

# Majority Consensus Algorithms and Spectral Optimisation

Moez Draief

Intelligent Systems and Networks Group  
Imperial College London

**Workshop on Large Evolving Networks**  
**Heilbronn Institute for Mathematical Research**

## Computational power of (social) networks

### Off to the movies...

- Friends who want to watch one of two movies (together)
- They interact in pairs in order to come to a common choice before the end of the day
- Majority prefers one theater

What protocol should they run to decide to go to the latter?

## Computational power of (social) networks

### Off to the movies...

- Friends who want to watch one of two movies (together)
- They interact in pairs in order to come to a common choice before the end of the day
- Majority prefers one theater

What protocol should they run to decide to go to the latter?

### Biased voting [Kearns et al '09]

- Players paid if consensus conforms with their preference
- For some network topologies, minority preference consistently wins
- Individual behavioral characteristics (stubbornness, awareness of opposing incentives) correlate with earnings

## Consensus on networks

### Distributed computing

- Information fusion/consistency in distributed networks
- Network awareness (Computing graph properties)
- Multi-agent coordination and flocking

## Consensus on networks

### Distributed computing

- Information fusion/consistency in distributed networks
- Network awareness (Computing graph properties)
- Multi-agent coordination and flocking

### Distributed algorithms

- Numerous impossibility results in the deterministic case [Lynch et al '90s]
- Random walks
- Gossiping algorithms

## Binary majority consensus

### Desired outcome and metrics

- Nodes end with opinion held by majority of nodes
- Node can probe neighbours and update opinion accordingly using little (constant) memory
- Probability of error (convergence to incorrect consensus)
- Time to convergence

## Binary majority consensus

### Desired outcome and metrics

- Nodes end with opinion held by majority of nodes
- Node can probe neighbours and update opinion accordingly using little (constant) memory
- Probability of error (convergence to incorrect consensus)
- Time to convergence

### Applications

- Occurrence of a given event in cooperative decision making
- Voting in distributed systems
- Routine to solve more elaborate distributed decision making instances

## Outline

- 1 Majority consensus: algorithms**
  - Voter Model
  - Averaging Process
  - Binary consensus
- 2 Examples**
  - Complete graph
  - Star, ER, Ring, Line
- 3 Optimization of Majority Consensus**
  - Faster convergence time
  - Implementation
  - Examples



# Outline

- 1 Majority consensus: algorithms**
  - Voter Model
  - Averaging Process
  - Binary consensus
- 2 Examples**
  - Complete graph
  - Star, ER, Ring, Line
- 3 Optimization of Majority Consensus**
  - Faster convergence time
  - Implementation
  - Examples

## Model

### Interaction Model

- Connected undirected graph  $G = (V, E)$ ,  $|V| = n$
- $\alpha n$  nodes hold 0 and  $(1 - \alpha)n$  nodes hold 1,  $\alpha \in (1/2, 1)$
- Nodes  $i$  and  $j$  interact at rate  $q_{ij} = q_{ji}$ ,  $q_{ij} \neq 0$  iff  $(i, j) \in E$

### Markov chain

- $(X_t)_{t \geq 0}$  with rate matrix  $Q$ ,  $q_{ii} = -\sum_{i \neq j} q_{ij}$
- $(\pi_i)_{i \in V}$  stationary distribution is uniform on  $V$ . Mixing time:

$$|\mathbb{P}_j(X_t = i) - 1/n| = O\left(e^{-\lambda_2(Q)t}\right)$$

where  $\lambda_2(Q) = \inf\{\sum_{i,j} q_{ij}(x_i - x_j)^2, \|x\| = 1, x^T \mathbf{1} = 0\}$

## Probability of error

### Interaction Model

- Connected undirected graph  $G = (V, E)$ ,  $|V| = n$
- $\alpha n$  nodes hold 0 and  $(1 - \alpha)n$  nodes hold 1,  $\alpha \in (1/2, 1)$
- Nodes  $i$  and  $j$  interact at rate  $q_{ij}$  and  $i$  updates to  $j$ 's state w.p.  $1/2$

### Theorem [Liggett '85, Hassin-Peleg '01]

- The number of nodes in state 1 is a martingale.
- Probability of reaching (wrong) consensus at 1 is  $1 - \alpha$ .

## Time to convergence [Aldous 2012]

### Complete graph

- Each edge has rate  $1/(n-1)$ . The number of agents with opinion 1 evolves as Birth and Death proces with

$$\lambda_{k,k+1} = \lambda_{k,k-1} = \frac{k(n-k)}{2(n-1)}.$$

- Time to convergence =  $O(n)$

### General graph

- Conductance  $\eta(Q) = \inf_{A \subset V} \frac{\sum_{i \in A, j \in A^c} q_{ij}}{|A||A^c|/n}$
- The Markov chain tracking the number of nodes in state 0 evolves at least  $\eta(Q)$  times as fast as on the complete graph,  
Time to convergence  $O(n/\eta(Q))$

## Time to convergence [Cooper et al 2012.]

### Cheeger's inequality

- Conductance:  $\eta(Q) = \inf_{A \subset V} \frac{\sum_{i \in A, j \in A^c} q_{ij}}{|A||A^c|/n}$

- Spectral Gap:

$$\lambda_2(Q) = \inf\{\sum_{i,j} q_{ij}(x_i - x_j)^2, \|x\| = 1, x^T \mathbf{1} = 0\}$$

$$\lambda_2(Q) \leq \eta(Q).$$

- Time to convergence of voter model  $O(n/(\lambda_2(Q)))$ .

## Time to convergence [Cooper et al 2012.]

### Cheeger's inequality

- Conductance:  $\eta(Q) = \inf_{A \subset V} \frac{\sum_{i \in A, j \in A^c} q_{ij}}{|A||A^c|/n}$

- Spectral Gap:

$$\lambda_2(Q) = \inf \left\{ \sum_{i,j} q_{ij} (x_i - x_j)^2, \|x\| = 1, x^T \mathbf{1} = 0 \right\}$$

$$\lambda_2(Q) \leq \eta(Q).$$

- Time to convergence of voter model  $O(n/(\lambda_2(Q)))$ .

Let  $S$  of size  $k$  be the subset realising the inf in  $\eta(Q)$  and let  $x$

such that  $x_i = -\sqrt{\frac{n-k}{kn}}$ ,  $i \in S$  and  $x_i = \sqrt{\frac{k}{(n-k)n}}$ ,  $i \in S^c$ .

## Distributed averaging

At each interaction of  $(i, j)$  occurring at rate  $q_{ij}$

$$x_i(t) = x_j(t) = \frac{x_i(t-) + x_j(t-)}{2}.$$

### Theorem [Boyd et al '06, Aldous-Lanoue '12]

- Algorithm converges to the average value, using  $O(\text{Poly}(\log(n)))$  memory per node
- Time to convergence to up  $O(1/n)$  error of the average is

$$\Theta(\log(n)/\lambda_2(Q)),$$

## Distributed averaging: Proof

Assume that  $\sum_i x_i(0) = 0$ .

Let  $Q(t) = \|x(t)\|^2$ . When an  $i, j$  interaction takes place  $Q(t)$  reduces by  $(x_i - x_j)^2/2$ .

$$\begin{aligned} \mathbb{E}(dQ(t) \mid x(t) = x) &= - \sum_{i,j} q_{ij} \frac{(x_i - x_j)^2}{2} dt \\ &\leq -\lambda_2(Q) \|x\|^2 / 2 dt \end{aligned}$$

In particular

$$\mathbb{E}\|x(t)\|^2 \leq \|x(0)\|^2 e^{-\lambda_2(Q)t/2}$$



Binary consensus

## Small memory

Could we use less memory and still guarantee small error?

## Small memory

Could we use less memory and still guarantee small error?

### Impossibility [Land-Belew '95]

- Connected undirected graph  $G = (V, E)$ ,  $|V| = n$ ,
- $\alpha n$  nodes in 0 and  $(1 - \alpha)n$  nodes in 1,  $\alpha \in (1/2, 1)$ ,  $2\alpha - 1$  is the *voting margin*.
- $i$  contacts  $j$  at rate  $q_{ij} > 0 \forall (i, j) \in E$

No 1-bit distributed algorithm can solve the majority consensus problem.

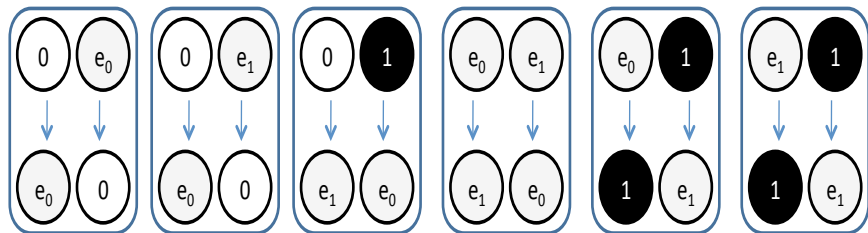
## Binary Consensus with two undecided states

Averaging-like updates: States  $0 < e_0 < e_1 < 1$ .

## Binary Consensus with two undecided states

Averaging-like updates: States  $0 < e_0 < e_1 < 1$ .

Rules: Swaps + Annihilation



## Mean-field analysis (Complete graph)

Let  $q_{ij} = \frac{1}{n-1}$ ,  $i \neq j$  and  $\mathbf{X}(t) = (|S_0(t)|, |S_{e_0}(t)|, |S_{e_1}(t)|, |S_1(t)|)$  is a Markov process with the following transition rates

$$\rightarrow \begin{cases} (|S_0(t)| - 1, |S_{e_0}(t)| + 1, |S_{e_1}(t)| + 1, |S_1(t)| - 1) & : \frac{|S_0(t)||S_1(t)|}{n-1} \\ (|S_0(t)|, |S_{e_0}(t)| - 1, |S_{e_1}(t)| + 1, |S_1(t)|) & : \frac{|S_{e_0}(t)||S_1(t)|}{n-1} \\ (|S_0(t)|, |S_{e_0}(t)| + 1, |S_{e_1}(t)| - 1, |S_1(t)|) & : \frac{|S_0(t)||S_{e_1}(t)|}{n-1}. \end{cases}$$

## Mean-field analysis (Complete graph)

Let  $q_{ij} = \frac{1}{n-1}$ ,  $i \neq j$  and  $\mathbf{X}(t) = (|S_0(t)|, |S_{e_0}(t)|, |S_{e_1}(t)|, |S_1(t)|)$  is a Markov process with the following transition rates

$$\rightarrow \begin{cases} (|S_0(t)| - 1, |S_{e_0}(t)| + 1, |S_{e_1}(t)| + 1, |S_1(t)| - 1) & : \frac{|S_0(t)||S_1(t)|}{n-1} \\ (|S_0(t)|, |S_{e_0}(t)| - 1, |S_{e_1}(t)| + 1, |S_1(t)|) & : \frac{|S_{e_0}(t)||S_1(t)|}{n-1} \\ (|S_0(t)|, |S_{e_0}(t)| + 1, |S_{e_1}(t)| - 1, |S_1(t)|) & : \frac{|S_0(t)||S_{e_1}(t)|}{n-1} \end{cases}$$

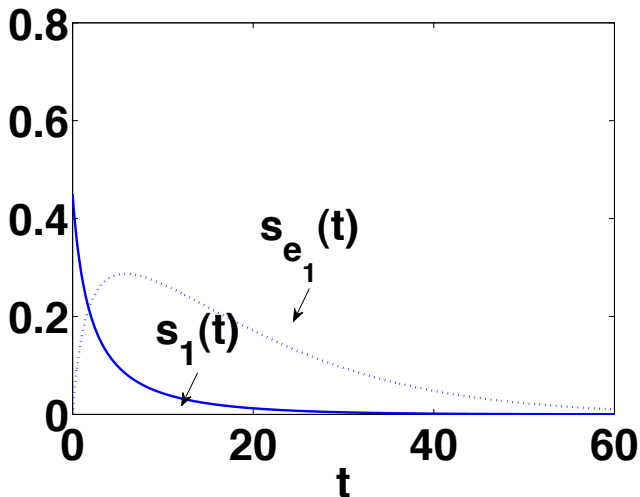
By Kurtz,  $\mathbf{X}(t)/n$  converges to  $(s_0(t), s_{e_0}(t), s_{e_1}(t), s_1(t))$ ,

$$s_{e_1}(t) \sim (2\alpha - 1) \frac{1 - \alpha}{\alpha} t e^{-(2\alpha-1)t}$$

$$s_1(t) \sim (2\alpha - 1) \frac{1 - \alpha}{\alpha} e^{-(2\alpha-1)t}.$$

Binary consensus

## Minority states



## General bound

### Theorem [Draief-Vojnovic '12]

Let  $T$  be the time until there are only nodes in state 0 and  $e_0$ .

$$\mathbb{E}(T) = O(\log n / \delta(G, \alpha))$$

where

$$\delta(Q, \alpha) = \min_{S \subset V, |S|=(2\alpha-1)n} \min_{\lambda \in \text{Spec}(Q_S)} |\lambda|$$

### Generalised conductance lemma [Babaee-Draief '13+]

We have

$$\delta(Q, \alpha) \geq c_\alpha \lambda_2(Q)$$

In particular,

$$\mathbb{E}(T) = O(\log(n) / \lambda_2(Q))$$



## Proof: Depletion of nodes in state 1

Let  $A_j$  and  $Z_j$  indicator node in 0 and 1 resp. The transitions of the Markov process  $(Z, A)$  is given by

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

## Proof: Depletion of nodes in state 1

Let  $A_i$  and  $Z_i$  indicator node in 0 and 1 resp. The transitions of the Markov process  $(Z, A)$  is given by

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

For  $t \in [t_k, t_{k+1})$  where  $\{S_0(t) = S_k\}$

$$\frac{d}{dt} \mathbb{E}_k(A_i(t)) = - \left( \sum_{l \in V} q_{i,l} \right) \mathbb{E}_k(A_i(t)) + \begin{cases} \sum_{j \in V} q_{i,j} \mathbb{E}_k(A_j(t)), & i \notin S_k \\ 0, & i \in S_k \end{cases}$$

where  $\mathbb{E}_k$  is the expectation conditional on  $\{S_0(t) = S_k\}$ .

**(random) Piecewise-linear dynamical system****Dynamics**

The dynamics of the system boils down to  $Y(t) = (Y_i(t))_{i \in V}$ ,

$$\frac{d}{dt} \mathbb{E}_k(Y(t)) = Q_{S_k} \mathbb{E}_k(Y(t)),$$

for  $t \in [t_k, t_{k+1})$  during which  $\{S_0(t) = S_k\}$  and  $Q_{S_k}$  is given by

$$Q_S(i, j) = \begin{cases} -\sum_{l \in V} q_{i,l}, & i = j \\ q_{i,j}, & i \notin S, j \neq i \\ 0, & i \in S, j \neq i. \end{cases}$$

# Solution

## Proposition

Solving the above differential equation and using the strong Markov property

$$\mathbb{E}(Y(t)) = \mathbb{E} \left[ e^{\lambda(t)} Y(0) \right]$$

where  $\lambda(t) = Q_{S_k}(t - t_k) + \sum_{l=0}^{k-1} Q_{S_l}(t_{l+1} - t_l)$ .

## Proof: Spectrum of $Q_S$

For any finite graph  $G$ , there exists  $\delta(G, \alpha) > 0$  such that, for any non-empty subset of vertices  $S$  with  $|S| \in [(2\alpha - 1)n, \alpha n]$ , if  $\lambda$  is an eigenvalue of the the matrix  $Q_S$  defined in, then

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

## Proof: Spectrum of $Q_S$

For any finite graph  $G$ , there exists  $\delta(G, \alpha) > 0$  such that, for any non-empty subset of vertices  $S$  with  $|S| \in [(2\alpha - 1)n, \alpha n]$ , if  $\lambda$  is an eigenvalue of the the matrix  $Q_S$  defined in, then

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

- First  $(-\sum_{l \in V} q_{i,l})$ ,  $i \in S$  are eigenvalues of  $Q_S$
- The remaining eigenvalues correspond to eigenvectors of the form  $(\underbrace{x}_{S^c}, \underbrace{0, \dots, 0}_S)^T$ . Let  $W \subset S^c$  such that for  $i \in W$ ,  $x_i \neq 0$

$$-\lambda = \sum_{i \in W} \sum_{j \in S} q_{i,j} x_i^2 + \sum_{i \in W, j \in S^c \setminus W} q_{i,j} x_i^2 + \frac{1}{2} \sum_{i,j \in W} q_{i,j} (x_i - x_j)^2$$

## Proof: The End

Note that

$$\mathbb{E}(Y(t)) = \mathbb{E} \left[ e^{\lambda(t)} Y(0) \right]$$

where  $\lambda(t) = Q_{S_k}(t - t_k) + \sum_{l=0}^{k-1} Q_{S_l}(t_{l+1} - t_l)$

# Proof: The End

Note that

$$\mathbb{E}(Y(t)) = \mathbb{E} \left[ e^{\lambda(t)} Y(0) \right]$$

where  $\lambda(t) = Q_{S_k}(t - t_k) + \sum_{l=0}^{k-1} Q_{S_l}(t_{l+1} - t_l)$

Hence

$$\|\mathbb{E}(Y(t))\|_2 \leq \mathbb{E} \left[ \|e^{Q_{S_k}(t-t_k)}\| \prod_{l=0}^{k-1} \|e^{Q_{S_l}(t_{l+1}-t_l)}\| \|Y(0)\|_2 \right] \leq \sqrt{n} e^{-\delta(G, \alpha)t}$$



## Proof: The End

Note that

$$\mathbb{E}(Y(t)) = \mathbb{E} \left[ e^{\lambda(t)} Y(0) \right]$$

where  $\lambda(t) = Q_{S_k}(t - t_k) + \sum_{l=0}^{k-1} Q_{S_l}(t_{l+1} - t_l)$

Hence

$$\|\mathbb{E}(Y(t))\|_2 \leq \mathbb{E} \left[ \|e^{Q_{S_k}(t-t_k)}\| \prod_{l=0}^{k-1} \|e^{Q_{S_l}(t_{l+1}-t_l)}\| \|Y(0)\|_2 \right] \leq \sqrt{n} e^{-\delta(G,\alpha)t}$$

Therefore, by Cauchy-Schwartz, we have

$$\mathbb{P}(Y(t) \neq \mathbf{0}) \leq \sum_{i \in V} \mathbb{E}(Y_i(t)) \leq n e^{-\delta(G,\alpha)t}$$

## Summary

- Upper bound on the expected convergence time for a number of distributed candidate dynamics for solving Majority consensus
- Bounds based on the location of the spectral gap of rate matrix (generalised-cut: quick for expander graphs).
- For binary consensus, expected convergence time critically depends on the voting margin

## Summary

- Upper bound on the expected convergence time for a number of distributed candidate dynamics for solving Majority consensus
- Bounds based on the location of the spectral gap of rate matrix (generalised-cut: quick for expander graphs).
- For binary consensus, expected convergence time critically depends on the voting margin
- Application to particular network topologies: complete graphs, stars, ER graph, paths, cycles.

## Outline

- 1 **Majority consensus: algorithms**
  - Voter Model
  - Averaging Process
  - Binary consensus
- 2 **Examples**
  - Complete graph
  - Star, ER, Ring, Line
- 3 **Optimization of Majority Consensus**
  - Faster convergence time
  - Implementation
  - Examples

## Upper Bounds

### Corollary

An application of the theorem to complete graph  $q_{i,j} = \frac{1}{n-1}$  for all  $i \neq j$ , yields

$$\mathbb{E}(T_i) \leq \frac{1}{2\alpha - 1} \log(n).$$

### Exact asymptotics

A direct analysis of the dynamics of the 1st phase tracking the interactions of nodes in state 1 and nodes in state 0 implies that

$$\mathbb{E}(T_1) = \frac{n-1}{|S_0| - |S_1|} (H_{|S_1|} + H_{|S_0| - |S_1|} - H_{|S_0|})$$

where  $H_k = \sum_{i=1}^k \frac{1}{i}$

Complete graph

## Various initial conditions

- $|S_0| - |S_n| = (2\alpha - 1)n$ ,  $\alpha$  a constant larger than  $1/2$

$$\mathbb{E}(T_1) = \frac{1}{2\alpha - 1} \log(n) + O(1).$$

- If  $|S_0| = |S_1|$

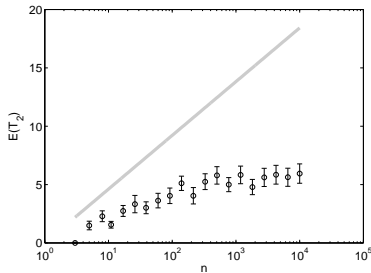
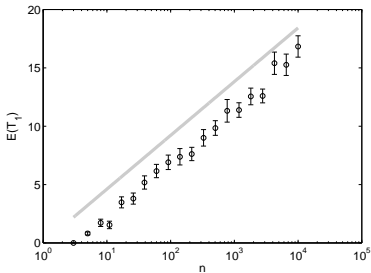
$$\mathbb{E}(T_1) = \frac{\pi^2}{6} n(1 + o(1)).$$

- $\mu_n = (|S_0| - |S_1|)/n$  is strictly positive but small ( $o(1)$ ),

$$\mathbb{E}(T_1) = \frac{1}{\mu_n} \log(n\mu_n) + O(1).$$

Complete graph

# Complete Graph: Theory v. Simulation



## Star

- **Star Network:**  $q_{1,i} = q_{i,1} = \frac{1}{n-1}$ ,  $i \neq 1$  and  $q_{i,j} = 0$ ,  $i, j \neq 1$ .  
 $\mathbb{E}(T_i) \leq \frac{1}{2\alpha-1} n \log(n)$ . Using, direct calculation

$$\mathbb{E}(T_1) = \frac{1}{(2\alpha-1)(3-2\alpha)} n \log(n) + O(n)$$

- **ER-graph:**  $q_{i,j} = \frac{1}{np_n} X_{i,j}$   $X_{i,j}$  i.i.d. Bernoulli r.v. with mean  $c \frac{\log(n)}{n}$ ,  $c > \frac{2}{2\alpha-1}$ , for  $h^{-1}$  the inverse of  $h(x) = x \log(x) + 1 - x$ ,

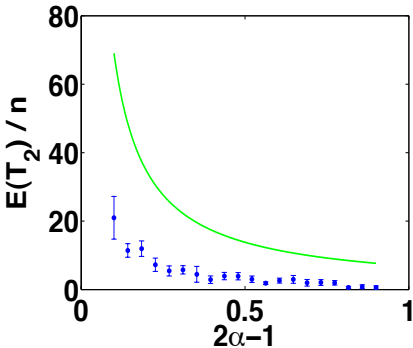
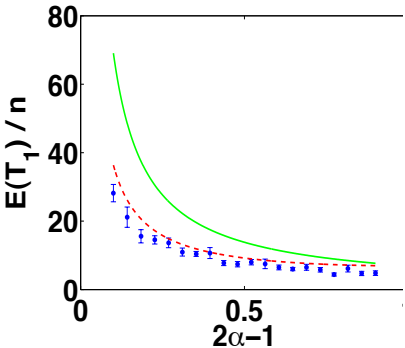
$$\mathbb{E}(T_i) \leq \frac{1}{(2\alpha-1)h^{-1}\left(\frac{2}{c(2\alpha-1)}\right)} \log(n) + O(1)$$

- **Path:**  $\mathbb{E}(T_i) \leq \frac{16(1-\alpha)^2}{\pi^2} n^2 \log(n) + O(1)$
- **Ring:**  $\mathbb{E}(T_i) \leq \frac{4(1-\alpha)^2}{\pi^2} n^2 \log(n) + O(1)$ .



Star, ER, Ring, Line

# Star



## ER-graph

- **Star Network:**  $q_{1,i} = q_{i,1} = \frac{1}{n-1}$ ,  $i \neq 1$  and  $q_{i,j} = 0$ ,  $i, j \neq 1$ .  
 $\mathbb{E}(T_i) \leq \frac{1}{2\alpha-1} n \log(n)$ . Using, direct calculation

$$\mathbb{E}(T_1) = \frac{1}{(2\alpha-1)(3-2\alpha)} n \log(n) + O(n)$$

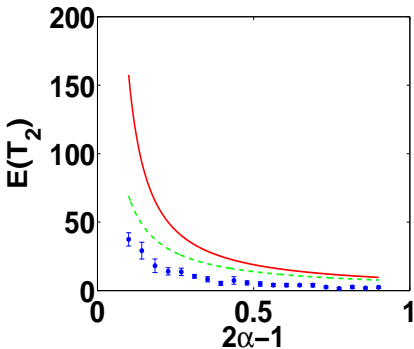
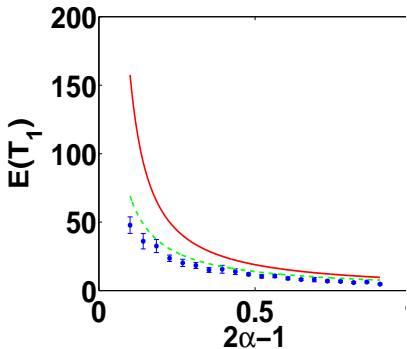
- **ER-graph:**  $q_{i,j} = \frac{1}{np_n} X_{i,j}$   $X_{i,j}$  i.i.d. Bernoulli r.v. with mean  $c \frac{\log(n)}{n}$ ,  $c > \frac{2}{2\alpha-1}$ , for  $h^{-1}$  the inverse of  $h(x) = x \log(x) + 1 - x$ ,

$$\mathbb{E}(T_i) \leq \frac{1}{(2\alpha-1)h^{-1}\left(\frac{2}{c(2\alpha-1)}\right)} \log(n) + O(1)$$

- **Path:**  $\mathbb{E}(T_i) \leq \frac{16(1-\alpha)^2}{\pi^2} n^2 \log(n) + O(1)$
- **Ring:**  $\mathbb{E}(T_i) \leq \frac{4(1-\alpha)^2}{\pi^2} n^2 \log(n) + O(1)$ .

Star, ER, Ring, Line

# ER-graph



# Path and Ring

- **Star Network:**  $q_{1,i} = q_{i,1} = \frac{1}{n-1}$ ,  $i \neq 1$  and  $q_{i,j} = 0$ ,  $i, j \neq 1$ .  
 $\mathbb{E}(T_i) \leq \frac{1}{2\alpha-1} n \log(n)$ . Using, direct calculation

$$\mathbb{E}(T_1) = \frac{1}{(2\alpha-1)(3-2\alpha)} n \log(n) + O(n)$$

- **ER-graph:**  $q_{i,j} = \frac{1}{np_n} X_{i,j}$   $X_{i,j}$  i.i.d. Bernoulli r.v. with mean  $c \frac{\log(n)}{n}$ ,  $c > \frac{2}{2\alpha-1}$ , for  $h^{-1}$  the inverse of  $h(x) = x \log(x) + 1 - x$ ,

$$\mathbb{E}(T_i) \leq \frac{1}{(2\alpha-1)h^{-1}\left(\frac{2}{c(2\alpha-1)}\right)} \log(n) + O(1)$$

- **Path:**  $\mathbb{E}(T_i) \leq \frac{16(1-\alpha)^2}{\pi^2} n^2 \log(n) + O(1)$
- **Ring:**  $\mathbb{E}(T_i) \leq \frac{4(1-\alpha)^2}{\pi^2} n^2 \log(n) + O(1)$ .

## Outline

- 1 Majority consensus: algorithms**
  - Voter Model
  - Averaging Process
  - Binary consensus
- 2 Examples**
  - Complete graph
  - Star, ER, Ring, Line
- 3 Optimization of Majority Consensus**
  - Faster convergence time
  - Implementation
  - Examples

Faster convergence time

**Convex Optimization [Boyd, Diaconis, Xiao '04]**

For technical reasons, let us assume

$$Q = P - I_n$$

where  $P$  is a symmetric stochastic matrix.

**Eigenvalue (convex) optimization**

Minimize the time it takes majority consensus to converge, i.e.

$$\begin{array}{ll} \text{minimize} & \lambda_2(P) = \sup\{x^T P x \mid x^T x = 1, x^T \mathbf{1} = 0\} \\ \text{subject to} & P_{ij} \geq 0, \quad P_{ij} = 0 \text{ if } i, j \notin E \\ & \text{and} \quad \sum_j P_{ij} = 1, \forall i \end{array}$$

## Subgradient method

- Let  $u$  be the eigenvector associated with  $\lambda_2(P)$ .
- Let  $E^\ell$ ,  $\ell = (i, j)$  an edge in the graph such that

$$E_{ij}^\ell = E_{ji}^\ell = 1, \quad E_{ii}^\ell = E_{jj}^\ell = -1$$

- The subgradient of the objective function  $\lambda_2(P)$  is

$$g(P) = \left( u^T E^1 u, \dots, u^T E^m u \right)$$

- In particular, for  $\ell = (i, j)$

$$u^T E^\ell u = (u_i - u_j)^2$$

- To compute eigenvector we could use Lanczos method or recent distributed algorithms [Kempe-McSherry '08].

## Projected subgradient method [Bertsekas '99]

$k \leftarrow 1$

**repeat**

Subgradient Step

Calculate  $g^{(k)}$  and update  $P \leftarrow P - \beta_k g^{(k)}$ ,  $\beta_k$  step size,  $\beta \rightarrow 0$ ,

$\sum_k \beta_k \rightarrow \infty$

Sequential Projection

Projection onto non-negative orthant

$P_\ell \leftarrow \max \{P_\ell, 0\}$ ,  $\ell = 1, \dots, m$

**For each node**  $i = 1, \dots, n$ ,  $\mathcal{L}(i) = \{\ell \mid \text{edge } \ell \text{ connected to } i\}$

Projection onto half-spaces

**While**  $\sum_{\ell \in \mathcal{L}(i)} P_\ell > 1$

$\mathcal{L}(i) \leftarrow \{\ell \mid \ell \in \mathcal{L}(i), P_\ell > 0\}$

$\gamma \leftarrow \min \left\{ \min_{\ell \in \mathcal{L}(i)} P_\ell, \left( \sum_{\ell \in \mathcal{L}(i)} P_\ell - 1 \right) / |\mathcal{L}(i)| \right\}$

$P_\ell \leftarrow P_\ell - \gamma$ ,  $\ell \in \mathcal{L}(i)$

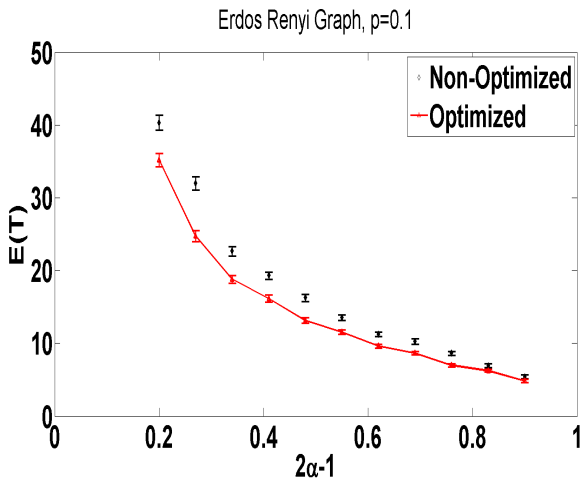
$k \leftarrow k + 1$

It can be implemented in a distributed fashion [Boyd et al '06].



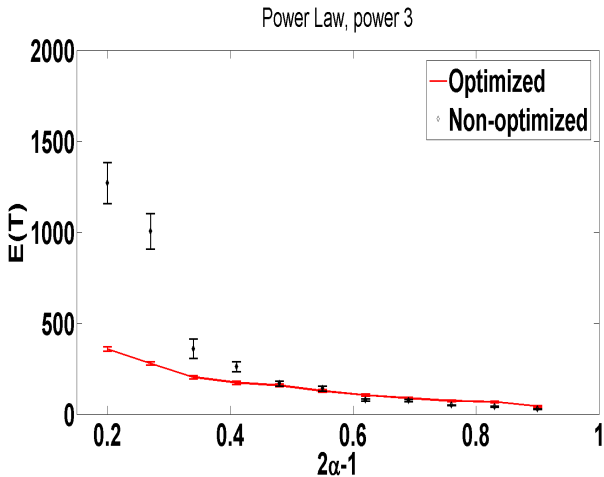
Examples

## ER-graph [Babae, Draief'13+]



Examples

# Preferential attachment [Babaee, Draief'13+]



## Summary

- Algorithms for solving Majority consensus
- Performance: memory, error, time to convergence
- Time to convergence related to spectral properties of rate matrix
- Speedingup convergence via convex optimisation

## Future Work

- Lower-bounds of convergence time
  - O. Ayaso, D. Shah and M. Dahleh, Information Theoretic Bounds for Distributed Computation over Networks of Point-to-Point Channels, IEEE IT, 2010.
  - M. Abdullah, M. Draief, Consensus on the Initial Global Majority by Local Majority Polling for a Class of Sparse Graphs, Arxiv1209.5025, 2013.
- Trade-off between memory, error, time to convergence.
- Distributed spectral computations
  - David Kempe, Frank McSherry: A Decentralized Algorithm for Spectral Analysis, Journal of Computer and System Sciences, 2008.
  - S. Korada, A. Montanari, and S. Oh, Gossip PCA, Sigmetrics 2011.