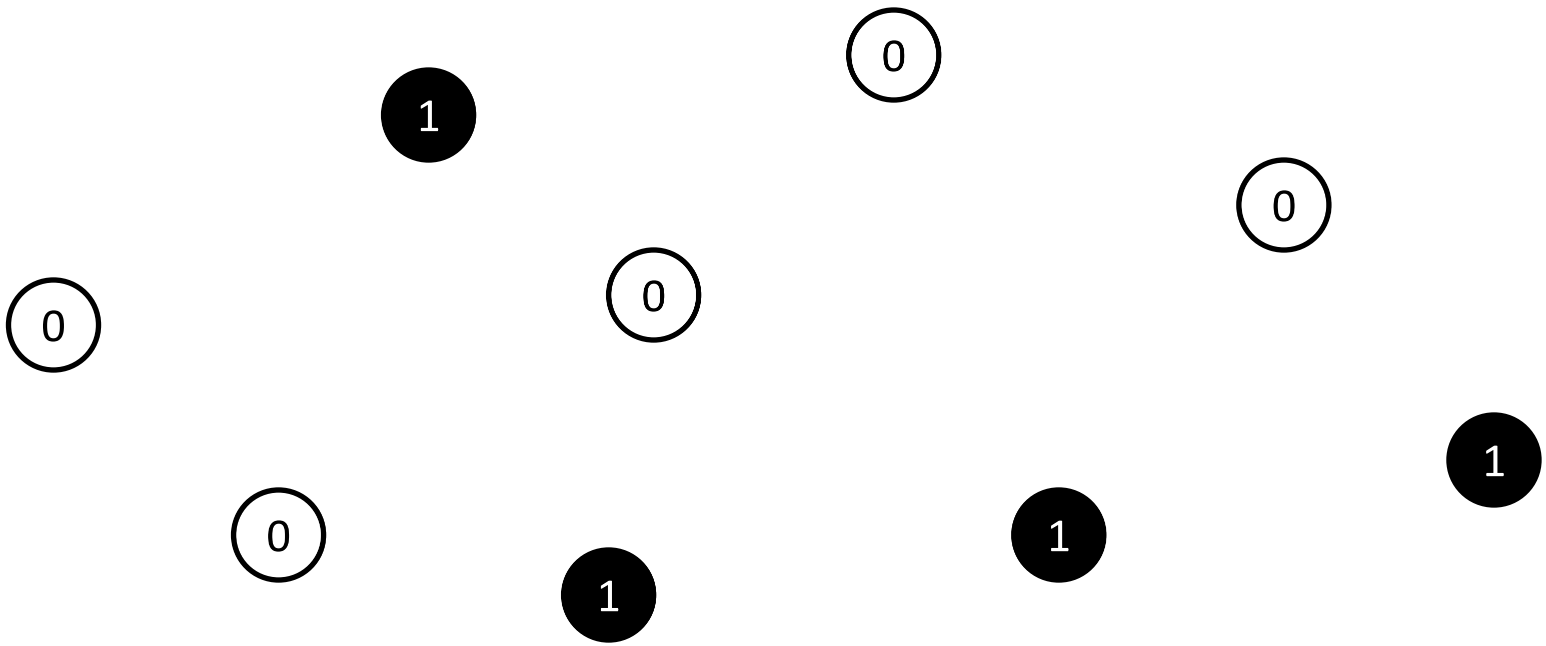
Requirement: limited memory per node and pairwise communication between nodes


1
0
0
0
0
0
1
1
1



1
0
0
0
0
0
1
1
1
0

Our notation



Approximate majority algorithms

2 states

- States: 0, 1
- Convergence time

  = $\Omega(n)$
- Probability of error

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

3 states

- States: 0, e, 1
- Convergence time

  = $O(\log n + \log 1/\epsilon)$
- Probability of error

  = $e^{-D(\frac{1+\epsilon}{2}||\frac{1}{2})n}$

4 states

- States: 0, e^0, e^1, 1
- Convergence time

  = $O(\frac{1}{\epsilon}\log n)$
- Probability of error

  = 0

n = number of nodes, ϵ = voting margin

Questions of interest

- **Correctness:** probability that each node identifies the initial majority state?
- **Convergence time:** time to reach consensus?
- Dependence on the number of nodes $\boldsymbol{n}$, voting margin ϵ, network structure?

Desiderata

- ☐ Reach correct consensus – initial majority
- ☐ Fast convergence
- ☐ Small communication overhead
- ☐ Small processing per node
- ☐ Decentralized

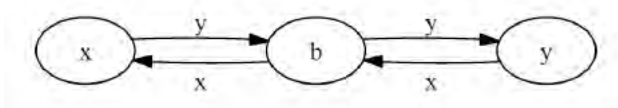
Outline

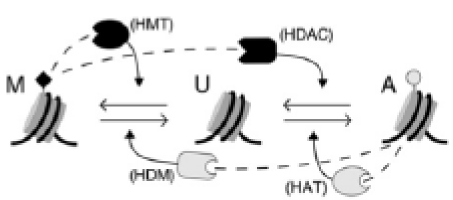
- Related work
- 3-state algorithm
- 4-state algorithm
- Conclusion

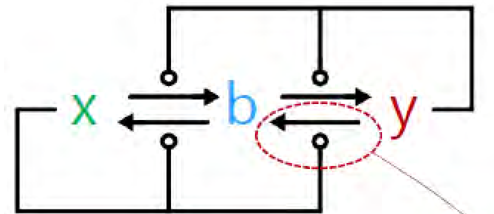
Some related work

2001	**Distributed Probabilistic Polling and Applications to Proportionate Agreement** Hassin and Peleg, Information and Computation	Probability of error for classical voter model
2004	**Computation in networks of passively mobile finite-state sensors** Angluin, Aspnes, Diamadi, Fischer, Peralta, PODC	6 state, exact majority
2007	**A Simple Population Protocol for Fast Robust Approximate Majority** Angluin, Aspnes, Eisenstat, DISC	3-state, approximate majority, discrete time, complete graph, correctness
2007	**Theoretical Analysis of Epigentic Cell Memory by Nucleosome Modification** Dodd, Micheelsen, Sneppen, Thon, Cell	3-state, approximate majority, discrete time
2009	**Artificial Biochemistry** Cardelli, Algorithmic Bioprocesses, Springer	3-state, approximate majority, continuous time
2009	**Using Three States for Binary Consensus on Complete Graphs** Perron, Vasudevan, V., IEEE Infocom	3-state, approximate majority, continuous time, complete graph, exact probability of error, convergence time
2009	**Interval Consensus: From Quantized Gossip to Voting** Benzit, Thiran and VetterliI, IEEE ASSP	Distributed m-ary hypothesis, exact computation
2010	**Convergence Speed of Binary Interval Consensus** Draief and V., IEEE Infocom (SIAM J. Control Optim., 2012)	4-state, exact majority, continuous time, arbitrary graph, convergence time
2012	**The Cell Cycle Switch Computes Approximate Majority** Cardelli, Csikasz-Nagy, Scientific Reports	Equivalence to approximate majority under certain conditions
2014	**Determining Majority in Networks with Local Interactions and very Small Local Memory** Mertzios, Nikoletseas, Raptopoulos, Spirakis, ICALP	Necessity of 4 states for exact computation

More references in this slide deck





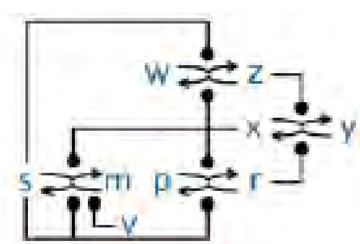


$$x + y \to^{r} y + b$$

$$y + x \to^{r} x + b$$

$$b + x \to^{r} x + x$$

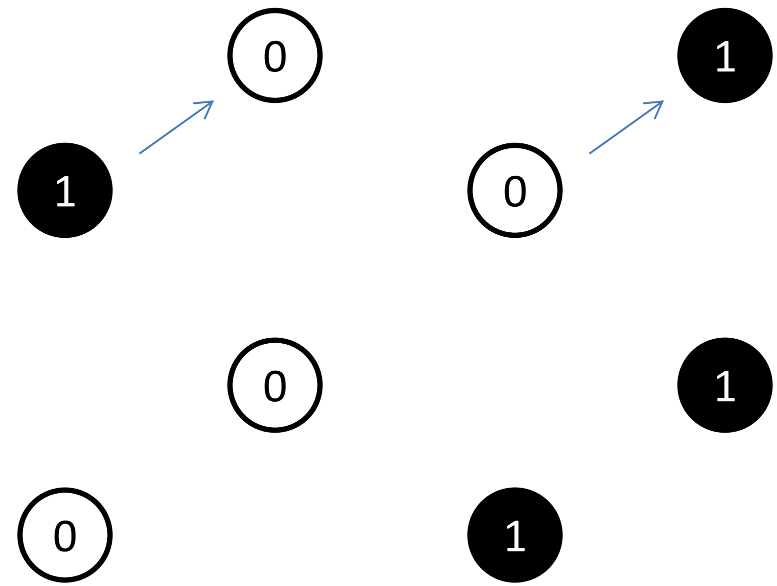
$$b + y \to^{r} y + y$$



Classical voter model

[Hassin-Peleg-01]

- Node takes over the state of the contacted node
- Binary state per node & binary signaling



- 0 initially held by v nodes, 1 initially held by u nodes
- Complete graph node interactions

Probability of incorrect consensus $f_{u,v} = \frac{u}{u+v}$

Statistical tests with limited memory

[Information Theory 70's]

$0 = a_0 \quad a_1 \quad \cdots \quad a_m = 1$

$H_i: \ \theta \in [a_i, a_{i+1}), \ i = 0, \dots, m-1$

000110111110100011

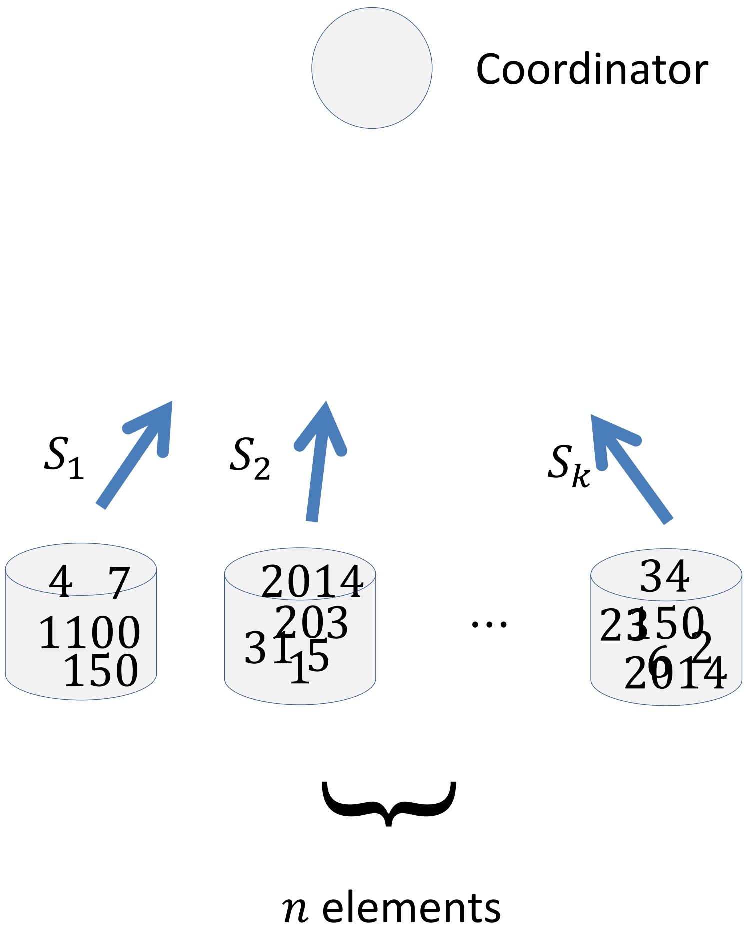
i. i. d. mean θ

- How many states S needs to identify the correct hypothesis with probability $1 - o(1)$ with the number of observations?

- *m+1 necessary and sufficient* [Koplowitz, IEEE Trans IT '75]

Quantile summaries

[Greenwald- Knanna-2004]



- Approximate quantile computation:

  Input:

  rank r

  rel. acc. par. $\epsilon > 0$

  Output:

  element of rank

  $\hat{r} \in [r - \epsilon n, r + \epsilon n]$

- Quantile summaries: max number of data elements communicated by any node

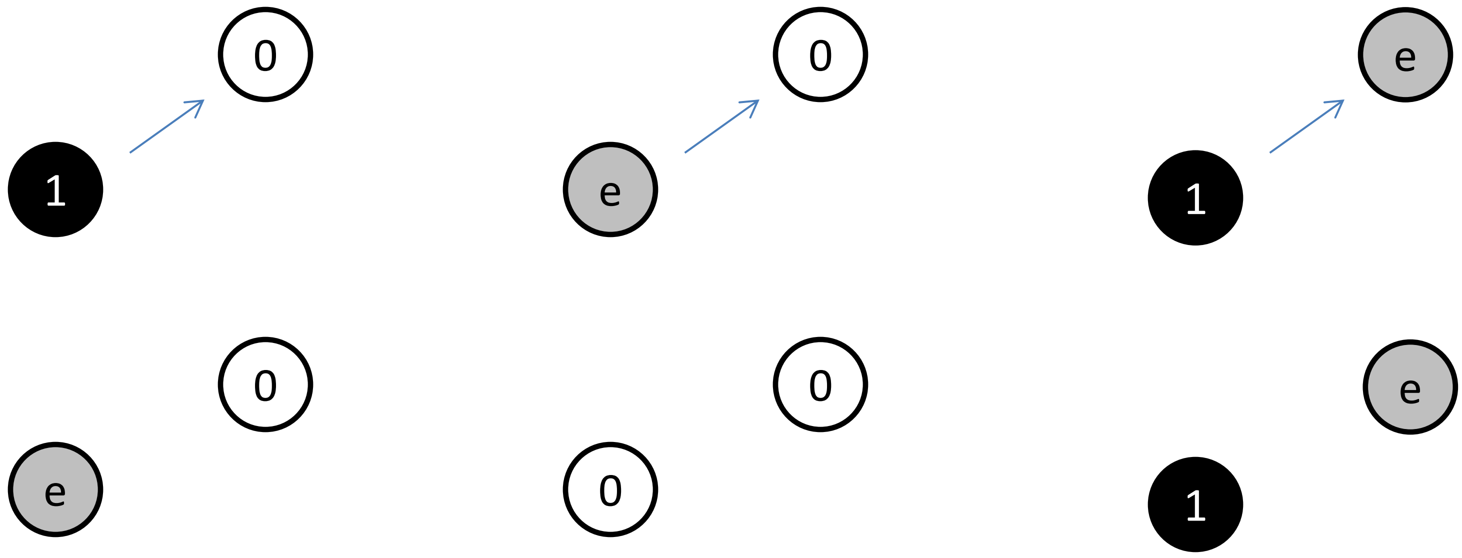
$$\max_{1 \le j \le k} |S_j| = O\left(\frac{1}{\epsilon} \log^2 n\right)$$

Outline

- Related work
- 3-state algorithm
- 4-state algorithm
- Conclusion

3-state algorithm

- Both *processing* and *signaling* take one of three states
  - 0 or 1 or e
  - e = “indecisive” state



Assumptions

- Interactions: asynchronous continuous-time, complete graph

  Each node samples another node uniformly at random at instances of a Poisson process with intensity 1

3-state algorithm: state evolution

- U = number of nodes in state 0
- V = number of nodes in state 1
- n = total number of nodes

- (U, V) Markov process:

$$(U,V) \to \begin{cases} (U+1, V) & \text{with rate} \quad (n-U-V)\dfrac{U}{n} \\ (U-1, V) & \text{with rate} \quad U\dfrac{V}{n} \\ (U, V+1) & \text{with rate} \quad (n-U-V)\dfrac{V}{n} \\ (U, V-1) & \text{with rate} \quad V\dfrac{U}{n} \end{cases}$$

Ternary protocol: probability of error

- (u, v) = initial point, $v > u$

- Theorem – probability of error:

$$f_{u,v} = \frac{1}{2}\sum_{j=1}^{u} \frac{a_{u,v}(j)}{2^{(u-j)+(v-j)}}$$

$$a_{u,v}(j) = \frac{v-u}{(u-j)+(v-j)}\binom{(u-j)+(v-j)}{(u-j)}$$

Probability of error (cont'd)

- Corollary – For initial state (u_n, v_n) such that $\left(\frac{u_n}{n}, \frac{v_n}{n}\right) \rightarrow \left(\frac{1-\epsilon}{2}, \frac{1+\epsilon}{2}\right)$, for $\epsilon \in (0,1]$, we have

$$\frac{1}{n}\log\left(f_{u_n,v_n}\right) \sim -D\left(\frac{1+\epsilon}{2} \,||\, \frac{1}{2}\right), \text{ large } n$$

- Exponentially decreasing in n
- Correctness with high probability if $\epsilon = \Omega\left(\sqrt{\frac{\log n}{n}}\right)$

Proof main ideas

- First-step analysis:

$$(wu + wv + 2uv)f_{u,v} = wuf_{u+1,v} + wvf_{u,v+1} + uvf_{u-1,v} + uvf_{u,v-1}$$

where $w = n - u - v$

with the boundary conditions:

$f_{0,v} = 0$, for $v > 0$
$f_{u,0} = 1$, for $u \geq 0$

Proof main ideas (cont'd)

- Lemma – $f_{u,v}$ solution of

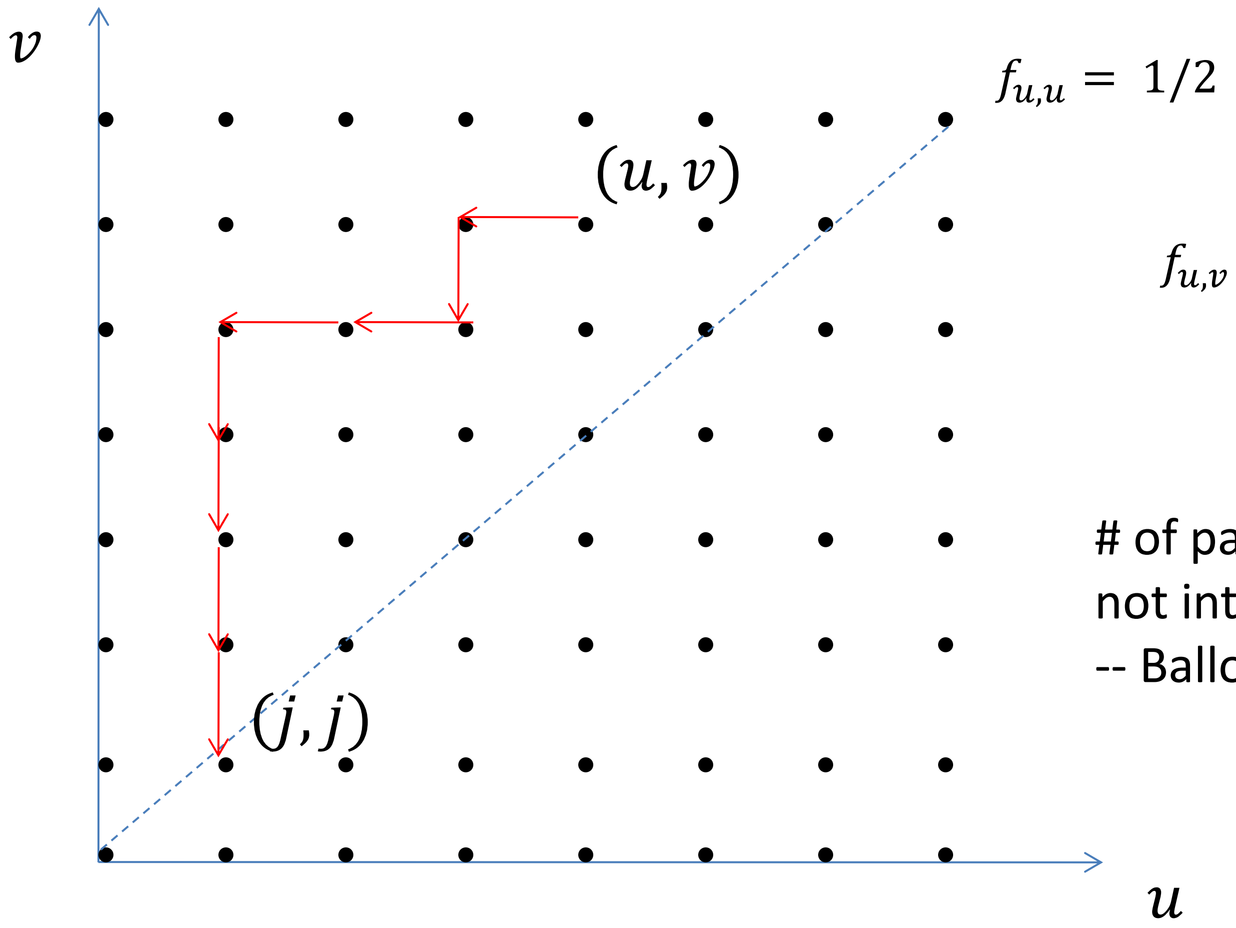
$$f_{u,v} = \frac{1}{2} f_{u,v-1} + \frac{1}{2} f_{u-1,v}$$

with the boundary conditions: $f_{0,v} = 0$ for $v > 0$, $f_{u,0} = 1$, for $u \geq 0$

i.e. $f_{u,v}$ is the error probability for

$$(u,v) \rightarrow \begin{cases} (u, v-1) & \frac{1}{2} 1_{v>0} \\ (u-1, v) & \frac{1}{2} 1_{u>0} \end{cases}$$

Proof main ideas (cont'd)



$$f_{u,v} = \sum_{j=1}^{u} \frac{1}{2^{(u-j)+(v-j)}} n_j$$

# of paths from (u, v) to (j, j) not intersecting $u = v$
-- Ballot theorem

Convergence time

- The limit ODE

$$\frac{d}{dt}u(t) = u(t)\big(1 - u(t) - 2v(t)\big)$$

$$\frac{d}{dt}v(t) = v(t)\big(1 - u(t) - 2u(t)\big)$$

- Def: $t_{n,\epsilon}$ = smallest time t such that $1 - v(t)$ and $u(t)$ are of order $1/n$ given that $v(0) = (1+\epsilon)/2$ and $u(0) = (1-\epsilon)/2$

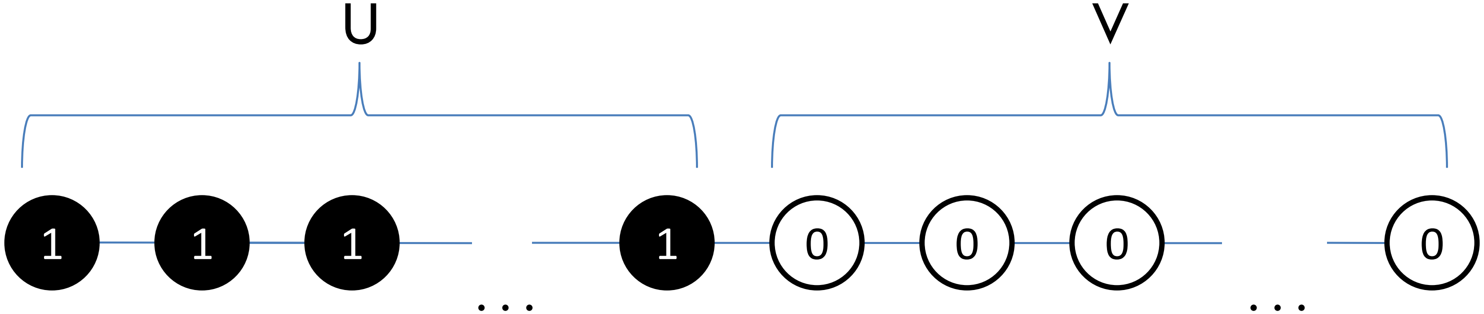
$$t_{n,\epsilon} = O\left(\log(n) + \log\left(\frac{1}{\epsilon}\right)\right)$$

Proof: $t = \log\left(\frac{\big(v(t)-u(t)\big)^3}{u(t)v(t)}\right) - \log\left(\frac{\big(v(0)-u(0)\big)^3}{u(0)v(0)}\right)$

Convergence time lower bound

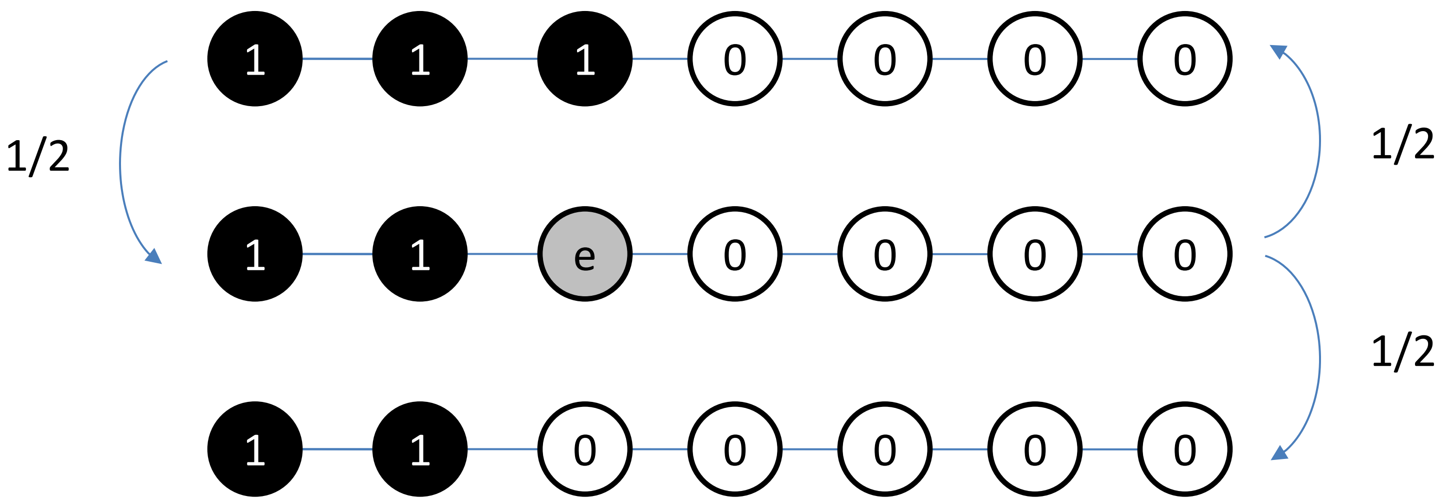
- Lower bound: $\Omega(n^2)$
- Example: path

  reduction to classical voter model



Convergence time lower bound (cont’d)

- Ternary protocol on a path corresponds to a classical voter model dynamics



Extension to plurality problem

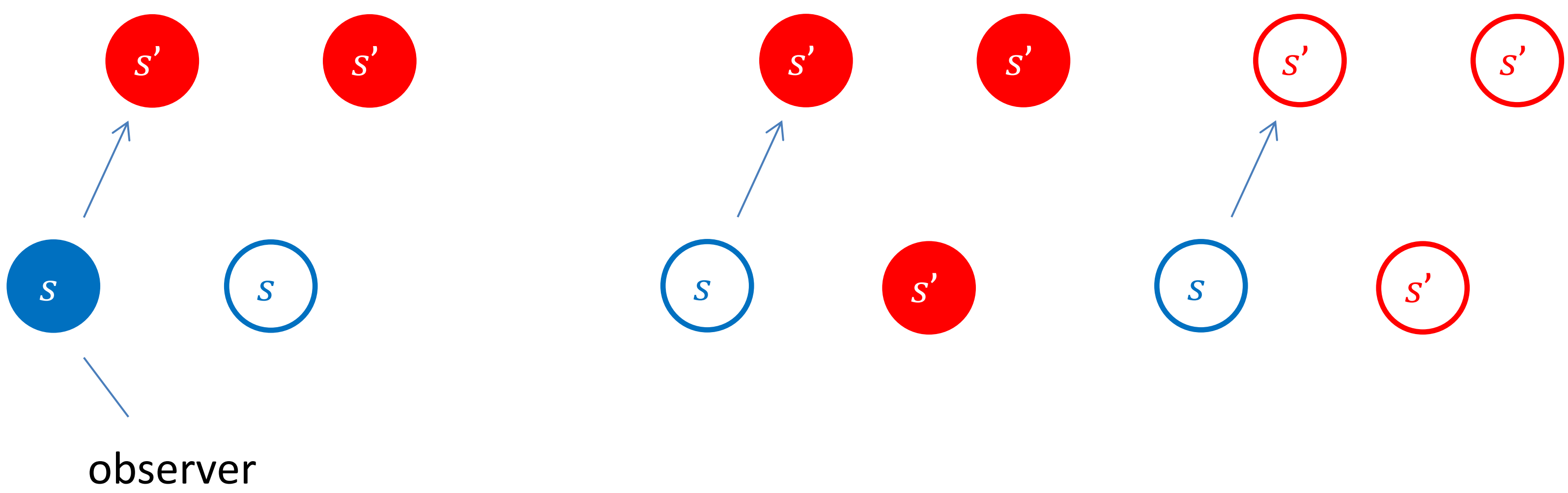
[Jung-Kim-V.-2012]

- $m \geq 2$ alternatives
  - Binary consensus as special case: $m = 2$

- Output: each node to correctly identify a state that is initially a plurality winner

Plurality algorithm

- m alternatives 1 2 … m
- 2m states: weak s strong s



State evolution

- (S, W) Markov process:

$$(S,W) \to \begin{cases} (S,W) + (-e_i, e_i) & \text{with rate } S_i \dfrac{\sum_{l \neq i} S_l}{n-1} \\ (S,W) + (0, -e_i + e_j) & \text{with rate } W_i \dfrac{W_j}{n-1} \\ (S,W) + (e_j, -e_i) & \text{with rate } W_i \dfrac{S_j}{n-1} \end{cases}$$

The limit ODE

- For every $i = 1,2,\dots,m$ and $t \geq 0$

$$\frac{d}{dt}s_i(t) = \big(1 - 2s(t) + s_i(t)\big)s_i(t)$$

$$\frac{d}{dt}u_i(t) = s_i(t) - s(t)u_i(t)$$

- $u_i(t) = s_i(t) + w_i(t)$

- $s(t) = \sum_i s_i(t)$

δ - convergence time

Given $\delta \in (0,1)$, t_δ defined as follows

$$\sum_{i=1}^{k} u_i(t_\delta) = 1 - \delta$$

Limit points

- Theorem – Suppose that for $1 \leq k \leq m$

  $$u_1(0) = \cdots = u_k(0) > u_{k+1}(0) \geq \cdots \geq u_m(0)$$

  and $\sum_i s_i(0) = 1$.

  Then

  $$u_1(t) = \cdots = u_k(t) > u_{k+1}(t) \geq \cdots \geq u_m(t), \text{for } \forall t \geq 0$$

  Moreover, we have

  $$\lim_{t\to\infty} u_i(t) = \begin{cases} \frac{1}{k} & \text{if } i = 1,2,\dots,k \\ 0 & \text{else} \end{cases}$$

Limit points (cont'd)

The last theorem follows as a corollary of the following claims:

- $s_1(t) = \cdots = s_k(t) > s_{k+1}(t) \geq \cdots \geq s_m(t), \text{for } \forall t \geq 0$
- $\lim_{t \to \infty} s_i(t) = \begin{cases} \frac{1}{2k-1} & \text{if } i = 1,2,\dots,k \\ 0 & \text{else} \end{cases}$
- $s(t) \geq \frac{1}{2-\frac{1}{m}}, \forall t \geq 0$

Rate of convergence

- For every non-plurality state $k < i \leq m$

$$\lim_{t \to \infty} \frac{1}{t} \log(s_i(t)) = \lim_{t \to \infty} \frac{1}{t} \log(w_i(t)) = -\frac{1}{2k-1}$$

- Exponential diminishing of non-plurality states

Convergence time

- Theorem: For $s_1(0) = \cdots = s_k(0) > s_{k+1}(0) \geq \cdots \geq s_m(0)$ such that $s_1(0) - s_{k+1}(0) \geq \epsilon > 0$ and $\sum_{i=1}^{m} s_i(0) = 1$, there exists a constant $c_m > 0$ such that

$$t_\delta \leq (2m-1)\left(\log\left(\frac{1}{\delta}\right) + \log\left(\frac{1}{\epsilon}\right)\right) + c_m$$

- Corollary: $t_\delta = O(m(\log(1/\delta) + \log(1/\epsilon) + \log(m))$

- Convergence time linear in the number of alternatives*
- Logarithmic in the voting margin

* Up to poly-log factors

Convergence lower bounds

- Theorem: For $u_1(0) = \cdots = u_k(0) > u_{k+1}(0) \geq \cdots \geq u_m(0)$

$$t_\delta \geq (2k-1)\log\left(\frac{1}{\delta}\right)$$

Convergence time lower bounds (cont'd)

- Theorem: For every $m > 1$ there exists an initial state with gap ϵ and constant c_m such that for ϵ and δ small enough

$$t_\delta \geq (m-1)\log\left(\frac{1}{\delta}\right) + (2m-1)\log\left(\frac{1}{\epsilon}\right) + c_m$$

- Take:

$$s_i(0) = \begin{cases} \frac{1}{m} + \frac{\epsilon}{2} & i = 1,2,\ldots,m/2 \\ \frac{1}{m} - \frac{\epsilon}{2} & i = m/2+1,\ldots,m \end{cases}$$

Probability of Error

[Babace-Draief-2013]

- Theorem - suppose that for $\alpha_1 > \alpha_2 \geq \cdots \geq \alpha_m > 0$,

$$\lim_{n\to\infty} \frac{1}{n}(U_1(0), U_2(0), \dots, U_m(0)) \to (\alpha_1, \alpha_2, \dots, \alpha_m)$$

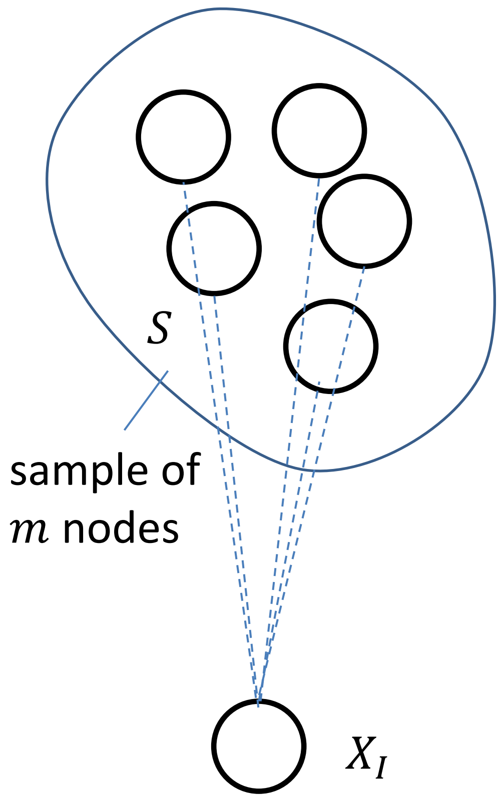
Then

$$\frac{1}{n}\log_2\, p_e \leq -(\alpha_1+\alpha_2) D\left(\frac{\alpha_1}{\alpha_1+\alpha_2} || \frac{1}{2}\right) + \frac{\log_2(m-1)}{n}$$

Polling algorithm

[Cruise-Ganesh-2013]

Y = number of nodes in S in state 1



do:

- Sample node I uniformly at random
- $S \leftarrow$ Sample of m nodes from the population with replacement
- $Y \leftarrow$ number of nodes in S in state 1

- If $X_I = 0 \,\&\, Y \geq d$

  $$X_I \leftarrow 1$$

- Else if $X_I = 1 \,\&\, Y \leq m - d$

  $$X_I \leftarrow 0$$

Polling algorithm (cont'd)

[Cruise-Ganesh-2013]

- Probability of error:

$$p_e \leq c_{\epsilon,m} exp\left(-(n-1)(m-1)D\left(\frac{1-\epsilon}{2} \,||\, \frac{1}{2}\right)\right)$$

- Expected convergence time:

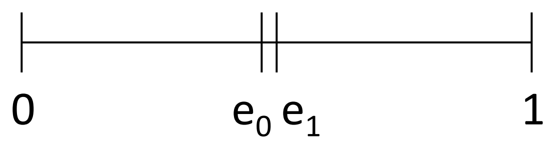
$$\Theta(\log n)$$

Outline

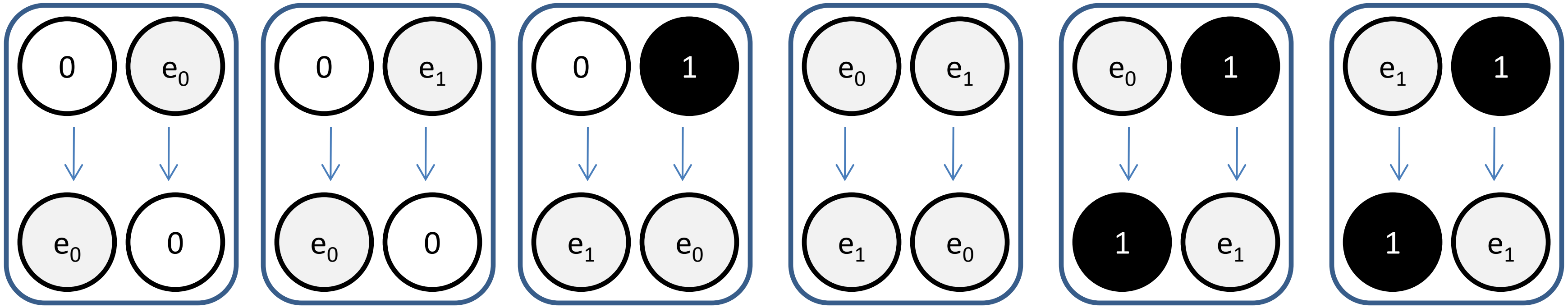
- Related work
- 3-state algorithm
- 4-state algorithm
- Conclusion

Quaternary protocol

- Four states



- Update rules: *swap* or *annihilate*



Correctness

[Benezit-Thiran-Vetterli-2010]

- Corollary - For any given connected graph, the binary interval consensus converges to the correct state with probability 1.

Convergence time

- Each edge (i, j) activated at instances of a Poisson point process of intensity $q_{i,j} \geq 0$
- Contract rate matrix: $Q = \left(q_{i,j}\right)$
- Family of matrices: for every non-empty subset of nodes S, $Q^S = (q_{i,j}^S)$ defined by

$$q_{i,j}^S = \begin{cases} -\sum_l q_{i,l} & \text{if } i = j \\ q_{i,j} & \text{if } i \notin S, j \neq i \\ 0 & \text{if } i \in S, j \neq i \end{cases}$$

Eigenvalue gap

- For any finite graph G, there exists $\delta(G, \alpha) > 0$ such that every eigenvalue λ of matrix Q_S satisfies

$$\lambda \leq -\delta(G, \alpha) < 0$$

Convergence time

- Two phases
  - Phase 1: time T_1 until depletion of state 1
  - Phase 2: time T_2 until depletion of state 2
- Theorem:

$$\mathbf{E}[T_1] \leq \frac{1}{\delta(G,\alpha)}(\log(n)+1)$$

$$\mathbf{E}[T_2] \leq \frac{1}{\delta(G,\alpha)}(\log(n)+1)$$

State evolution in Phase 1

- Phase 1

$$(Z, A) \to \begin{cases} (Z - e_i, A - e_j) & \text{with rate } q_{i,j} Z_i A_j \\ (Z - e_i + e_j, A) & \text{with rate } q_{i,j} Z_i (1 - A_j - Z_j) \\ (Z, A - e_i + e_j) & \text{with rate } q_{i,j} A_i (1 - A_j - Z_j) \end{cases}$$

1 if node *i* in state 1

1 if node *i* in state 0

State evolution in Phase 1 (cont'd)

- Probability that a node is in state 1 evolves as

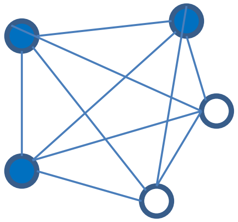
$$\frac{d}{dt}\mathbf{E}[A_i(t)] = -\left(\sum_l q_{i,l}\right)\mathbf{E}[A_i(t)] + \sum_j q_{i,j}\mathbf{E}\big[A_j(t)\big(1 - Z_i(t)\big)\big]$$

- System of linear ODEs:

$\frac{d}{dt}\mathbf{E}_k[A(t)] = Q_{S_k}\mathbf{E}_k[A(t)]$, S_k = set of nodes in state 0

- Bounds on the expected convergence time follow using a spectral bound

Complete graph



- Each edge activate at rate $\frac{1}{n-1}$
- $\delta(G,\alpha) = \epsilon \frac{n}{n-1}$
- $E[T_i] \leq \frac{1}{\epsilon}(\log(n) + 1)$, for $i = 1,2$

Complete graph: upper bound is tight

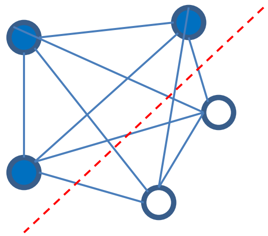
- By direct analysis:

$$\mathbf{E}[T_1] = \frac{n-1}{|S_0| - |S_1|}\left(H_{|S_1|} + H_{|S_0|-|S_1|} - H_{|S_0|}\right)$$

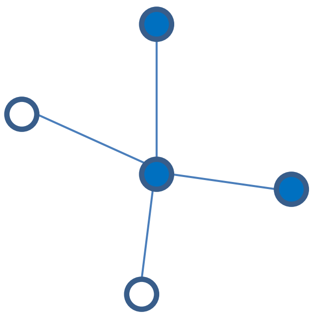
where $H_k = \sum_{i=1}^{k} 1/i$ is the k-th harmonic number

$$\Rightarrow \quad \boxed{\mathbf{E}[T_1] = \frac{1}{\epsilon}\log(n) + O(1)}$$

- 0 and 1 states annihilate after a random time with exponential distribution with parameter $\mathrm{E}(S_0(t), S_1(t)) \;/\; (n-1)$



Star



- Each edge activate at rate $\frac{1}{n-1}$

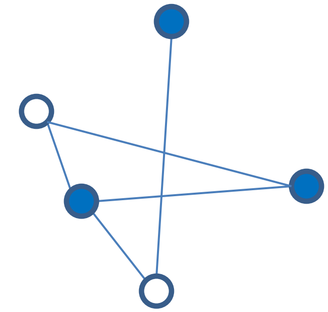
- $\delta(G,\alpha) = \frac{n-\sqrt{n^2-4\epsilon n}}{2(n-1)}$

$\Rightarrow$ $\mathbf{E}[T_i] \le \frac{1}{\epsilon} n \log(n)$, for $i = 1,2$

- Tight: by direct analysis

$$E[T_1] = \frac{1}{\epsilon(2-\epsilon)} n \log(n) + O(n)$$

Erdos-Renyi graph



- Edge (u,v) activated at rate $\frac{X_{(u,v)}}{np_n}$, $X_{(u,v)} \sim \mathrm{Ber}(p_n)$

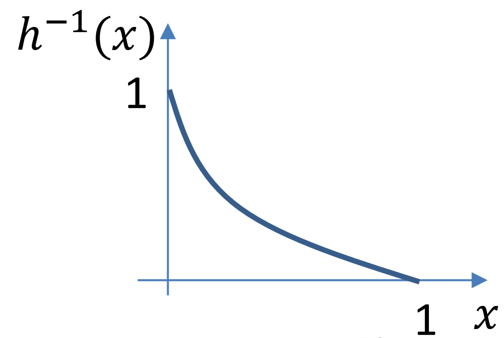
$$p_n = c\frac{\log(n)}{n}, \text{ for } c > 0$$

- If $c\epsilon > 2$

$$\frac{1}{\delta(g,\alpha)} \leq \frac{1}{\epsilon h^{-1}\left(\frac{2}{c\epsilon}\right)} + O\left(\frac{1}{\log(n)}\right) \text{ w.h.p.}$$

where $h^{-1}(x)$ is the inverse function of

$$h(x) = x\log(x) + 1 - x$$



Conclusion

- 3-state algorithm
  - Complete graph: correctness with high probability (exponentially decreasing error probability in n), fast convergence $O(\log(n) + \log(1/\epsilon))$
  - Extensions to plurality problem

- 4-state algorithm
  - Arbitrary connected graph: guaranteed correctness, expected convergence time upper bounds
  - Complete graph: expected convergence time $O((1/\epsilon)\log(n))$

Some open problems

- Lower bounds? - given memory and communication constraints and a probability of error budget, lower bounds for the expected convergence time?

- Better upper bounds?

- Tradeoff accuracy-convergence time: dependence on the memory and communication constraints and the network structure?

References

- J. Koplowitz, Necessary and Sufficient Memory Size for m-hypothesis Testing, IEEE Trans. on Information Theory, Vol 21, No 1, 1975
- M. Greenwald and S. Khanna, Space-efficient Online Computation of Quantile Summaries, ACM SIGMOD 2001
- Y. Hassin and D. Peleg, Distributed Probabilistic Polling and Applications to Proportionate Agreement, Information and Computation, 171, 2001
- D. Kempe, J. Kleinberg and E. Tardos, Maximizing Influence through a Social Network, ACM KDD 2003
- M. Greenwald and S. Khanna, Power-conserving Computation of Order-Statistics over Sensor Networks, ACM PODS 2004
- T. M. Liggett, Interacting Particle Systems, Springer, 2006
- S. Boyd, A. Ghosh, B. Prabhakar and D. Shah, Randomized gossip algorithms, IEEE Trans. on Information Theory, Vol 52, No 6, 2006
- D. Angluin, J. Aspnes, D. Eisenstat, A Simple Population Protocol for Fast Robust Approximate Majority, DISC, 2007
- F. Kuhn, T. Locher, R. Wattenhofer, Tight Bounds for Distributed Selection, ACM SPAA 2007

References (cont'd)

- W. P. Tay, J. N. Tsitsiklis and M. Z. Win, On the Subexponential Decay of Detection Error Probabilities in Long Tandems, IEEE Trans. on Info. The., Vol 54, No 10, 2008
- A. Nedic, A. Olshevsky, A. Ozdaglar and J. N. Tsitsiklis, Distributed Averaging Algorithms and Quantization Effects, IEEE Conf. on Decision and Control, 2008
- F. Benzit, P. Thiran and M. Vetterli, Interval Consensus: From Quantized Gossip to Voting, IEEE Int'l Conf. on Acoustics, Speech, and Signal Processing, 2009
- E. Perron, D. Vasudevan, M. V., Using Three States for Binary Consensus on Complete Graphs, IEEE Infocom 2009
- J. Cruise and A. Ganesh, Probabilistic Consensus via Polling and Majority Rules, Proc. of Allerton Conference, 2010
- D. Acemoglu, M. A. Dahleh, I. Lobel and A. Ozdaglar, Bayesian Learning in Social Networks, forthcoming Review of Economic Studies, 2011
- F. Benezit, P. Thiran and M. Vetterli, The Distributed Multiple Voting Problem, IEEE Journal on Selected Topics in Signal Processing, Vol 5, No. 4, 2011
- M. Draief and M. V., Convergence Speed of Binary Interval Consensus, SIAM J. Control Optim., vol 50, pp 1087-1109

References (cont'd)

- E. Mossel, J. Neeman and O. Tamuz, Majority Dynamics and Aggregation of Information in Social Networks, 2012
- F. Chierichetti and J. Kleinberg, Voting with Limited Information and Many Alternatives, ACM SODA 2012
- M. A. Abdullah and M. Draief, Majority Consensus on Random Graphs of a Given Degree Sequence, ArXiv, 2012
- K. Jung, B. Y. Kim, M. V., Distributed Ranking in Networks with Limited Memory and Communication, IEEE ISIT 2012
- S. Shang, P. W. Cuff, S. R. Kulkarni and P. Hui, An Upper Bound on the Convergence Time for Distributed Binary Consensus, 15th Int'l Conf. on Information Fusion, 2012
- Z. Huang, K. Yi, and Q. Zhang, Randomized Algorithms for Tracking Distributed Count, Frequencies and Ranks, ACM PODS 2012
- A. Babaee and M. Draief, Distributed Multivalued Consensus, Computer and Information Sciences III, 2013

References (cont'd)

- G. B. Mertzios, S. E. Nikoletseas, C. L. Raptopoulos, P. G. Spirakis, Determining Majority in Networks with Local Interactions and very Small Local Memory, ICALP 2014
- M. Feldman, N. Immorlica, B. Lucier, S. M. Weinberg, Information Aggregation in Social Networks, working paper, 2014