Congestion and Consensus on non-Complete Graphs

Emanuele Natale

joint work (mainly) with
Luca Becchetti, Andrea Clementi,
Francesco Pasquale and Luca Trevisan

Research Retreat on Graph Analytics
Bertinoro, 9 – 11 December 2015

Summary of the Talk

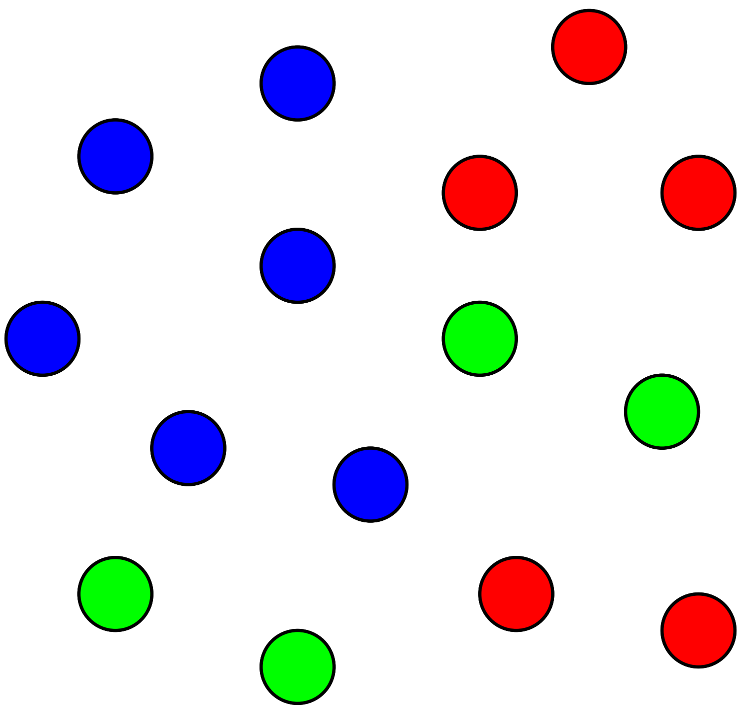
1. Majority Consensus
   (a) 3-Majority (take I)
   (b) Undecided-State
2. Congestion of $\mathcal{GOSSIP}$ random walks
3. Stabilizing Consensus
   (a) 3-Majority (take II)

Part 1: Majority Consensus

1. Majority Consensus
   (a) 3-Majority (take I)
   (b) Undecided-State
2. Congestion of $\mathcal{GOSSIP}$ random walks
3. Stabilizing Consensus
   (a) 3-Majority (take II)

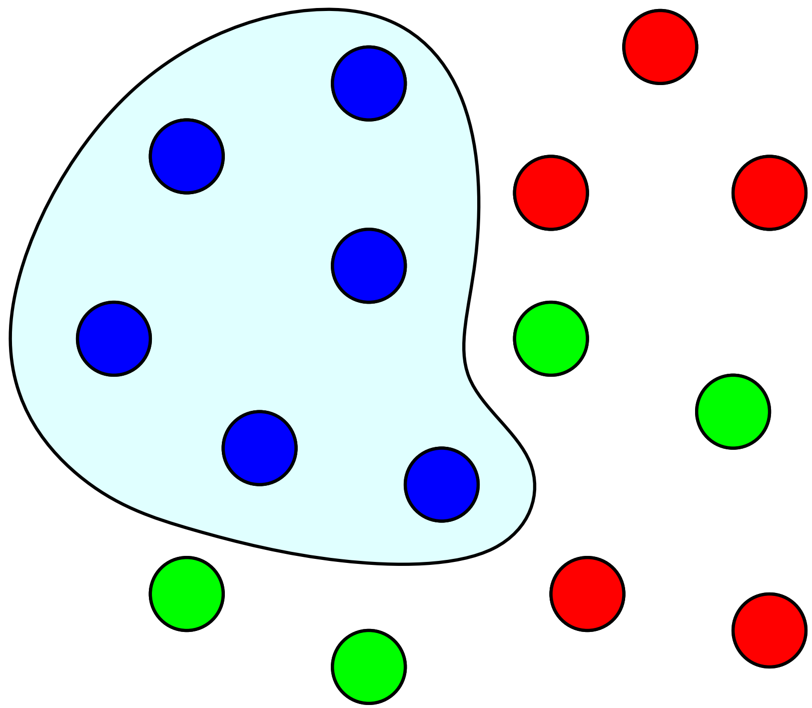
The (Plurality) Consensus Problem

We have a set of nodes each having one color out of $\{1, \ldots, k\}$.



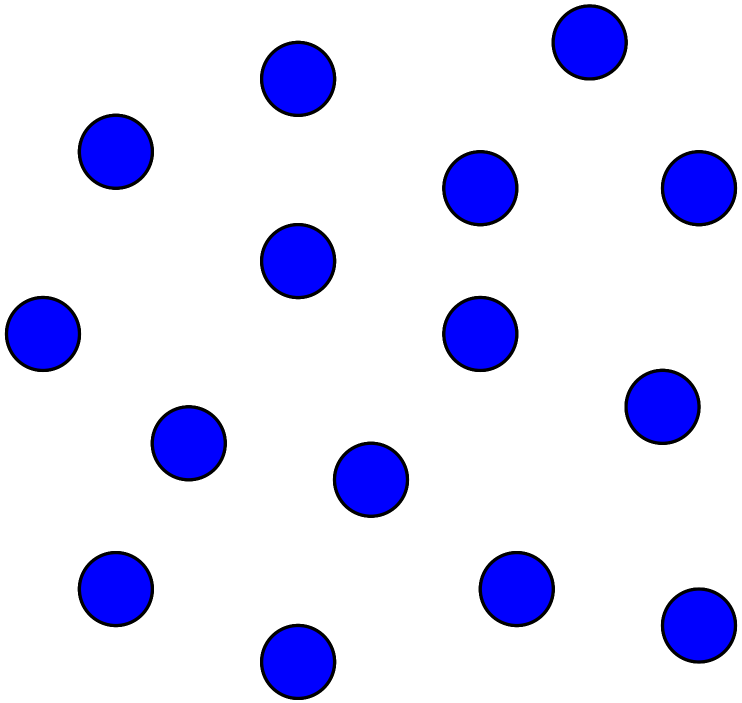
The (Plurality) Consensus Problem

(There is a plurality of nodes having the same color.)



The (Plurality) Consensus Problem

We want to reach consensus (on the plurality color).



Pre-CS History

Voter Model ('70).
Each node with a Poisson clock. When rings, takes the opinion of a random neighbor.

Pre-CS History

Voter Model ('70). Each node with a Poisson clock. When rings, takes the opinion of a random neighbor.

$\Longrightarrow$

Continuos time (sequential/asynchronous) process. Well studied in statistical physics (constant number of particle types).

Pre-CS History

Voter Model ('70). Each node with a Poisson clock. When rings, takes the opinion of a random neighbor.

$\Longrightarrow$

Continuos time (sequential/asynchronous) process. Well studied in statistical physics (constant number of particle types).

Probabilistic Polling (Peleg '01). Time divided in discrete rounds. All nodes *simultaneously* take the opinion of a random neighbor.

Pre-CS History

Voter Model ('70). Each node with a Poisson clock. When rings, takes the opinion of a random neighbor.

$\Longrightarrow$

Continuos time (sequential/asynchronous) process. Well studied in statistical physics (constant number of particle types).

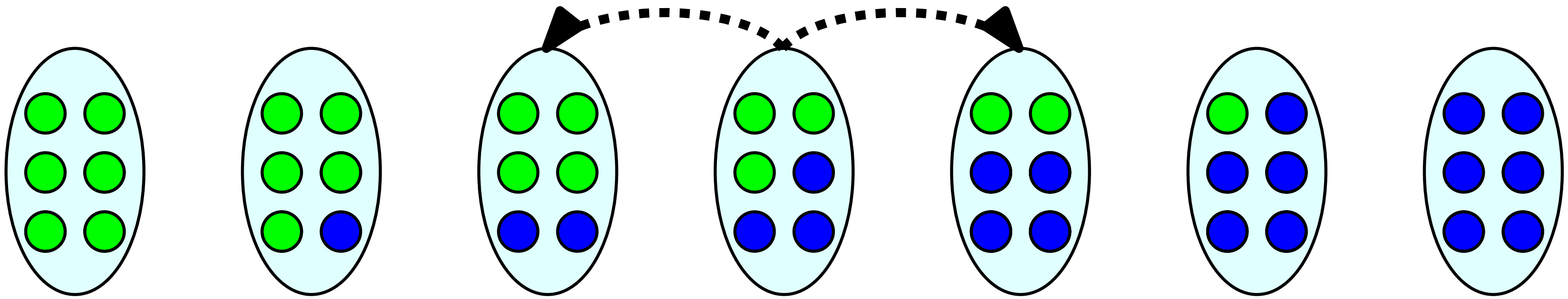
Probabilistic Polling (Peleg '01).

Time divided in discrete rounds. All nodes *simultaneously* take the opinion of a random neighbor.

$\Longrightarrow$

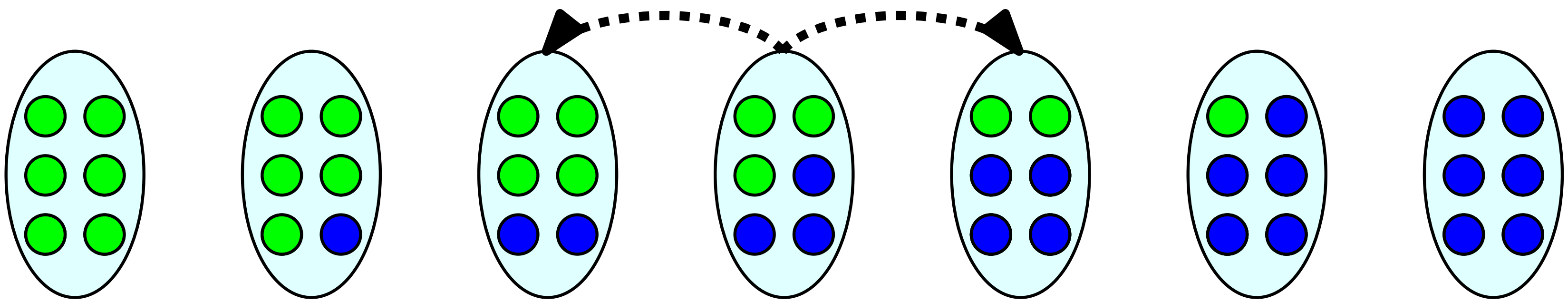
Discrete time (parallel/synchronous) process. Initiated the study of Plurality Consensus in Computer Science.

Asynchronous vs Synchronous

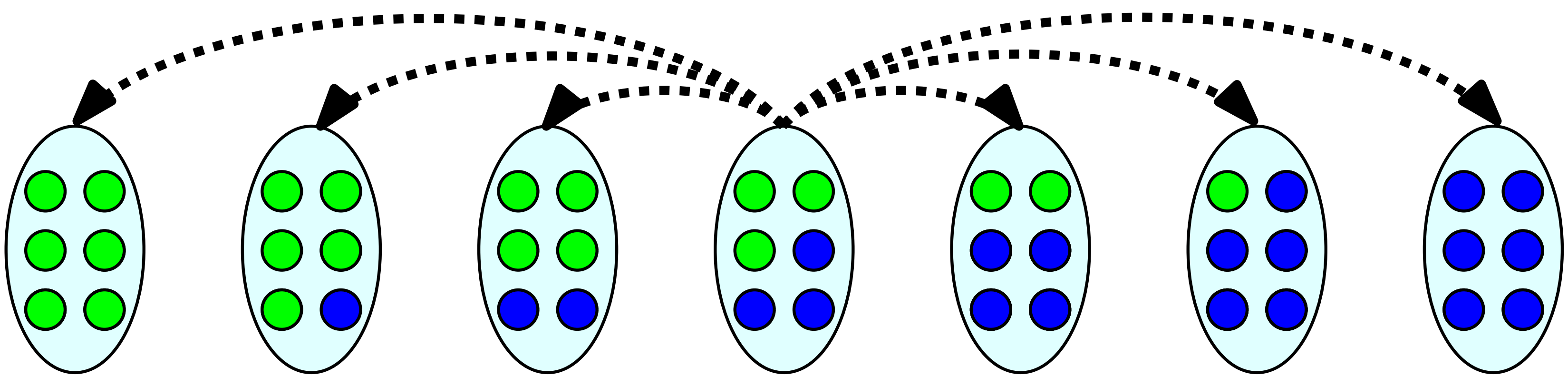


Asynchronous Case

Asynchronous vs Synchronous



Asynchronous Case



Synchronous Case

Our Setting

- **Initial bias**: the plurality is at least $(1 + \epsilon)$ times any other color.

Our Setting

- **Initial bias**: the plurality is at least $(1+\epsilon)$ times any other color.

- **Topology**: complete graph (and regular expanders).

Our Setting

- **Initial bias**: the plurality is at least $(1+\epsilon)$ times any other color.

- **Topology**: complete graph (and regular expanders).

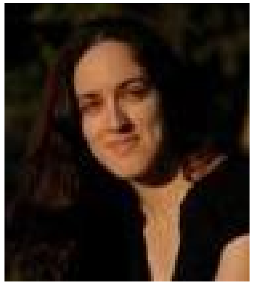
- **Communication model**: $\mathcal{GOSSIP}$ model [Censor-Hillel et al., STOC '12]. Each node in one round can exchange messages with only one neighbor.

Our Setting

- **Initial bias**: the plurality is at least $(1+\epsilon)$ times any other color.

- **Topology**: complete graph (and regular expanders).

- **Communication model**: $\mathcal{GOSSIP}$ model [Censor-Hillel et al., STOC '12]. Each node in one round can exchange messages with only one neighbor.



- **Local memory and message size**: $O(\log n)$.

Relationships to Other Communication Models

$\mathcal{GOSSIP}$ model with neighbors chosen randomly: Telephone Call, Push&Pull, Uniform Gossip…

Relationships to Other Communication Models

$\mathcal{GOSSIP}$ model with neighbors chosen randomly: Telephone Call, Push&Pull, Uniform Gossip…

$\mathcal{LOCAL}$ model [Peleg, SIAM '00]: each node in one round can exchange messages with all its neighbors.

Relationships to Other Communication Models

$\mathcal{GOSSIP}$ model with neighbors chosen randomly: Telephone Call, Push&Pull, Uniform Gossip…

$\mathcal{LOCAL}$ model [Peleg, SIAM '00]: each node in one round can exchange messages with all its neighbors.
…on the complete graph, plurality consensus can be achieved in one round.

Relationships to Other Communication Models

$\mathcal{GOSSIP}$ model with neighbors chosen randomly: Telephone Call, Push&Pull, Uniform Gossip…

$\mathcal{LOCAL}$ model [Peleg, SIAM '00]: each node in one round can exchange messages with all its neighbors.
…on the complete graph, plurality consensus can be achieved in one round.

Censor-Hillel et al. (STOC '12):
Every task that can be solved in the $\mathcal{LOCAL}$ model in T rounds, can be solved in $O(T + \text{polylog} n)$ rounds in the $\mathcal{GOSSIP}$ model.
But…

Relationships to Other Communication Models

$\mathcal{GOSSIP}$ model with neighbors chosen randomly:
Telephone Call, Push&Pull, Uniform Gossip...

$\mathcal{LOCAL}$ model [Peleg, SIAM '00]: each node in one round can exchange messages with all its neighbors.
...on the complete graph, plurality consensus can be achieved in one round.

Censor-Hillel et al. (STOC '12):
Every task that can be solved in the $\mathcal{LOCAL}$ model in T rounds, can be solved in $O(T + \text{polylog} n)$ rounds in the $\mathcal{GOSSIP}$ model.
But...using the preceding theorem, message size grows dramatically!

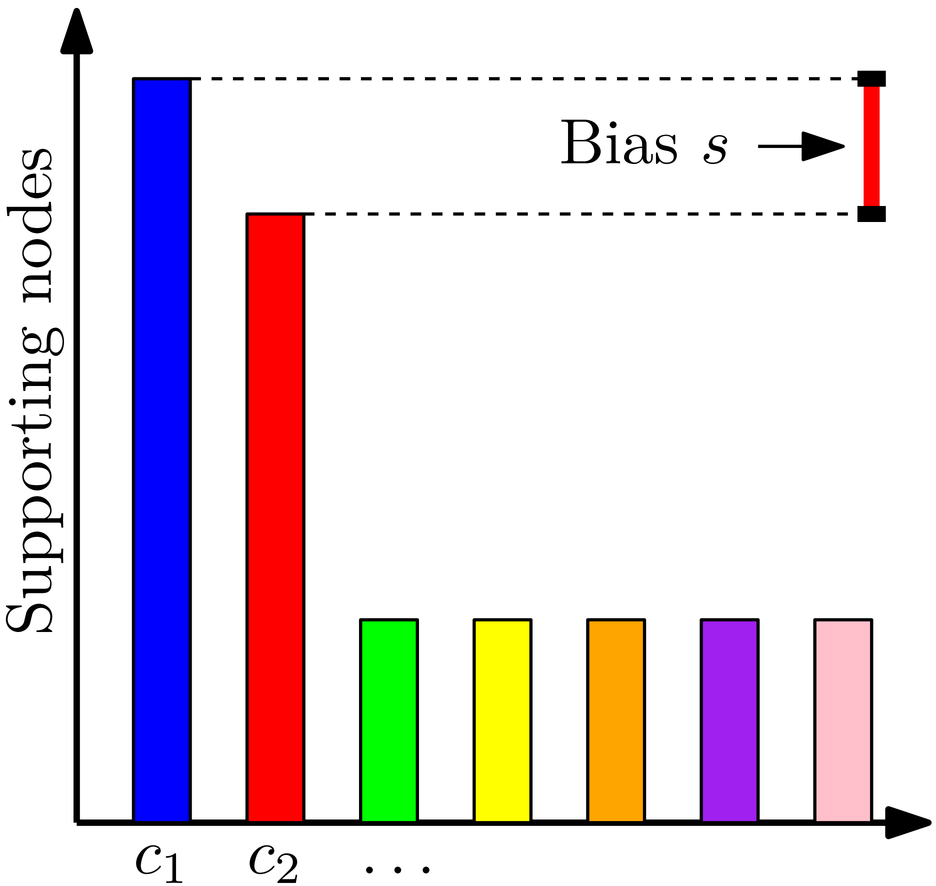
(Some) Related Works

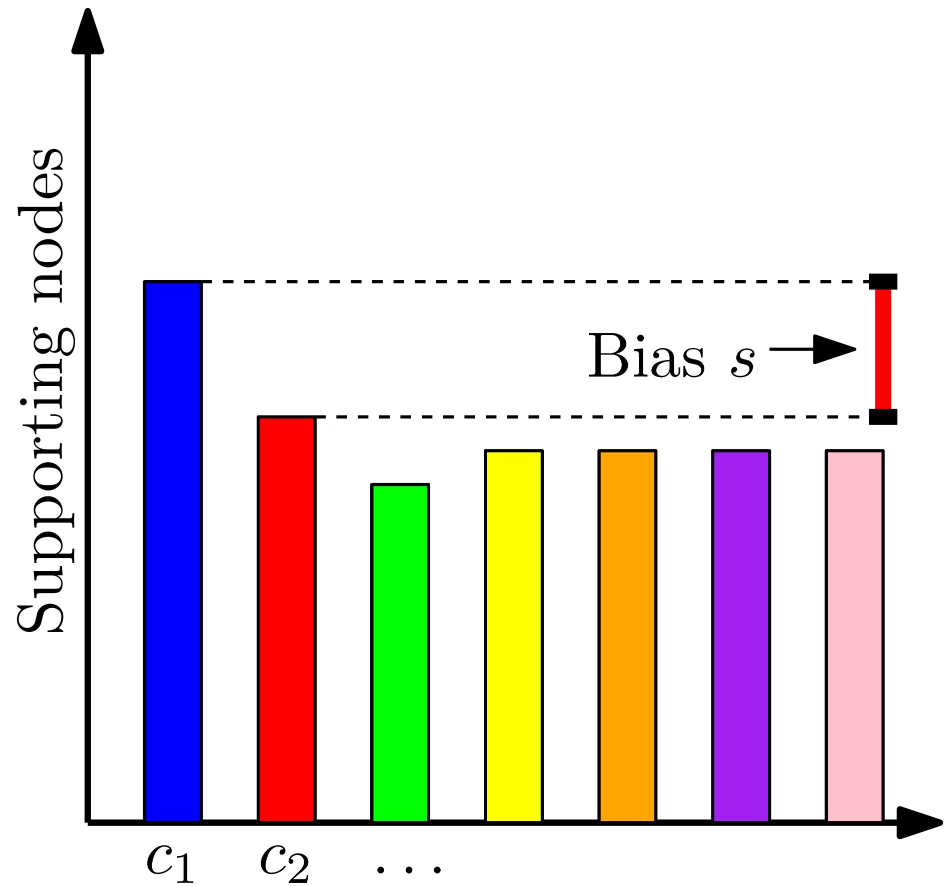
	Mem. & mess. size	**# of colors**	**Time efficiency**	**Comm. Model**
Kempe et al. FOCS '03	$O(k \log n)$ ✗	any ✓	$O(\log n)$ ✓	$\mathcal{GOSSIP}$ ✓
Angluin et al. DISC '07 Perron et al. INFOCOM '09	$\Theta(1)$ ✓	2 ✗	$O(\log n)$ ✓	Sequential ✗
Doerr et al. SPAA '11	$\Theta(1)$ ✓	2 ✗	$O(\log n)$ ✓	$\mathcal{GOSSIP}$ ✓
Babaee et al. Comp. J. '12 Jung et al. ISIT '12	$O(\log k)$ ✓	Constant ✗	$O(\log n)$ ✓	Sequential ✗

Characterizing the Initial Bias

$c_i^{(t)} := |\{i\text{-colored nodes}\}|$

color 1 is the plurality

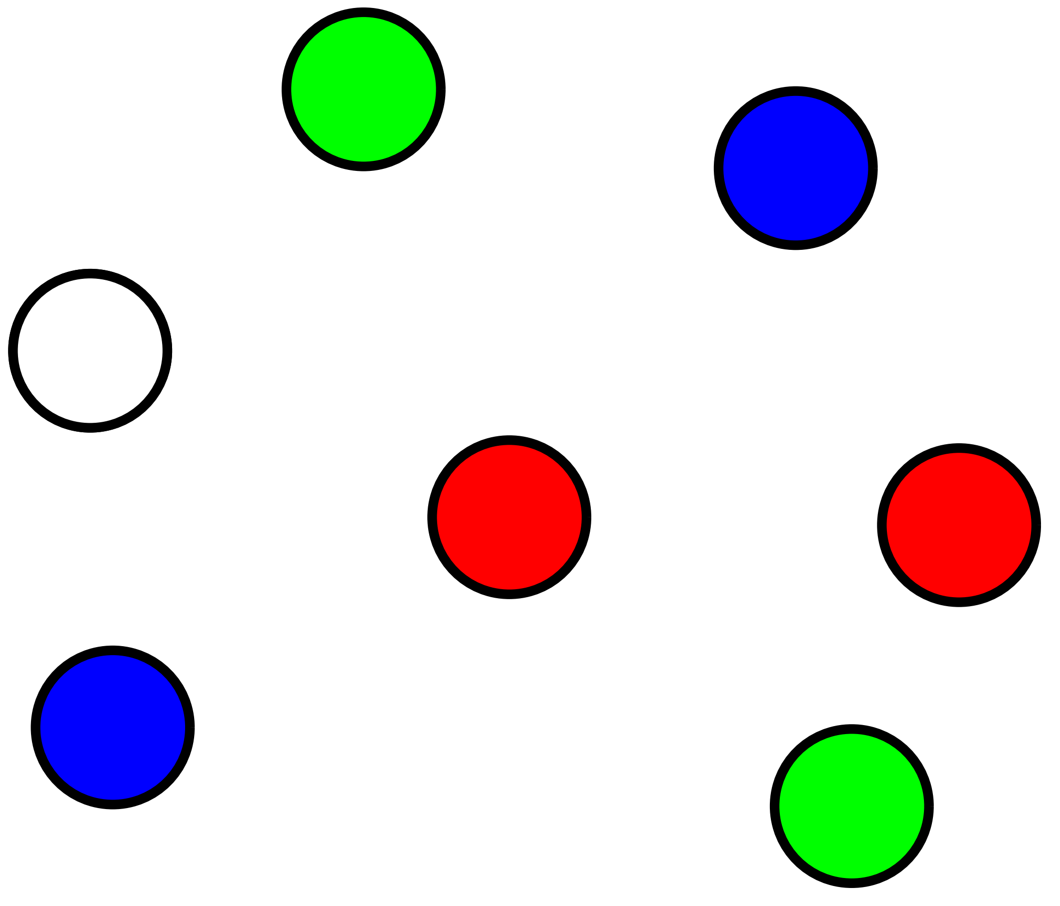




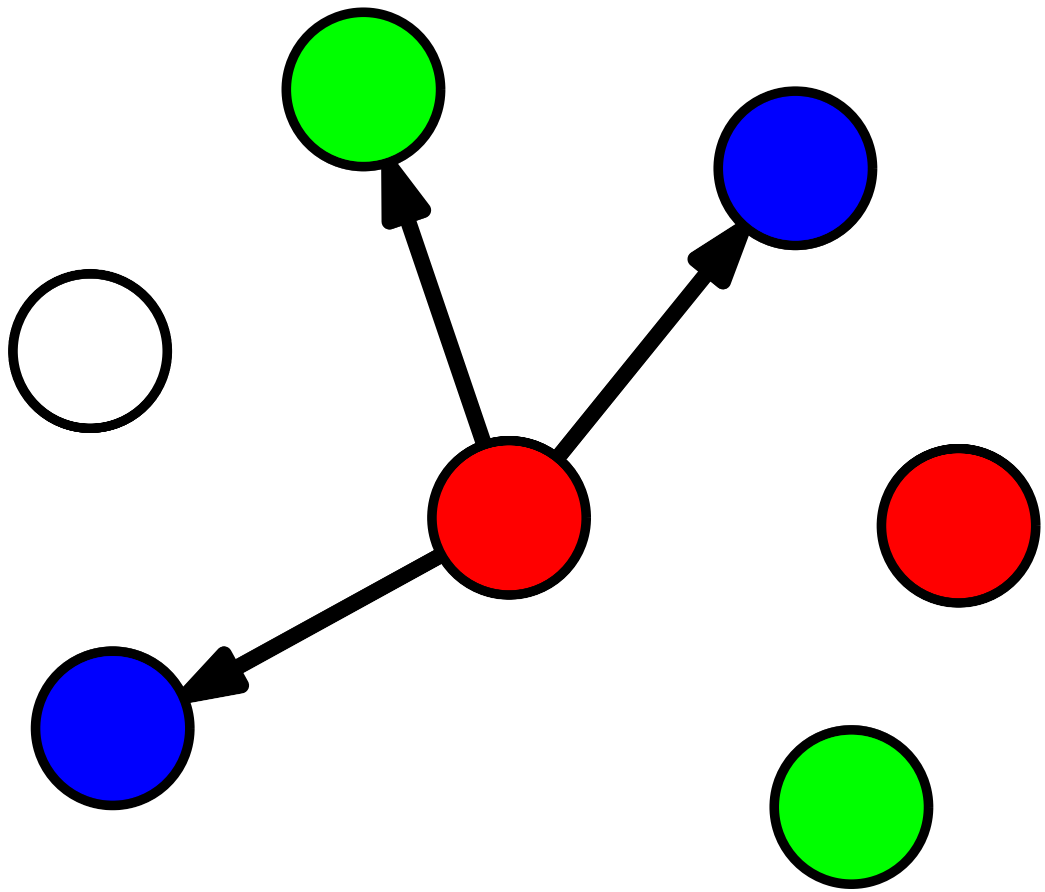
Part 1-a: 3-Majority (take I)

1. Majority Consensus
   (a) 3-Majority (take I)
   (b) Undecided-State
2. Congestion of $\mathcal{GOSSIP}$ random walks
3. Stabilizing Consensus
   (a) 3-Majority (take II)

The 3-Majority Dynamics

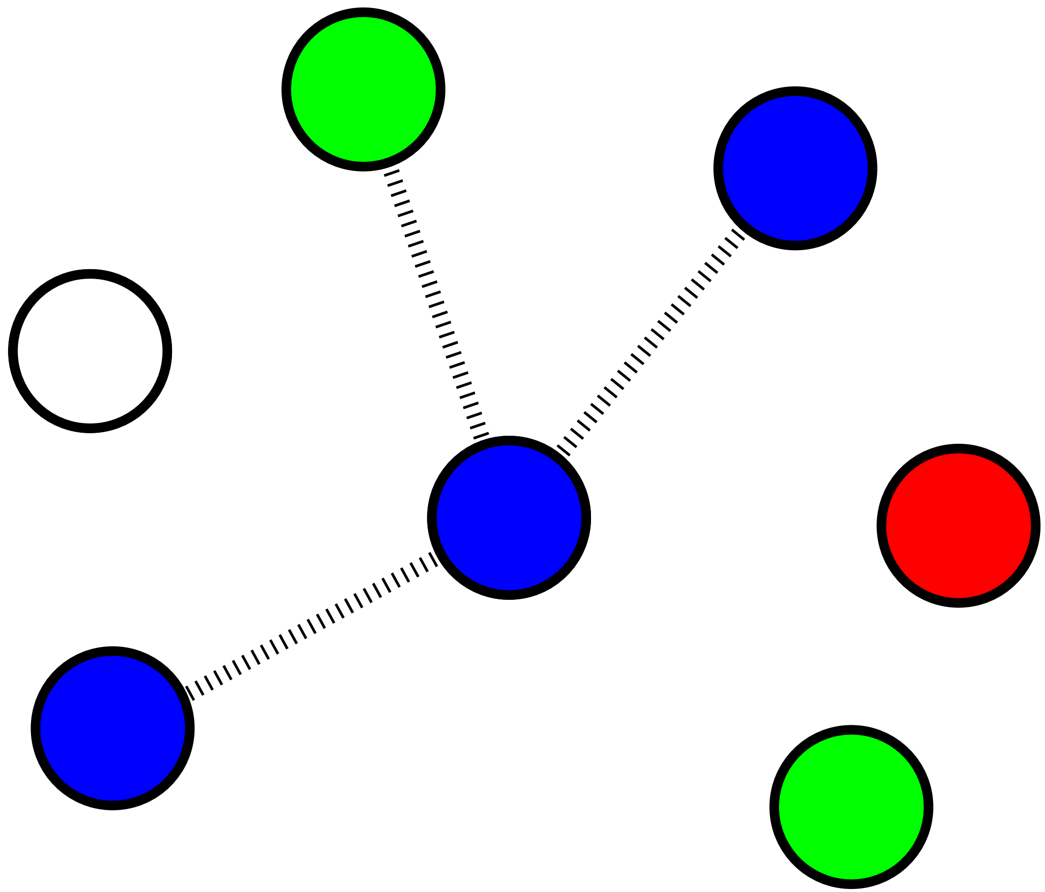


The 3-Majority Dynamics



Each node observes the color of three other nodes chosen u.a.r....

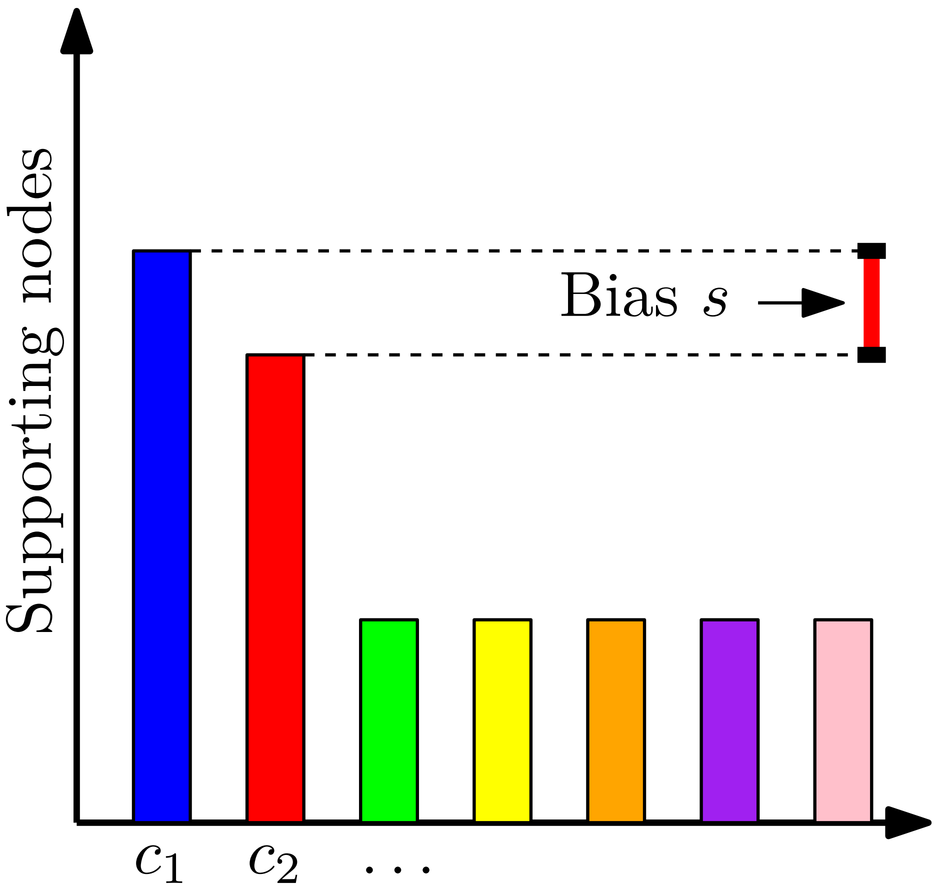
The 3-Majority Dynamics



...and changes its color according to the majority of these three (breaking ties u.a.r.).

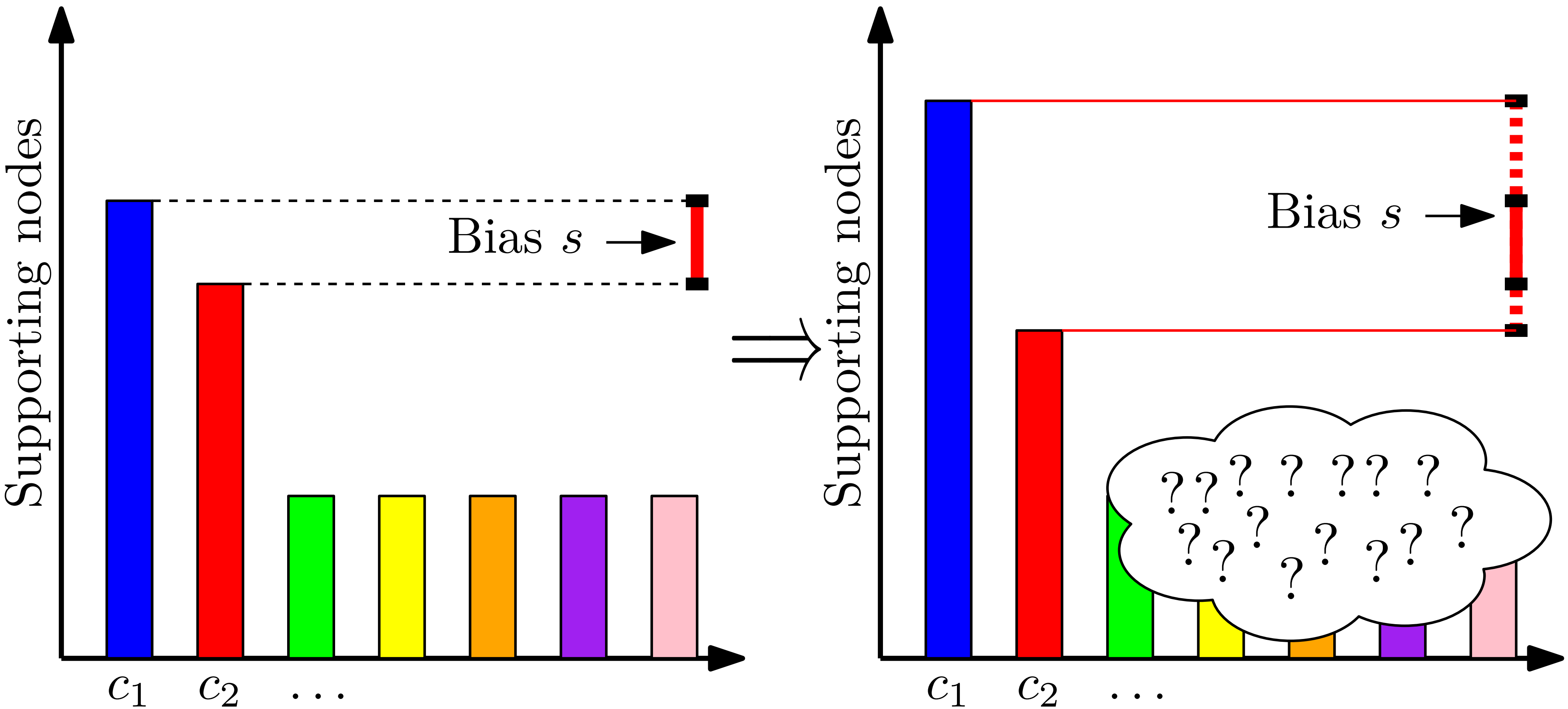
Upper Bound for the 3-Majority

$c_i^{(t)} := |\{i\text{-colored nodes}\}|$ color 1 is the plurality



Upper Bound for the 3-Majority

$c_i^{(t)} := |\{i\text{-colored nodes}\}|$ color 1 is the plurality



Upper Bound for the 3-Majority

$C_i^{(t)}$:= number of nodes supporting opinion i at round t.

$\mu_j(\mathbf{c}) = \mathbf{E}[C_j^{(t+1)} \mid \mathbf{C}^{(t)} = \mathbf{c}]$

Lemma 1. For any opinion j it holds

$$\mu_j(\mathbf{c}) = c_j \left(1 + \frac{c_j}{n} - \frac{1}{n^2} \sum_{h \in [k]} c_h^2\right).$$

Lemma 2. Let 1 be the plurality opinion, then

$$\mu_1 - \mu_j \geqslant s(\mathbf{c}) \left(1 + \frac{c_1}{n} \left(1 - \frac{c_1}{n}\right)\right).$$

Upper Bound for the 3-Majority

Lemma 1. For any opinion j it holds

$$\mu_j(\mathbf{c}) = c_j\left(1 + \frac{c_j}{n} - \frac{1}{n^2}\sum_{h\in[k]} c_h^2\right).$$

Proof.

$$\begin{aligned}
&P\,(\text{A node chooses color } j) \\
&= \left(\frac{c_j}{n}\right)^3 + 3\left(\frac{c_j}{n}\right)^2\left(\frac{n-c_j}{n}\right) \\
&\quad + \left(\frac{c_j}{n}\right)\left[1 - \left(\frac{\sum_{h=1}^k c_h^2}{n^2} + 2\left(\frac{c_j}{n}\right)\left(\frac{n-c_j}{n}\right)\right)\right] \\
&= c_j\left(1 + \frac{1}{n^2}\left(nc_j - \sum_{h\in[k]} c_h^2\right)\right).
\end{aligned}$$

Upper Bound for the 3-Majority

Lemma 2. Let 1 be the plurality opinion, then

$$\mu_1 - \mu_j \geqslant s(\mathbf{c}) \left(1 + \frac{c_1}{n}\left(1 - \frac{c_1}{n}\right)\right).$$

Proof.

$$\begin{aligned}
\mu_1 - \mu_j \geqslant \mu_1 - \mu_2 &= (c_1 - c_2) + \frac{\left(c_1^2 - c_2^2\right)}{n} - \frac{c_1 - c_2}{n^2}\sum_{h \in k} c_h^2 \\
&= s(\mathbf{c}) \left(1 + \frac{c_1 + c_2}{n} - \frac{1}{n^2}\sum_{h \in k} c_h^2\right) \\
&\geqslant s(\mathbf{c}) \left(1 + \frac{c_1 + c_2}{n} - \frac{c_1^2 + n c_2}{n^2}\right) \\
&= s(\mathbf{c}) \left(1 + \frac{c_1}{n}\left(1 - \frac{c_1}{n}\right)\right).
\end{aligned}$$

Convergence of 3-Majority [SPAA '14]

Theorem. From any configuration with $k < \sqrt[3]{n}$ colors, such that

$$s \geq 22\sqrt{2kn\log n},$$

the 3-majority protocol converges to the majority opinion in $O(2k\log n)$ rounds w.h.p., even in the presence of a $O(\sqrt{n})$-bounded dynamic adversary.

Proof. Plurality is preserved and the gap between plurality and others increses.

Convergence of 3-Majority [SPAA '14]

Theorem. From any configuration with $k < \sqrt[3]{n}$ colors, such that

$$s \geq 22\sqrt{2kn \log n},$$

the 3-majority protocol converges to the majority opinion in $O(2k \log n)$ rounds w.h.p., even in the presence of a $O(\sqrt{n})$-bounded dynamic adversary.

	Mem. & mess. size	**# of colors**	**Time efficiency**	**Comm. Model**
SPAA '14	$O(\log k)$	$n^{\Theta(1)}$	$O(k \log n)$	$\mathcal{GOSSIP}$

Convergence of 3-Majority [SPAA '14]

Theorem. From any configuration with $k < \sqrt[3]{n}$ colors, such that

$$s \geq 22\sqrt{2kn\log n},$$

the 3-majority protocol converges to the majority opinion in $O(2k\log n)$ rounds w.h.p., even in the presence of a $O(\sqrt{n})$-bounded dynamic adversary.

	Mem. & mess. size	**# of colors**	**Time efficiency**	**Comm. Model**
SPAA '14	$O(\log k)$ ✓	$n^{\Theta(1)}$ ✓	$O(k\log n)$	$\mathcal{GOSSIP}$ ✓

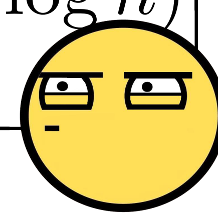
Convergence of 3-Majority [SPAA '14]

Theorem. From any configuration with $k < \sqrt[3]{n}$ colors, such that

$$s \geq 22\sqrt{2kn\log n},$$

the 3-majority protocol converges to the majority opinion in $O(2k\log n)$ rounds w.h.p., even in the presence of a $O(\sqrt{n})$-bounded dynamic adversary.

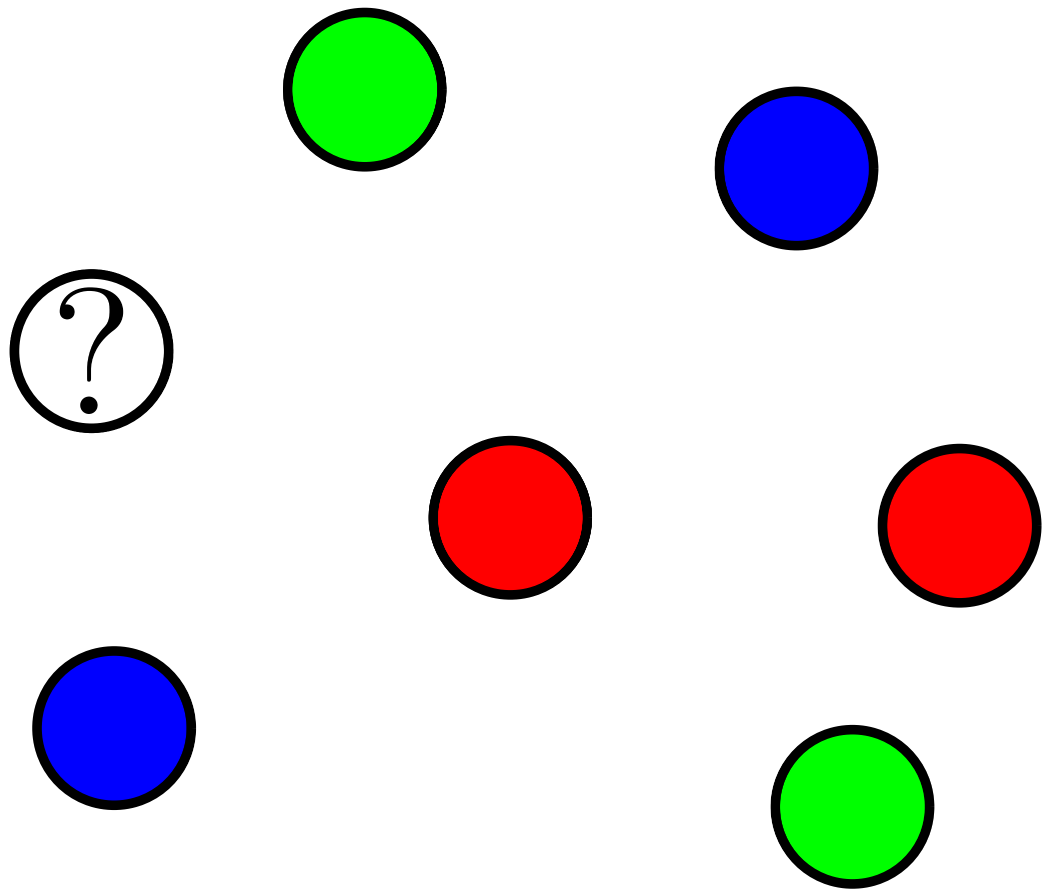
	Mem. & mess. size	**# of colors**	**Time efficiency**	**Comm. Model**
SPAA '14	$O(\log k)$	$n^{\Theta(1)}$	$O(k\log n)$	$\mathcal{GOSSIP}$



Part 1-b: Undecided-State

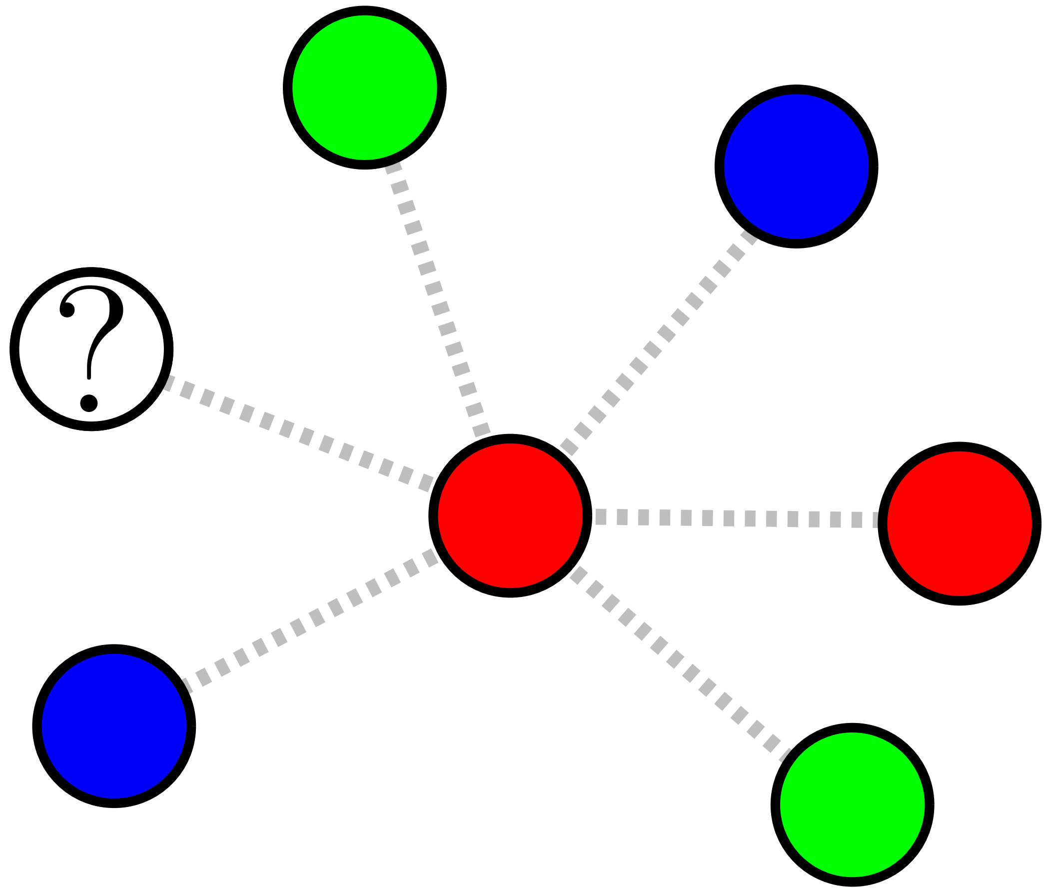
1. Majority Consensus
   (a) 3-Majority (take I)
   (b) Undecided-State
2. Congestion of $\mathcal{GOSSIP}$ random walks
3. Stabilizing Consensus
   (a) 3-Majority (take II)

The Undecided-State Dynamics



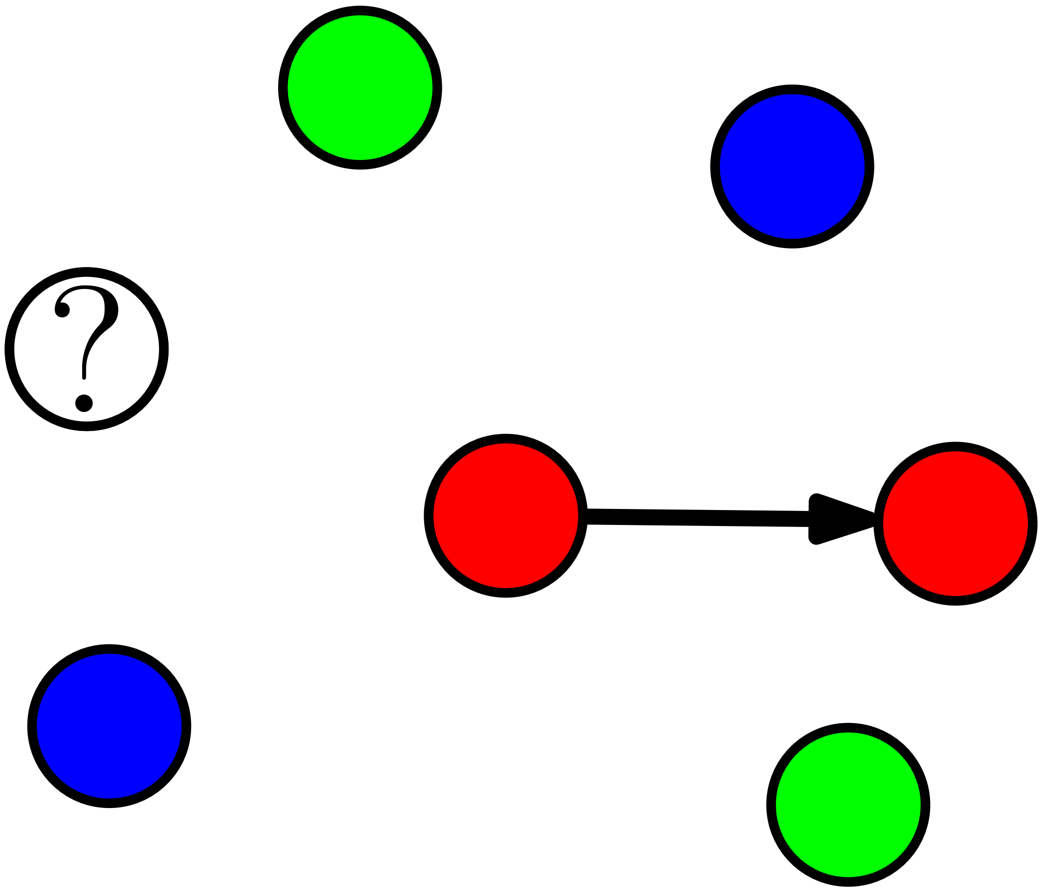
Some nodes can be "undecided".

The Undecided-State Dynamics



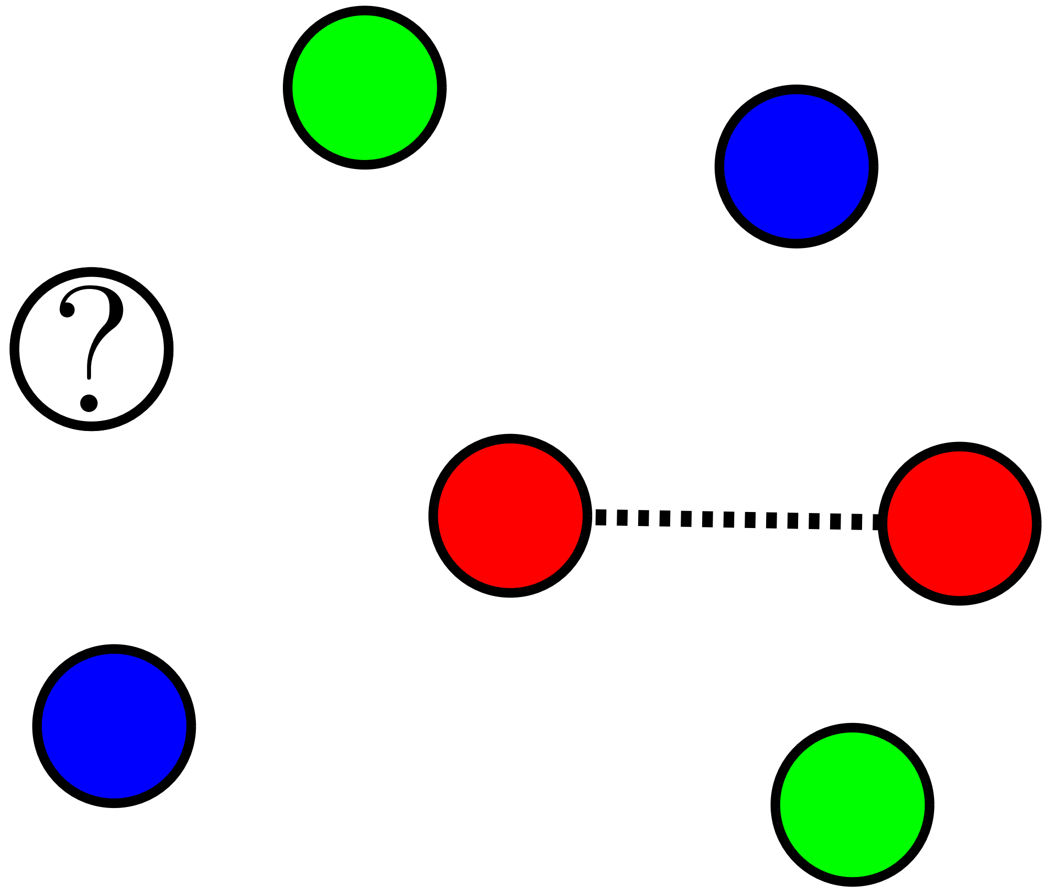
At the beginning of each round, each node observes a neighbor picked uniformly at random.

The Undecided-State Dynamics



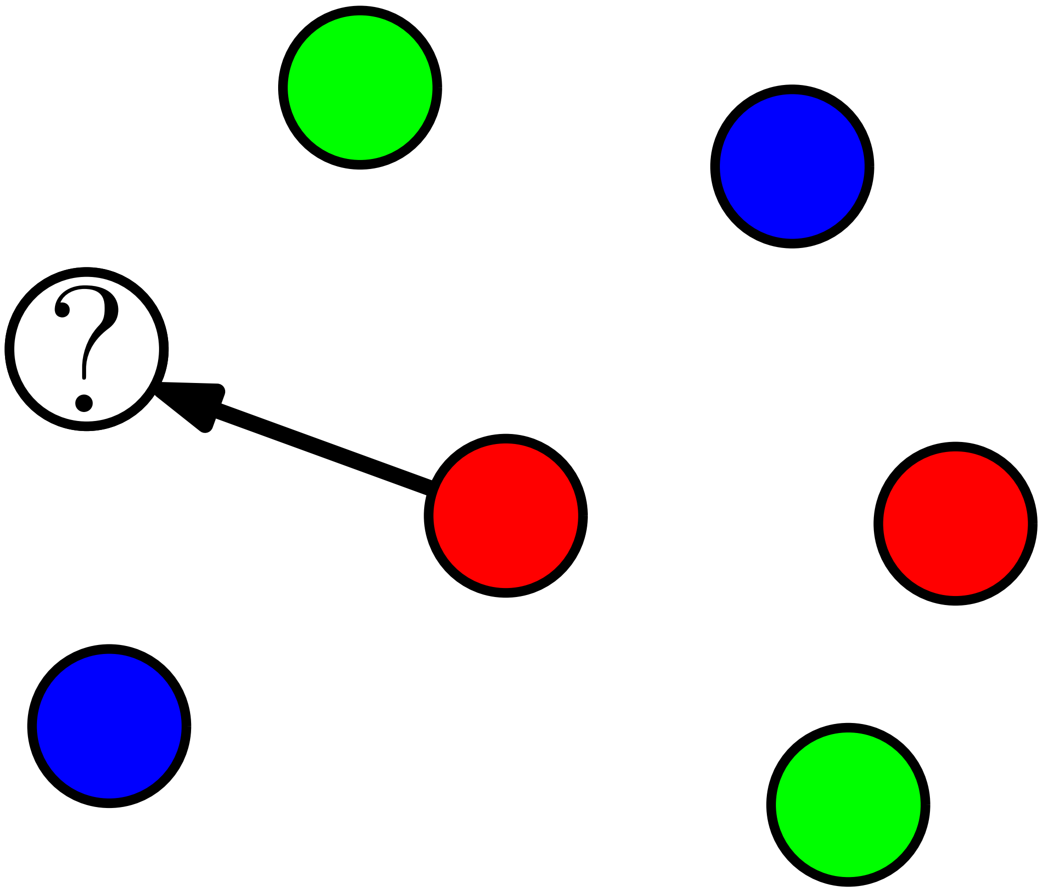
If the observed node shares the same color. . .

The Undecided-State Dynamics



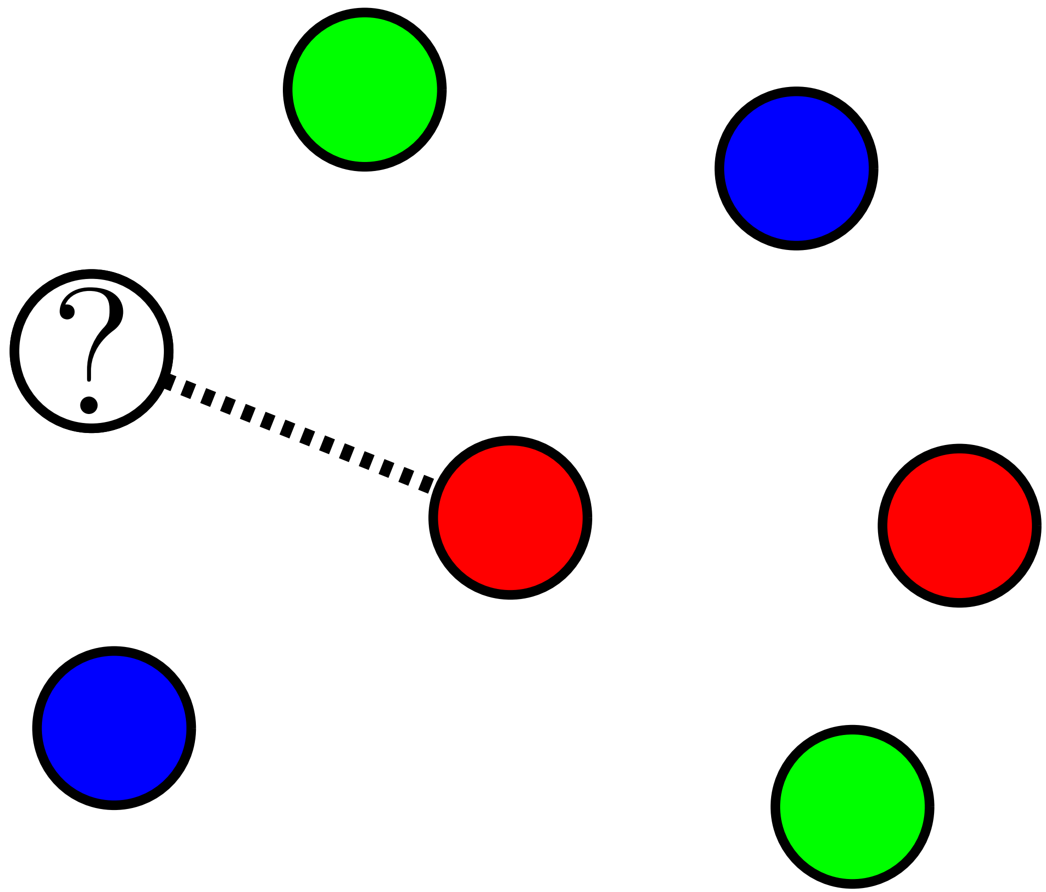
…nothing happens;

The Undecided-State Dynamics



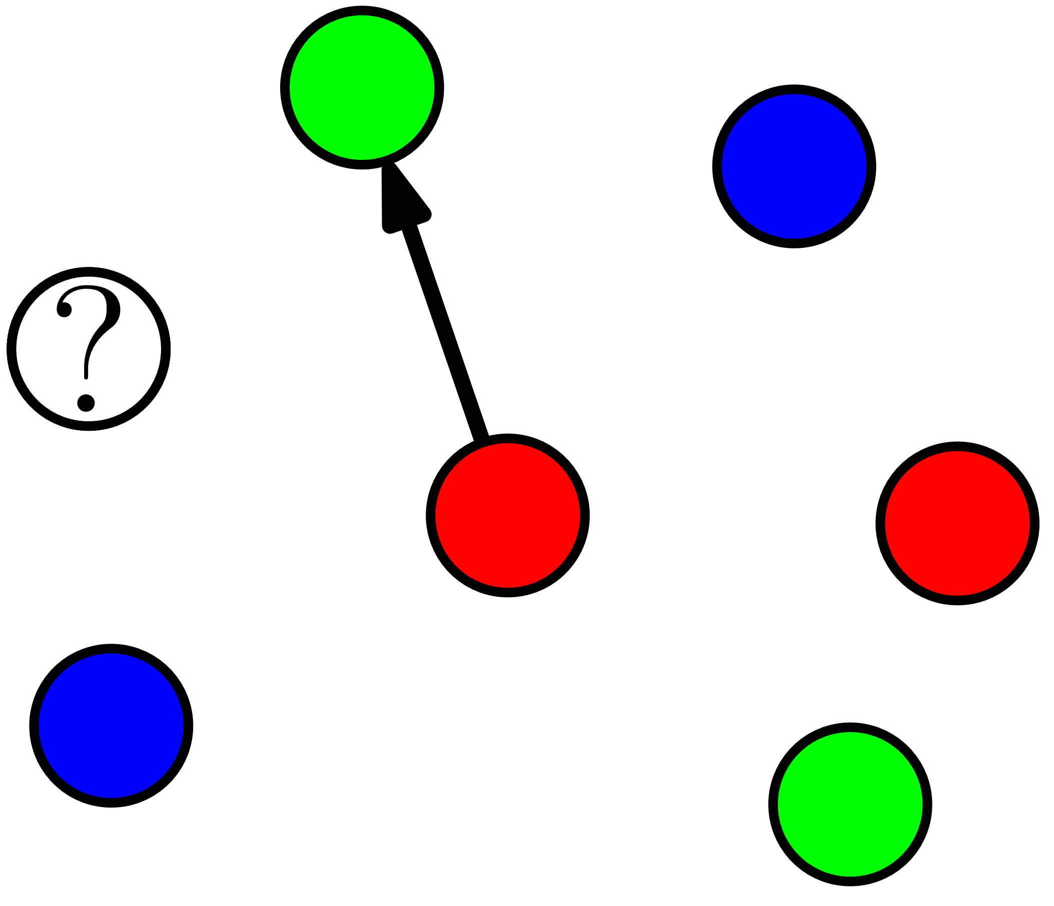
if the node observes an undecided one…

The Undecided-State Dynamics



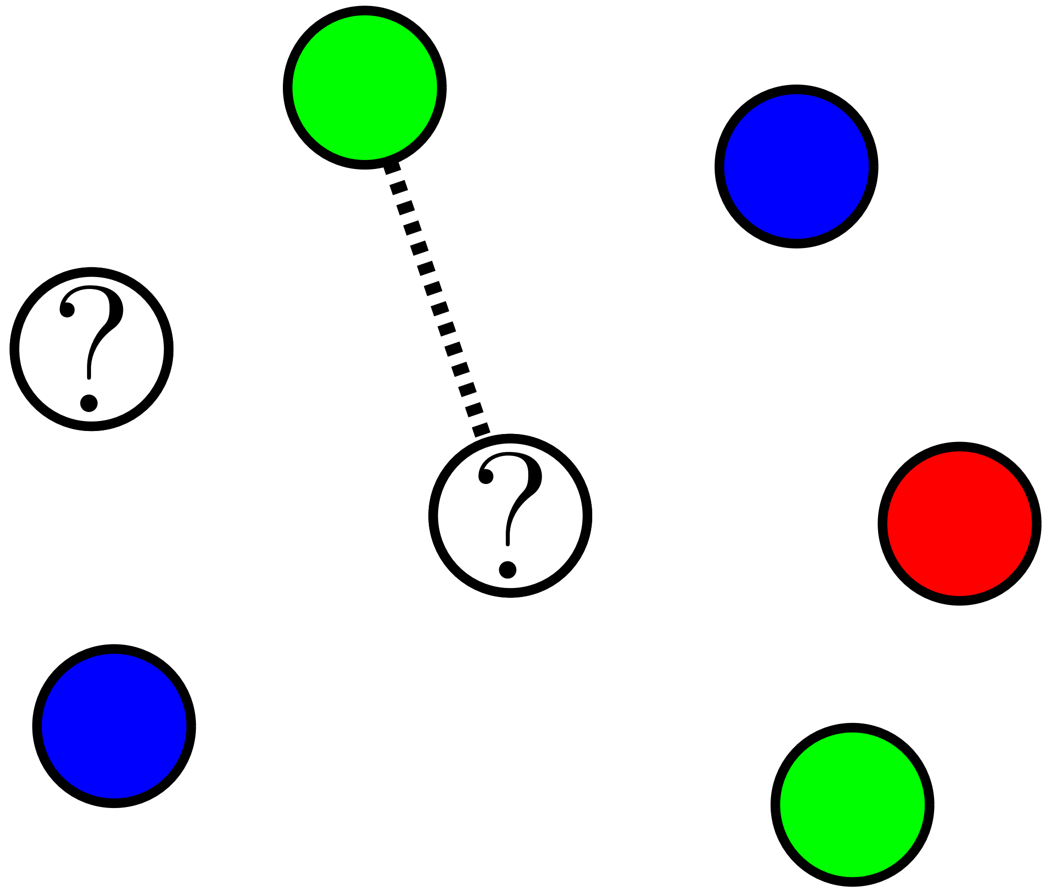
. . . nothing happens too;

The Undecided-State Dynamics



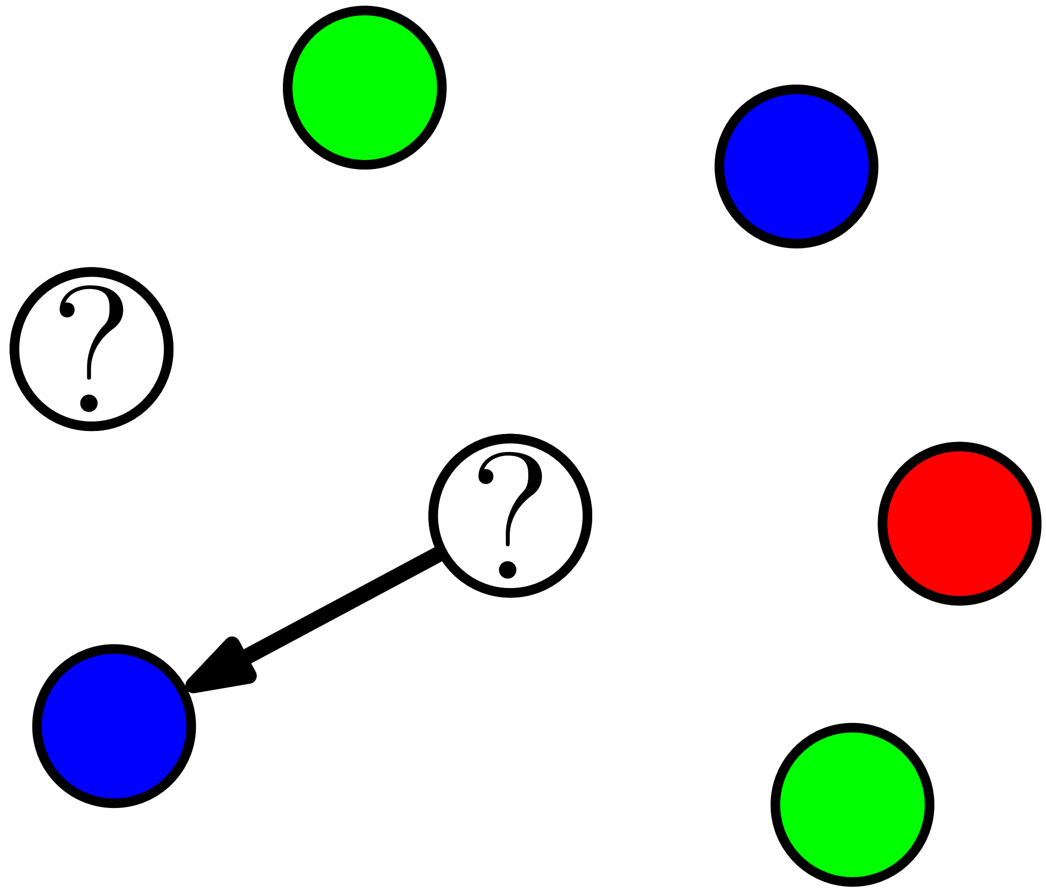
but, if the observed node has a different color…

The Undecided-State Dynamics



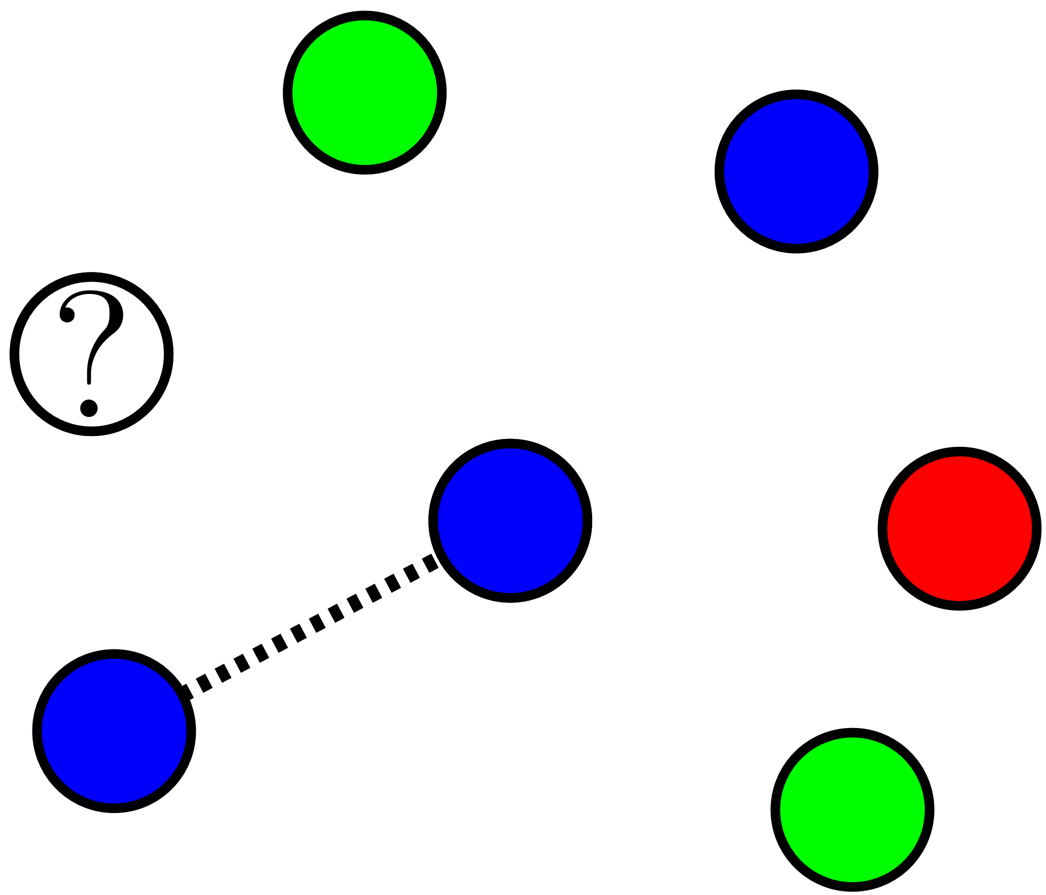
. . . then the node becomes undecided.

The Undecided-State Dynamics



Once undecided. . .

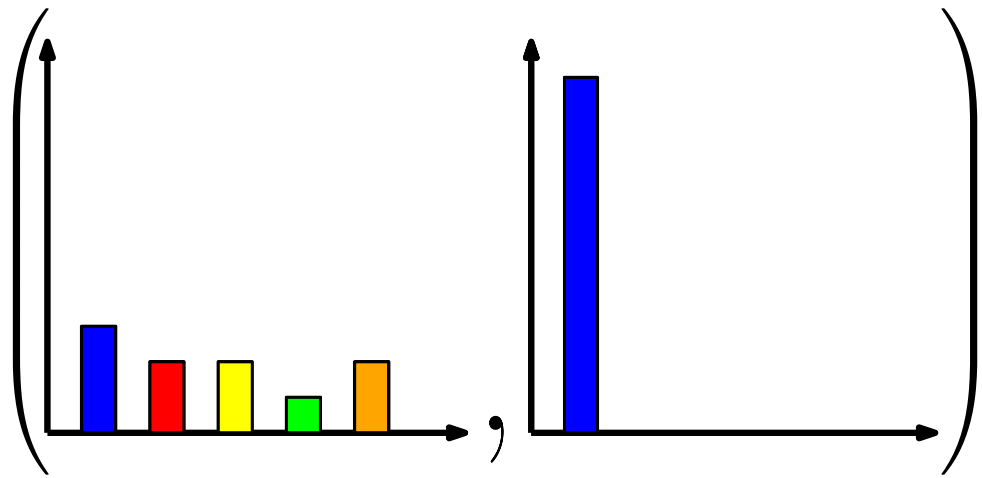
The Undecided-State Dynamics



. . . the node copies the first color it sees.

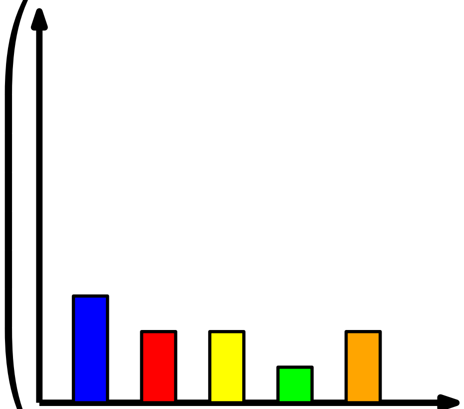
The Monochromatic Distance

$$\mathrm{md}(\mathbf{c}^{(\mathbf{0})}) := \sum_{i=1}^{k} \left(\frac{c_i^{(0)}}{c_1^{(0)}} \right)^2 = 1 + \mathcal{D}$$

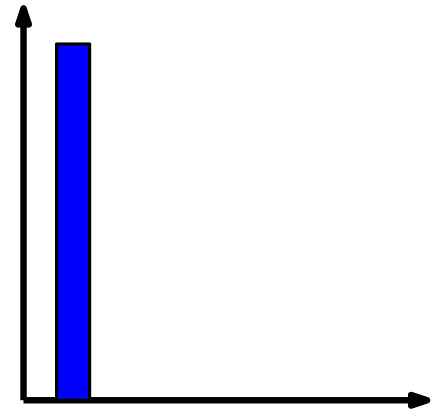


The Monochromatic Distance

$\mathrm{md}(\mathbf{c}^{(\mathbf{0})}) := \sum_{i=1}^{k} \left(\frac{c_i^{(0)}}{c_1^{(0)}} \right)^2 = 1 + \mathcal{D}\Bigg($

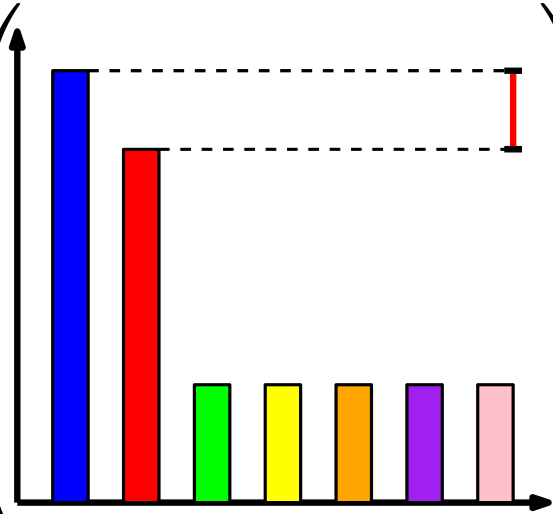


,

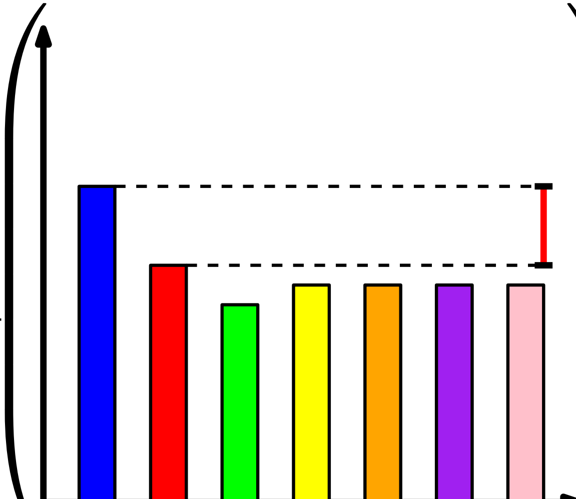


$\Bigg)$

$1 \leq \mathrm{md}\Bigg($



$\Bigg) \ll \mathrm{md}\Bigg($



$\Bigg) \leq k$

Convergence of the Undecided-State [SODA '15]

First analysis for $k = \omega(1)$ of the Undecided-State Dynamics (Angluin et al., Perron et al., Babaee et al., Jung et al.).

Theorem.
If $k = O\left((n/\log n)^{1/3}\right)$ and $c_1 \geq (1+\epsilon) \cdot c_2$ with $\epsilon > 0$, then w.h.p. the Undecided-State Dynamics reaches plurality consensus in $O\left(\mathrm{md}(\mathbf{c}^{(0)}) \cdot \log n\right) \cap \Omega(\mathrm{md}(\mathbf{c}^{(0)}))$ rounds.

Extension to d-Regular Expanders

Theorem
Given a d-regular expander graph,
$k = O\left((n/\log n)^{1/3}\right)$ and $c_1 \geq (1+\epsilon) \cdot c_2$ with $\epsilon > 0$,
using polylogarithmic memory and message size the plurality consensus problem can be solved in w.h.p. $O(\mathrm{md}(\mathbf{c})\mathrm{polylog}(n))$ rounds.

Extension to d-Regular Expanders

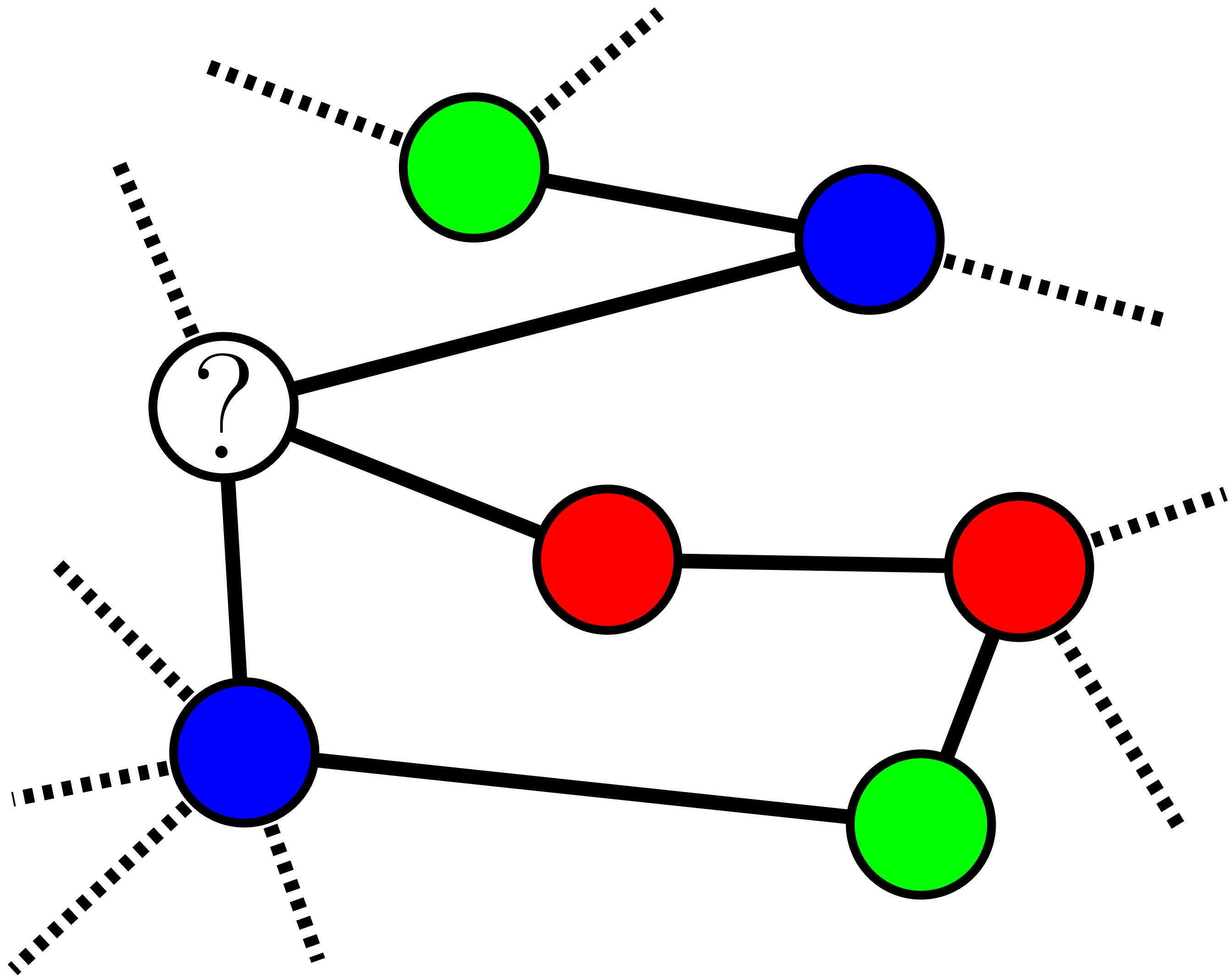
Theorem

Given a d-regular expander graph,
$k = O\left((n/\log n)^{1/3}\right)$ and $c_1 \geq (1+\epsilon) \cdot c_2$ with $\epsilon > 0$, using polylogarithmic memory and message size the plurality consensus problem can be solved in w.h.p. $O(\mathrm{md}(\mathbf{c})\mathrm{polylog}(n))$ rounds.

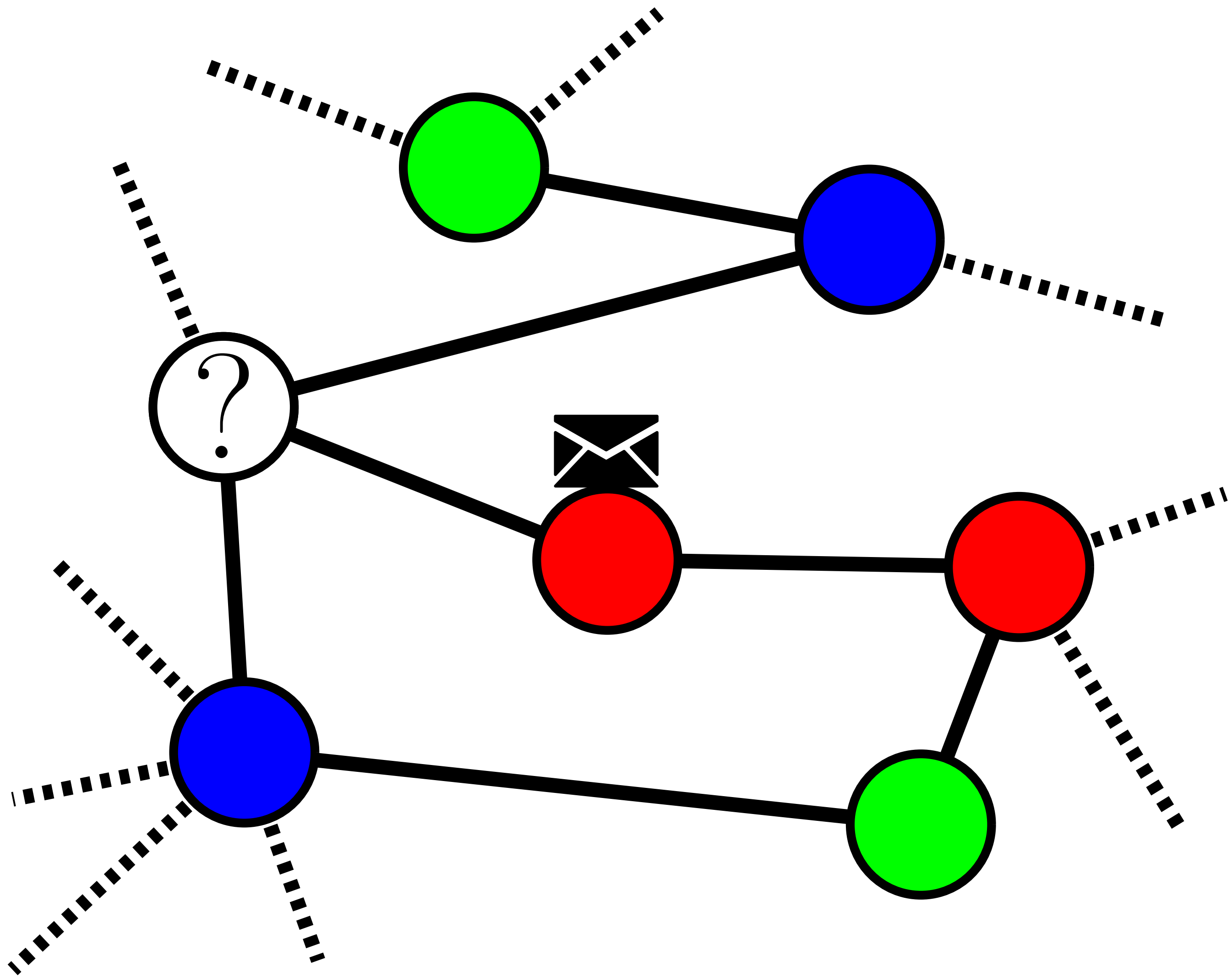
Idea

Simulate Undecided-State Dynamics on complete graph by sampling via n parallel random walks. (Rapidly mixing property)

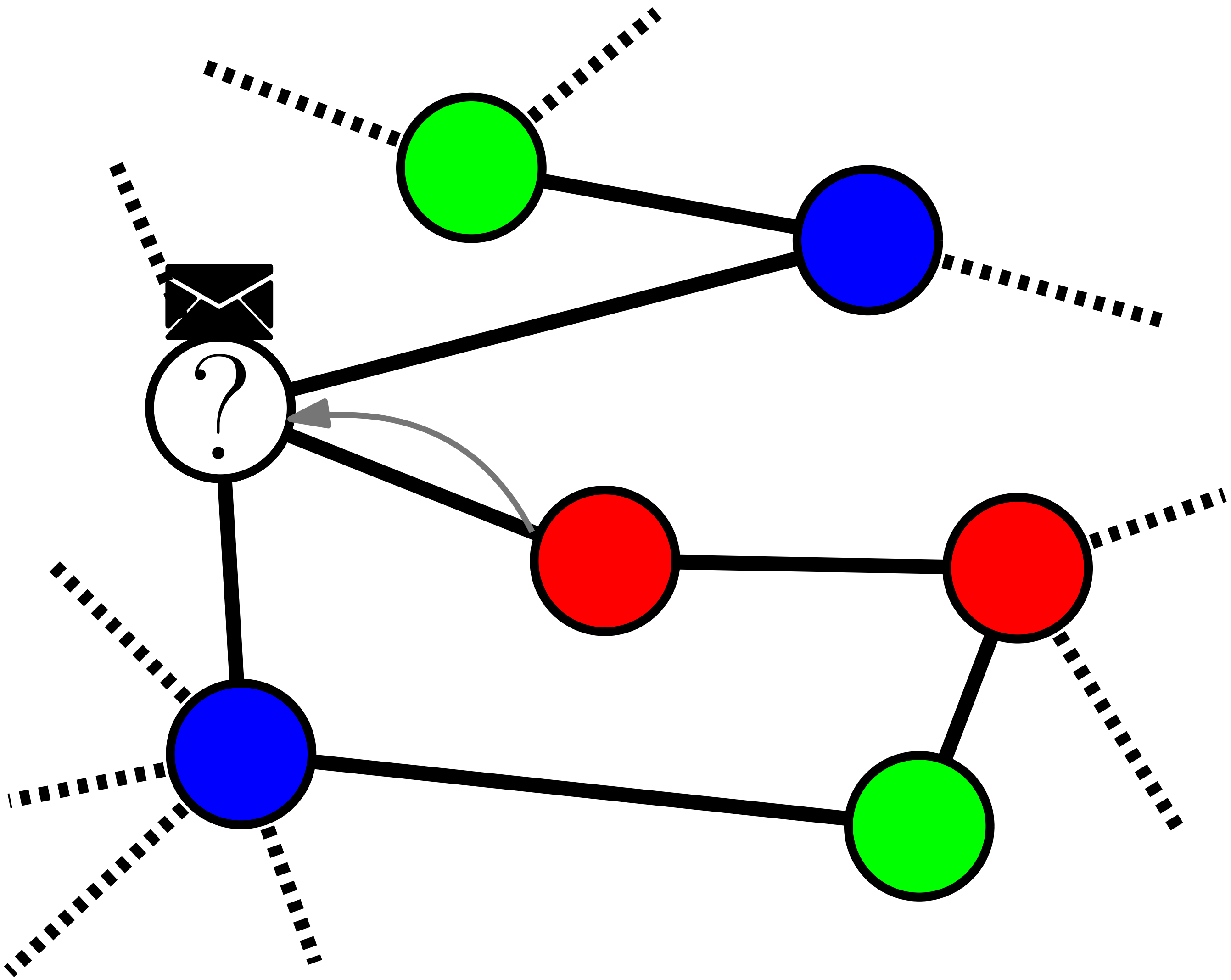
Extension to d-Regular Expanders



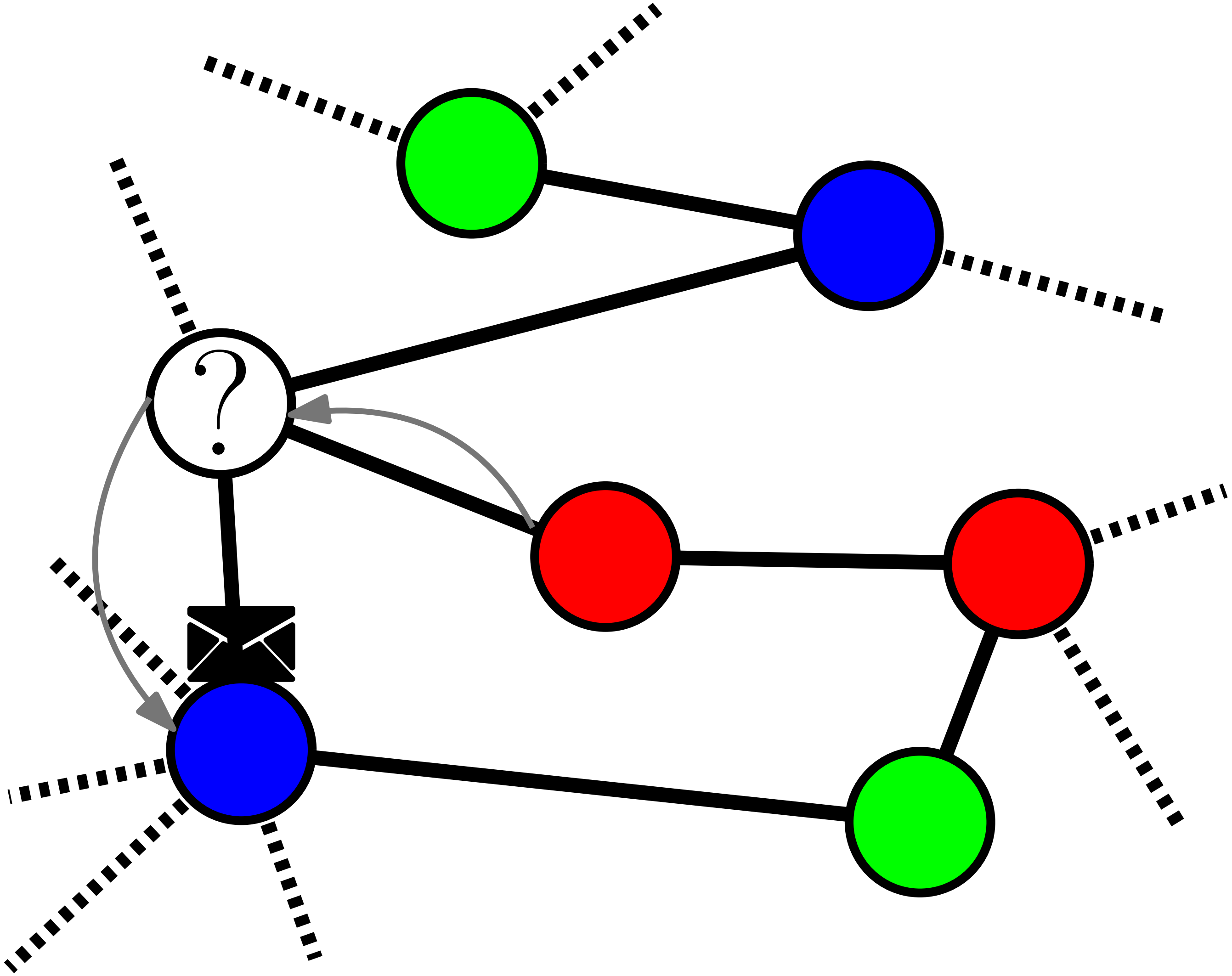
Extension to d-Regular Expanders



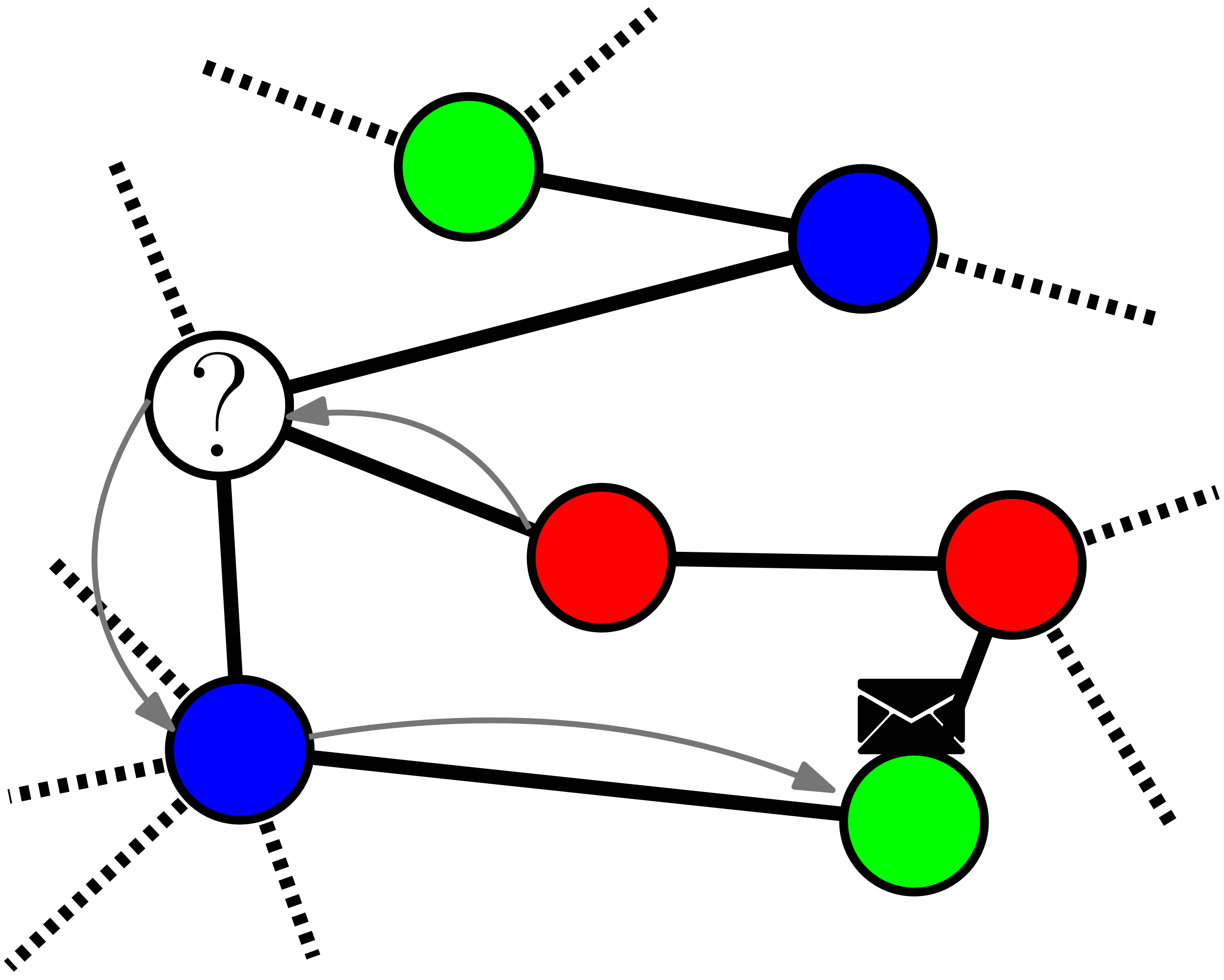
Extension to d-Regular Expanders



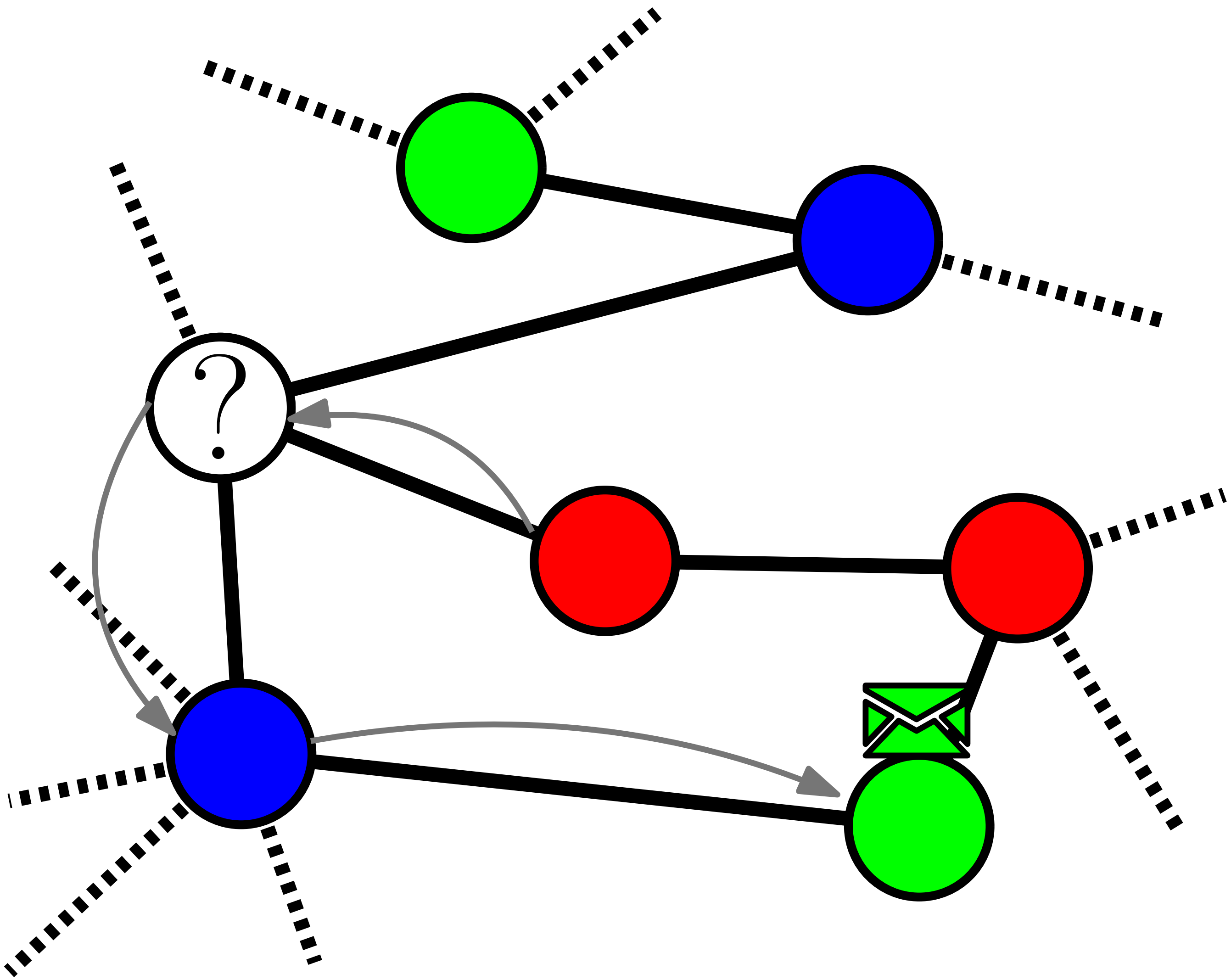
Extension to d-Regular Expanders



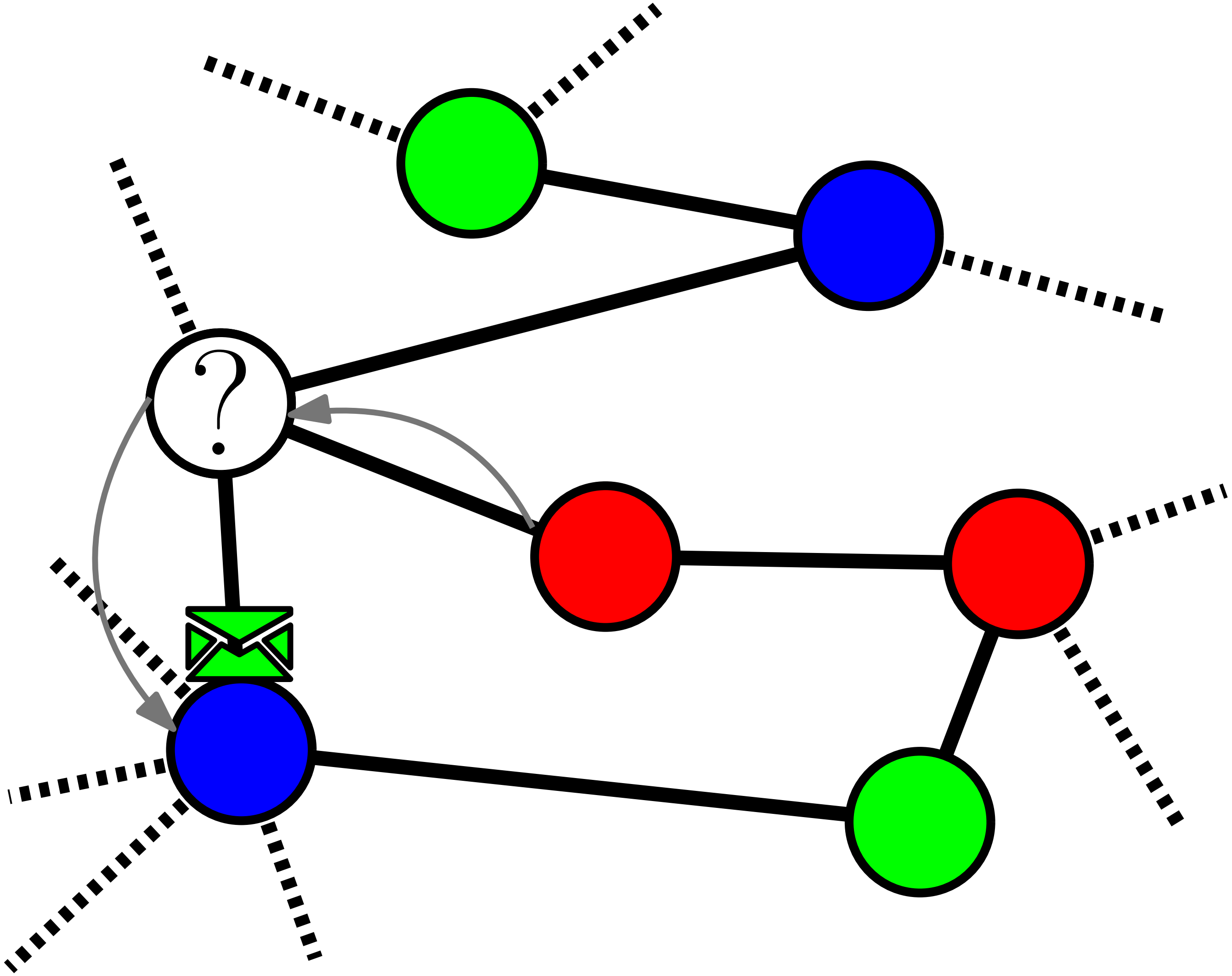
Extension to d-Regular Expanders



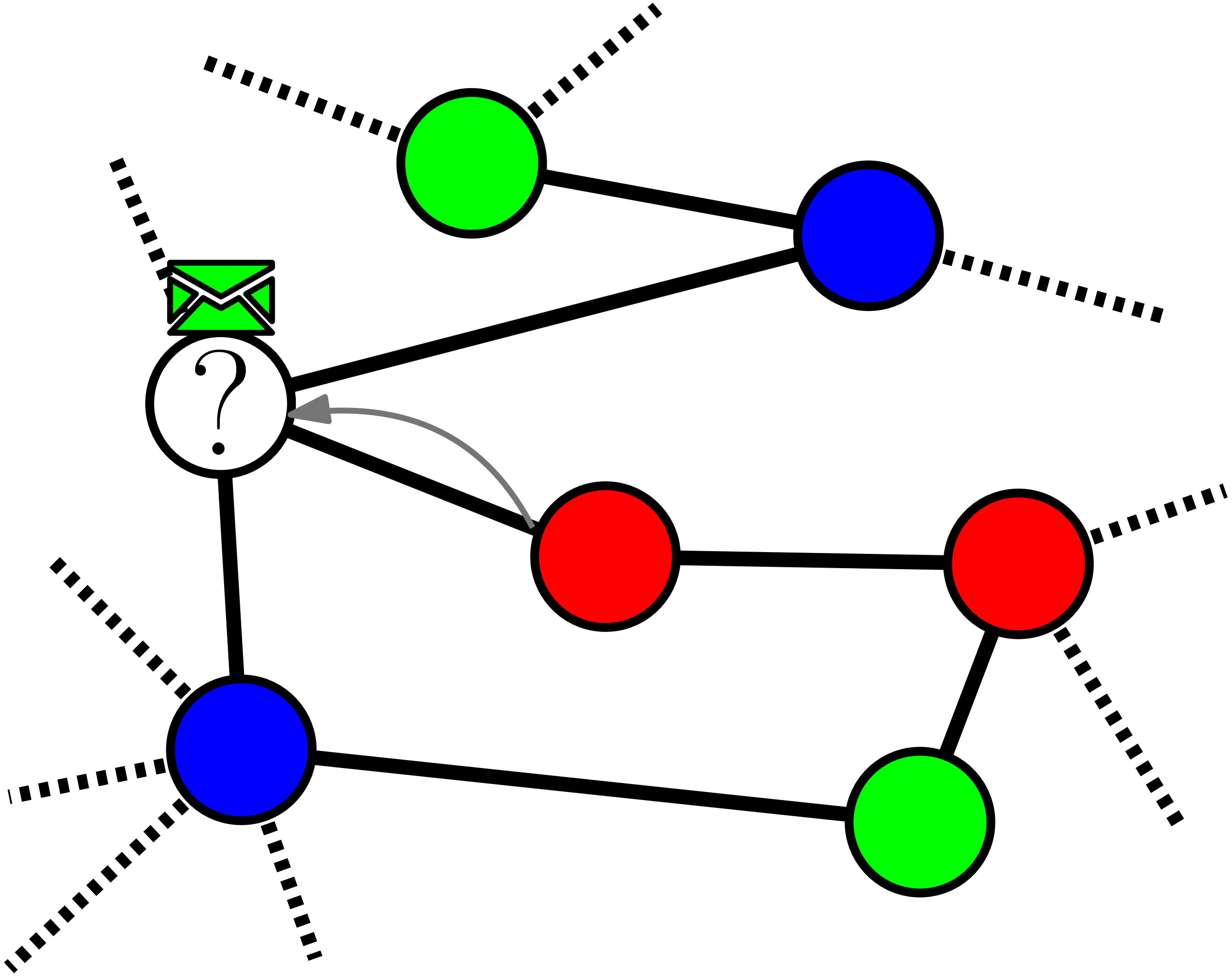
Extension to d-Regular Expanders



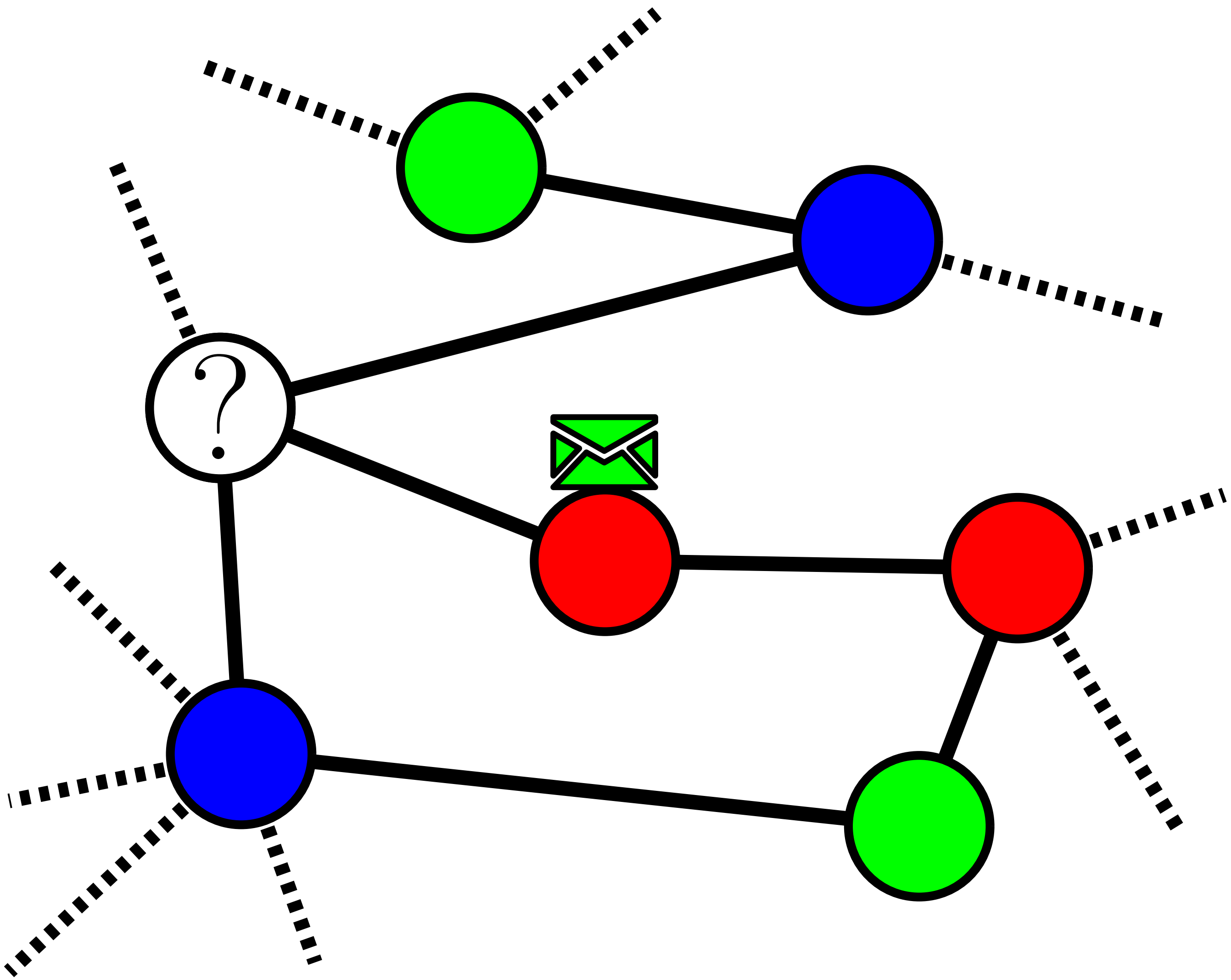
Extension to d-Regular Expanders



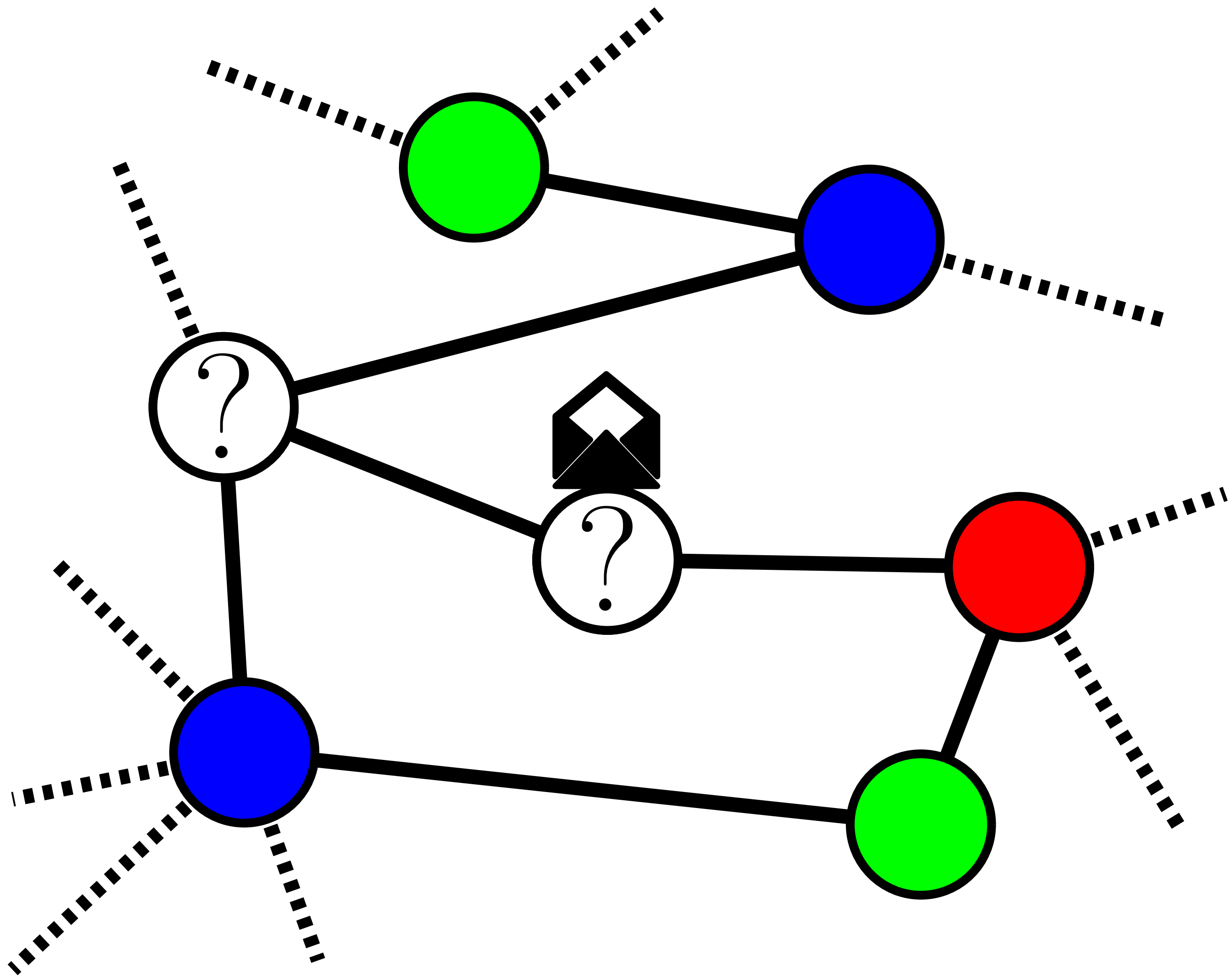
Extension to d-Regular Expanders



Extension to d-Regular Expanders

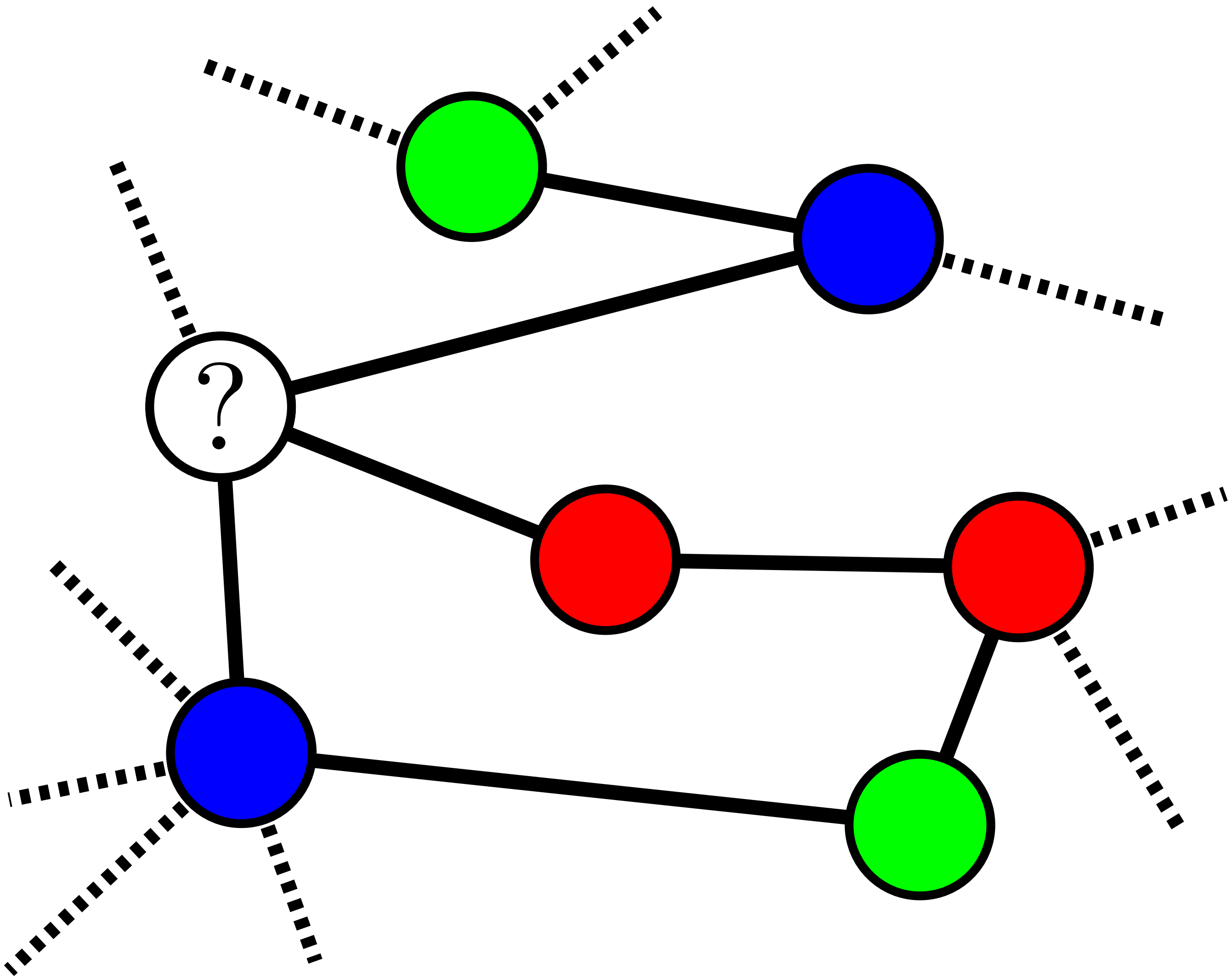


Extension to d-Regular Expanders



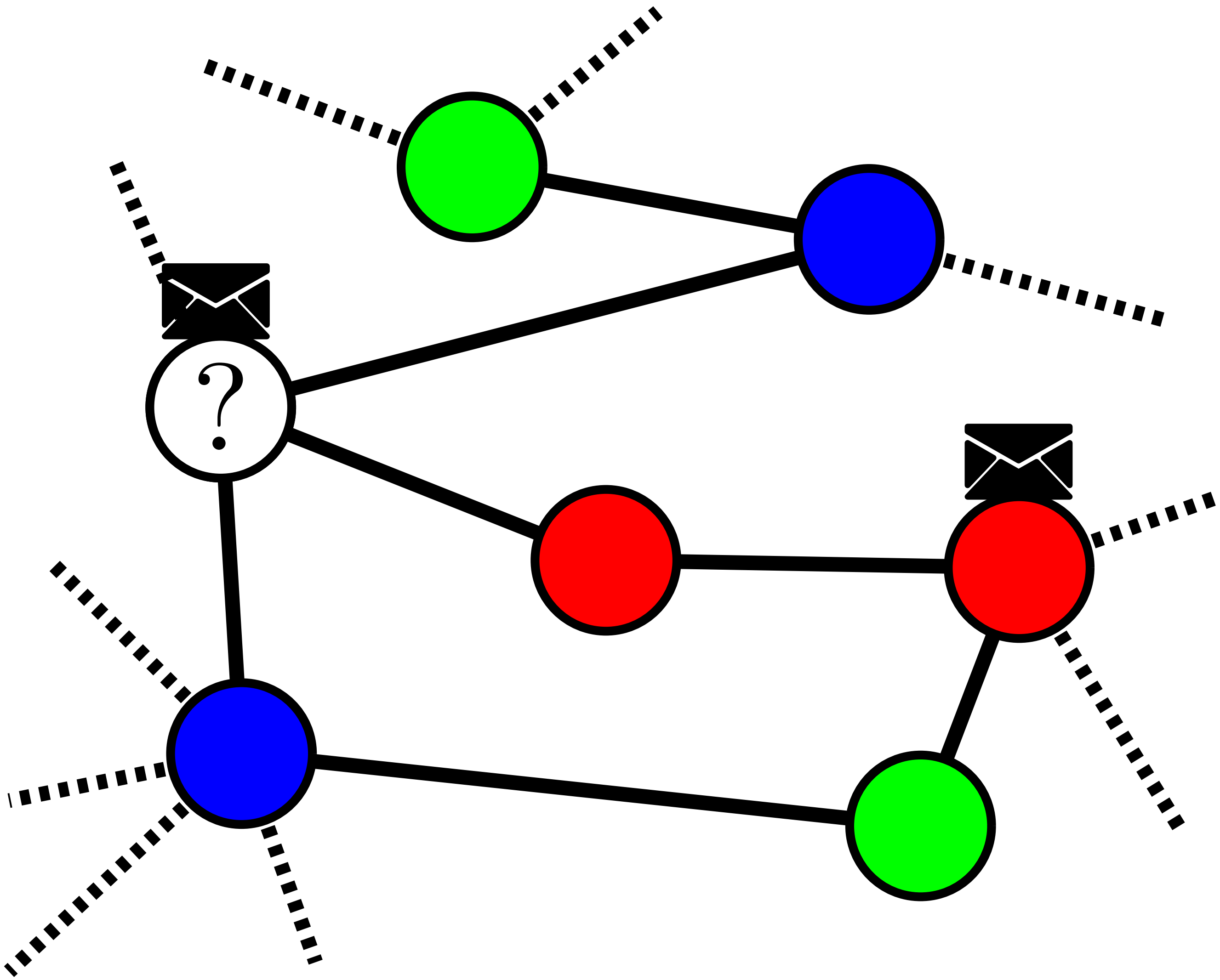
Random Walks in the $\mathcal{GOSSIP}$ Model

Issue. The $\mathcal{GOSSIP}$ model with $O(\text{polylog}\,n)$ limit on message size: congestion when random walks meet.



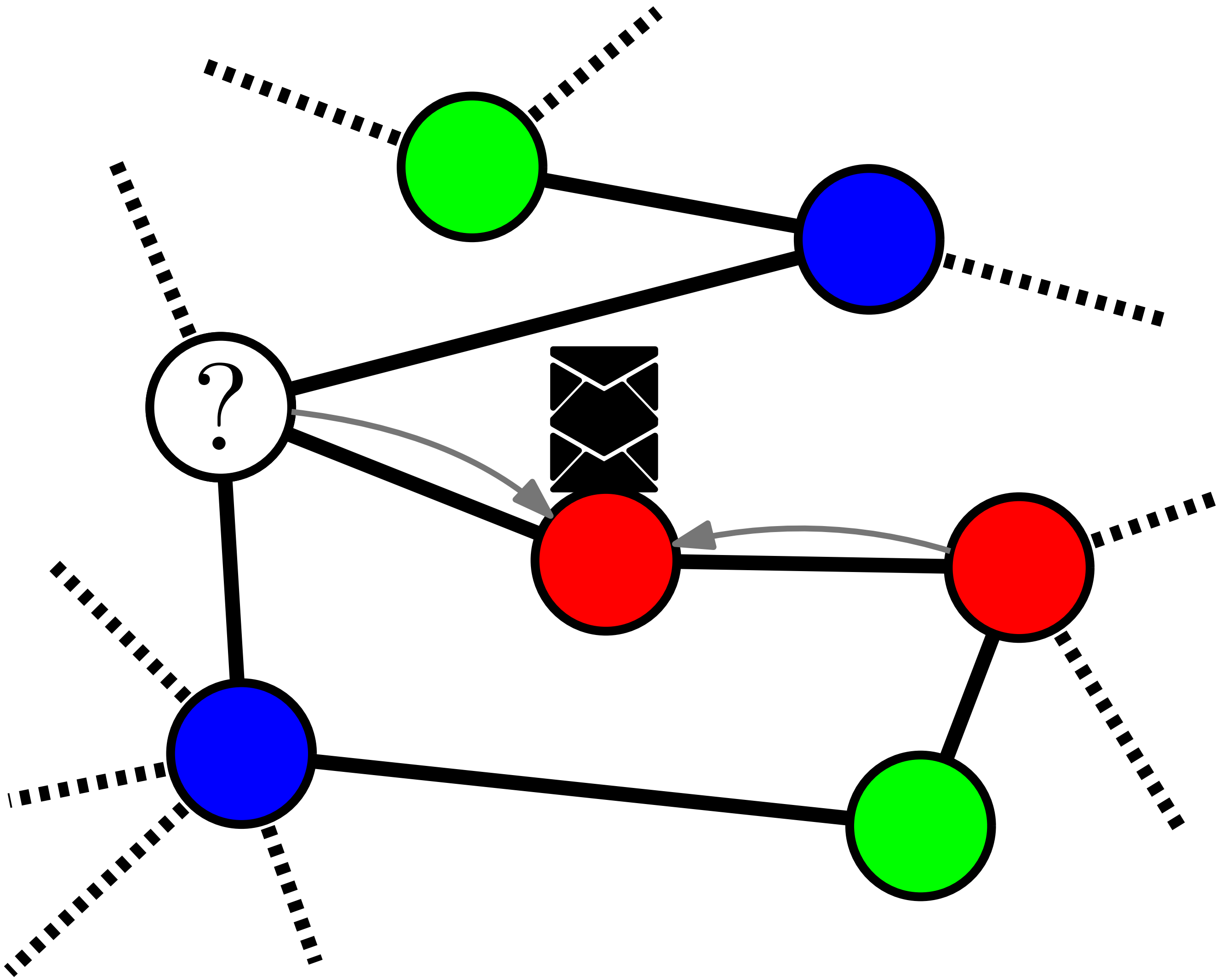
Random Walks in the $\mathcal{GOSSIP}$ Model

Issue. The $\mathcal{GOSSIP}$ model with $O(\text{polylog} n)$ limit on message size: congestion when random walks meet.



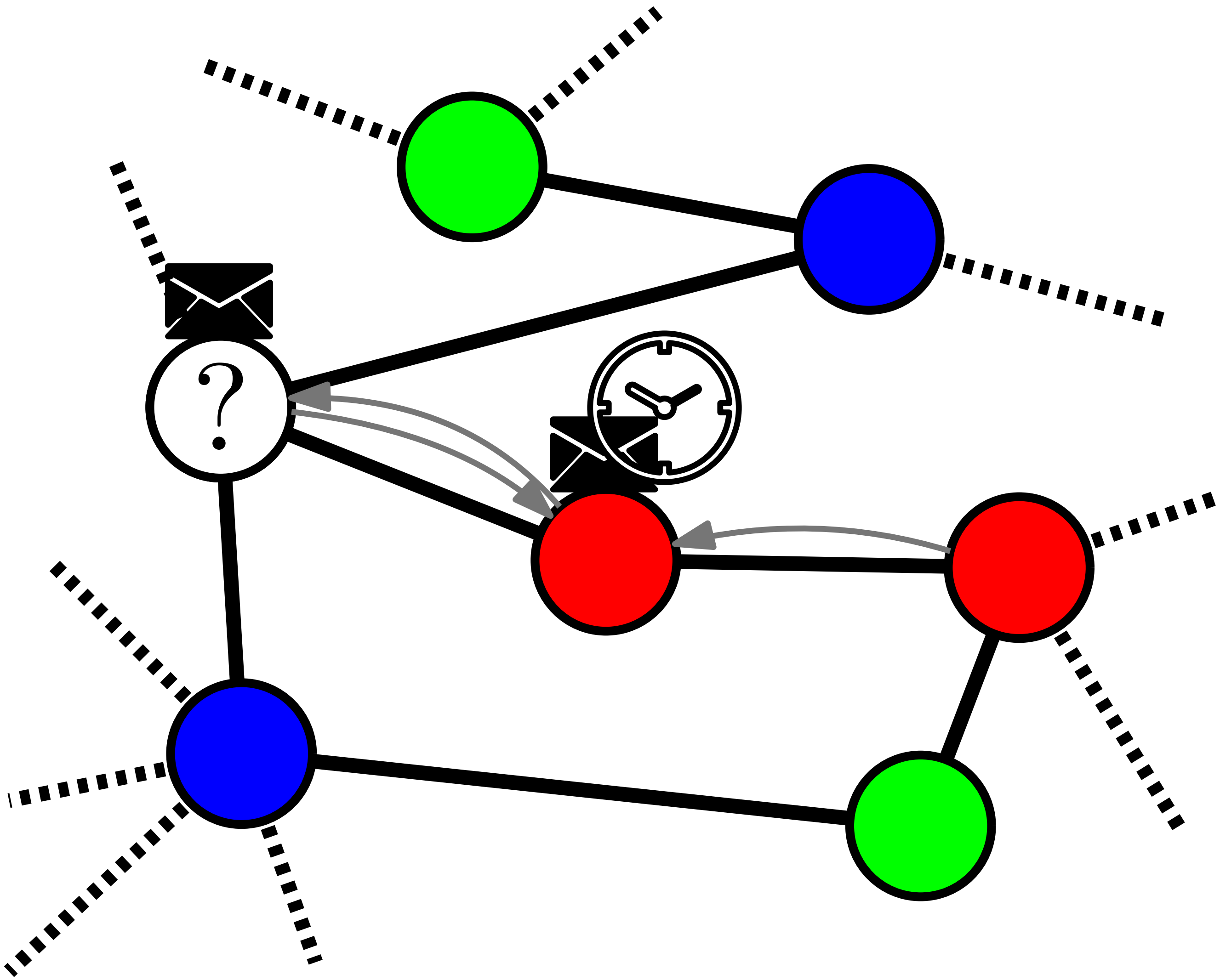
Random Walks in the $\mathcal{GOSSIP}$ Model

Issue. The $\mathcal{GOSSIP}$ model with $O(\text{polylog} n)$ limit on message size: congestion when random walks meet.



Random Walks in the $\mathcal{GOSSIP}$ Model

Issue. The $\mathcal{GOSSIP}$ model with $O(\text{polylog}n)$ limit on message size: congestion when random walks meet.



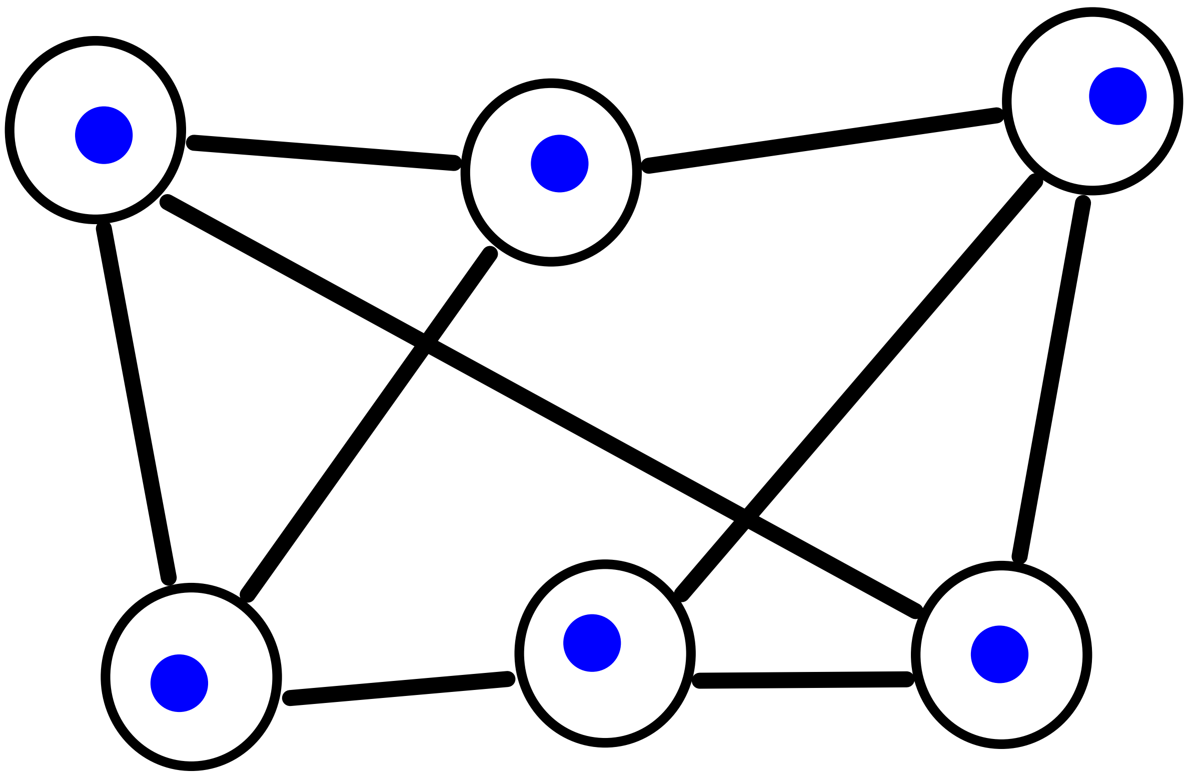
Part 2: Congestion of $\mathcal{GOSSIP}$ random walks

1. Majority Consensus
   (a) 3-Majority (take I)
   (b) Undecided-State
2. Congestion of $\mathcal{GOSSIP}$ random walks
3. Stabilizing Consensus
   (a) 3-Majority (take II)

Congestion in $\mathcal{GOSSIP}$ Random Walks

Goal: keep max load below $\mathcal{O}(\log n)$.

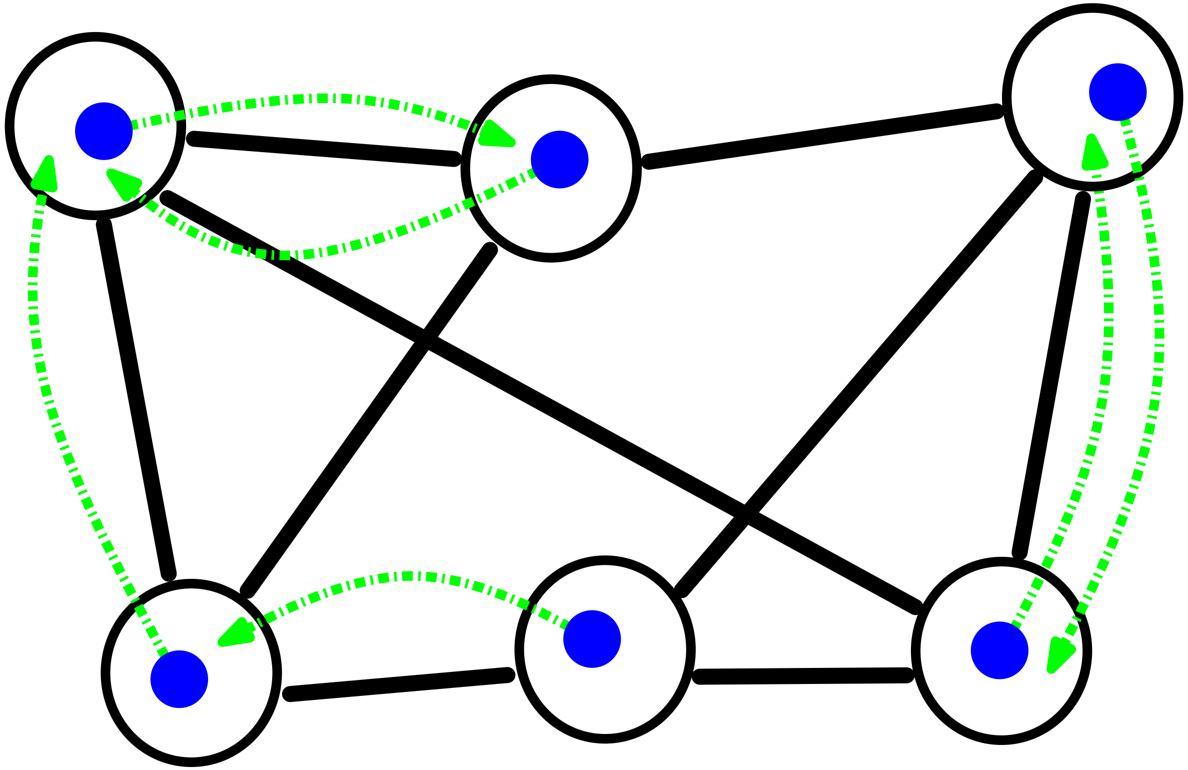
max # of tokens on each node



Congestion in $\mathcal{GOSSIP}$ Random Walks

Goal: keep max load below $\mathcal{O}(\log n)$.

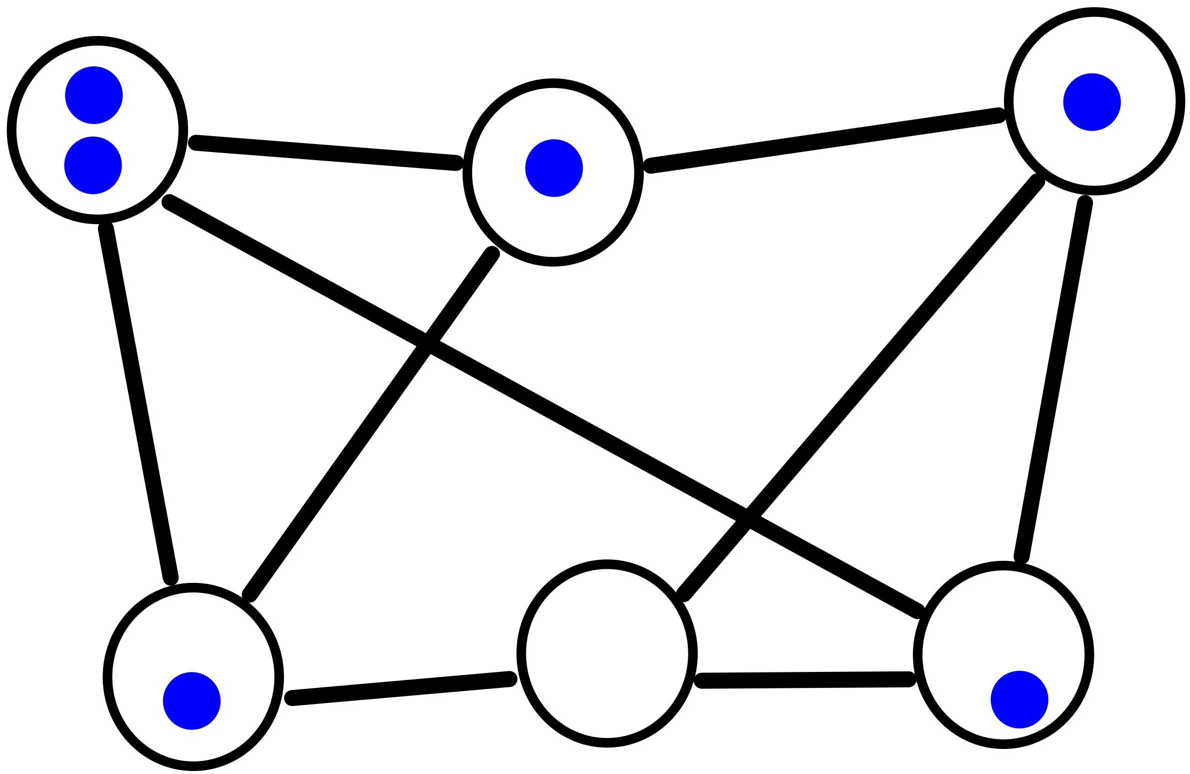
Simple random walks: max load $\mathcal{O}(\log n)$ w.h.p.



Congestion in $\mathcal{GOSSIP}$ Random Walks

Goal: keep max load below $\mathcal{O}(\log n)$.

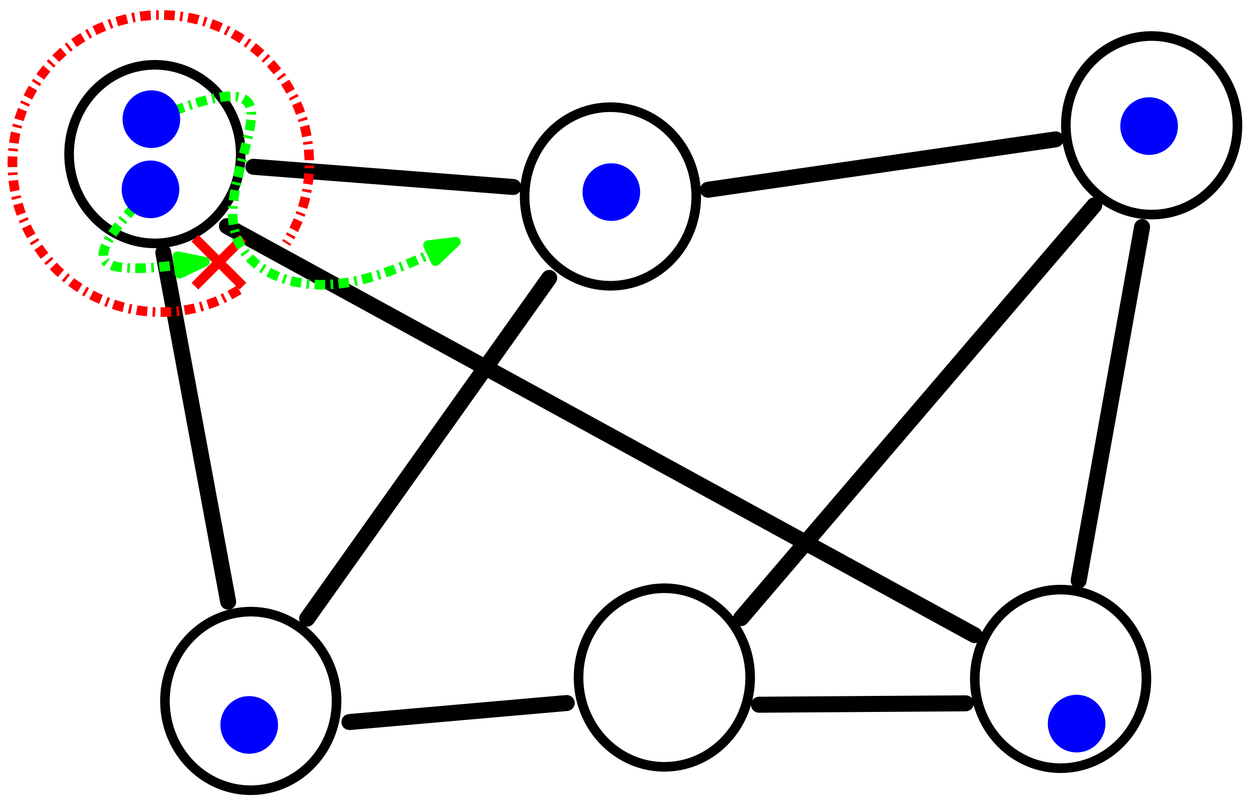
Simple random walks: max load $\mathcal{O}(\log n)$ w.h.p.



Congestion in $\mathcal{GOSSIP}$ Random Walks

Goal: keep max load below $\mathcal{O}(\log n)$.

Simple random walks: max load $\mathcal{O}(\log n)$ w.h.p.



$\mathcal{GOSSIP}$ model [Censor-Hillel et al. '12]: only one token moves from each node (limited communication).

Max load of $\mathcal{GOSSIP}$ random walks: $\mathcal{O}(\log n)$?

Some Related Work

Information exchange in phone-call model [Berenbrink et al. 2010, Elsässer et al. 2015]: analysis for polylog(n) rounds.

Some Related Work

Information exchange in phone-call model [Berenbrink et al. 2010, Elsässer et al. 2015]: analysis for polylog(n) rounds.

Mixing time on regular expanders [Becchetti et al. 2015]: maximum load $\sqrt{t}$ (t rounds).

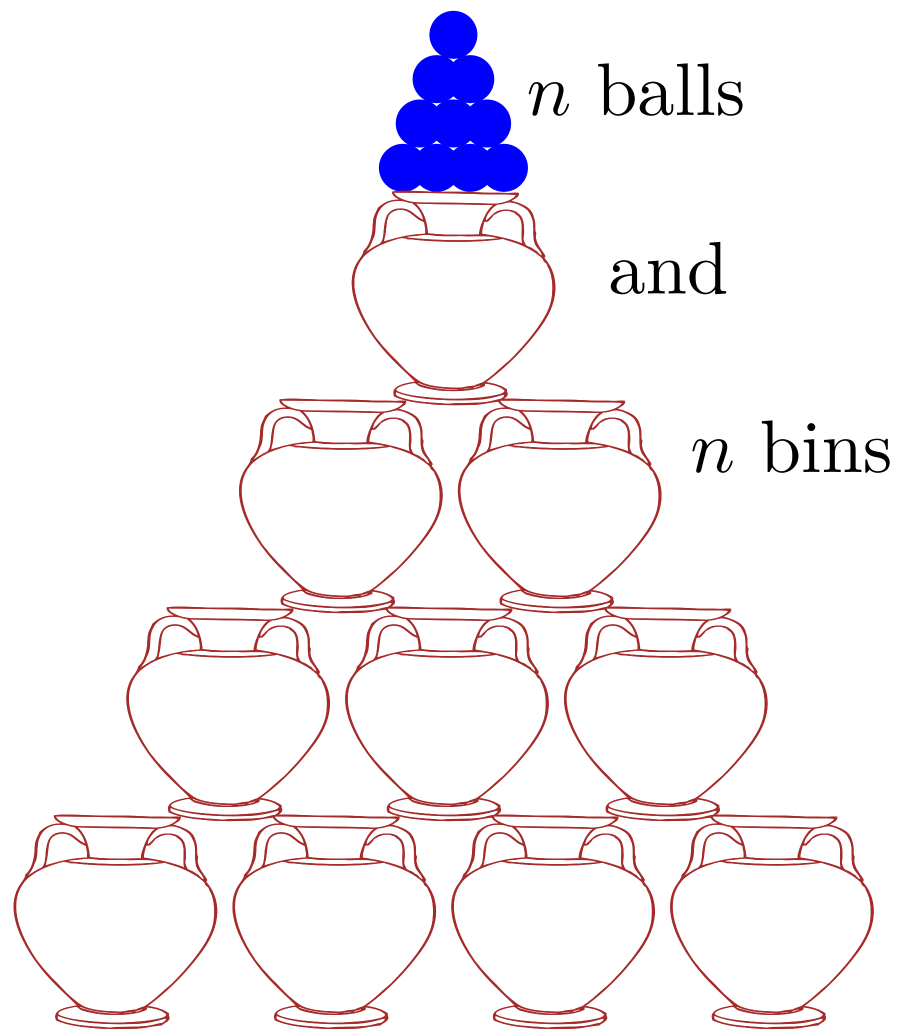
Some Related Work

Information exchange in phone-call model [Berenbrink et al. 2010, Elsässer et al. 2015]: analysis for polylog(n) rounds.

Mixing time on regular expanders [Becchetti et al. 2015]: maximum load $\sqrt{t}$ (t rounds).

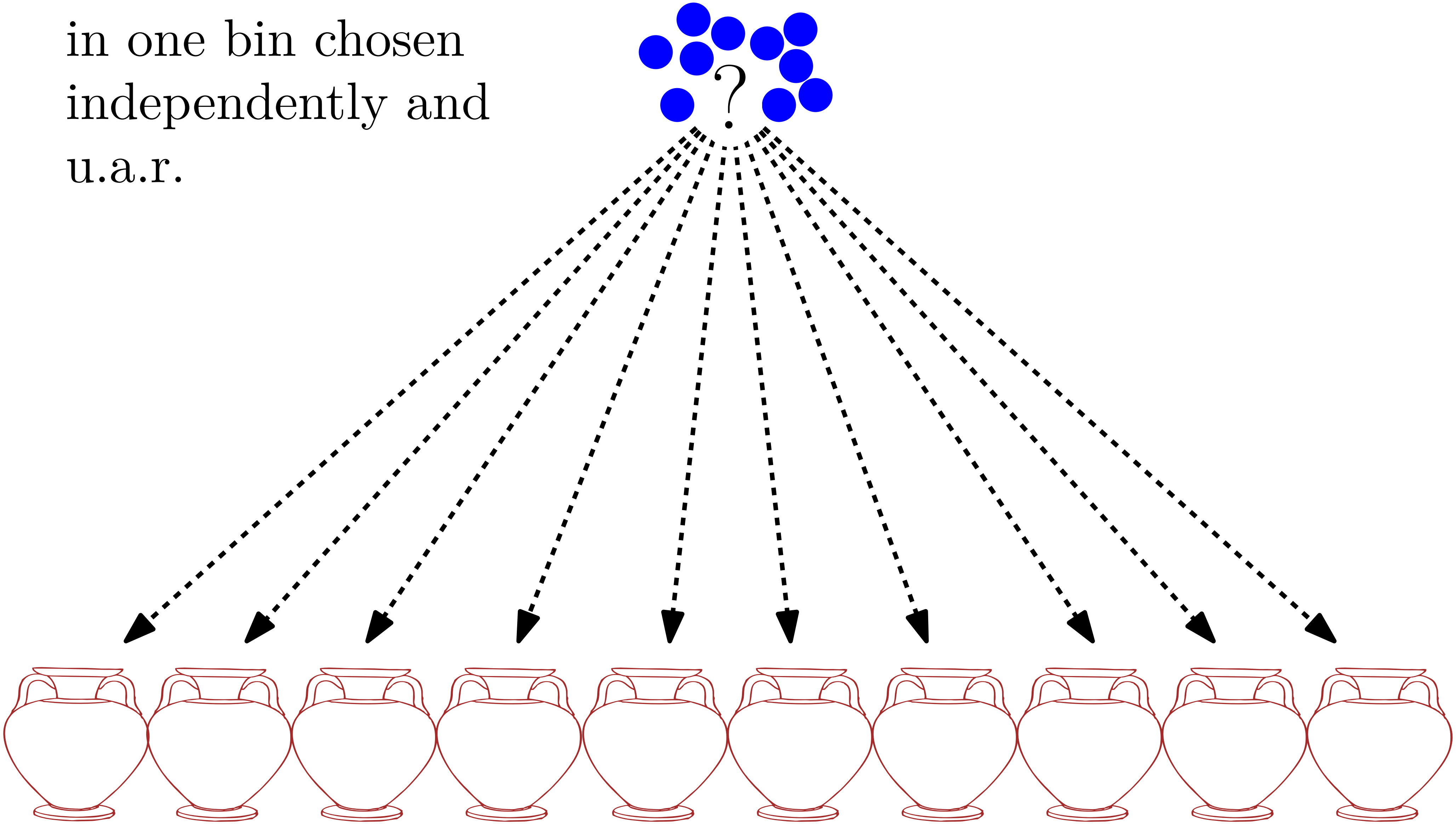
Closed Jackson networks in queueing theory: asynchronous version of $\mathcal{GOSSIP}$ r.w.s (admits closed form solution).

Seemingly Off Topic: Balls-into-Bins



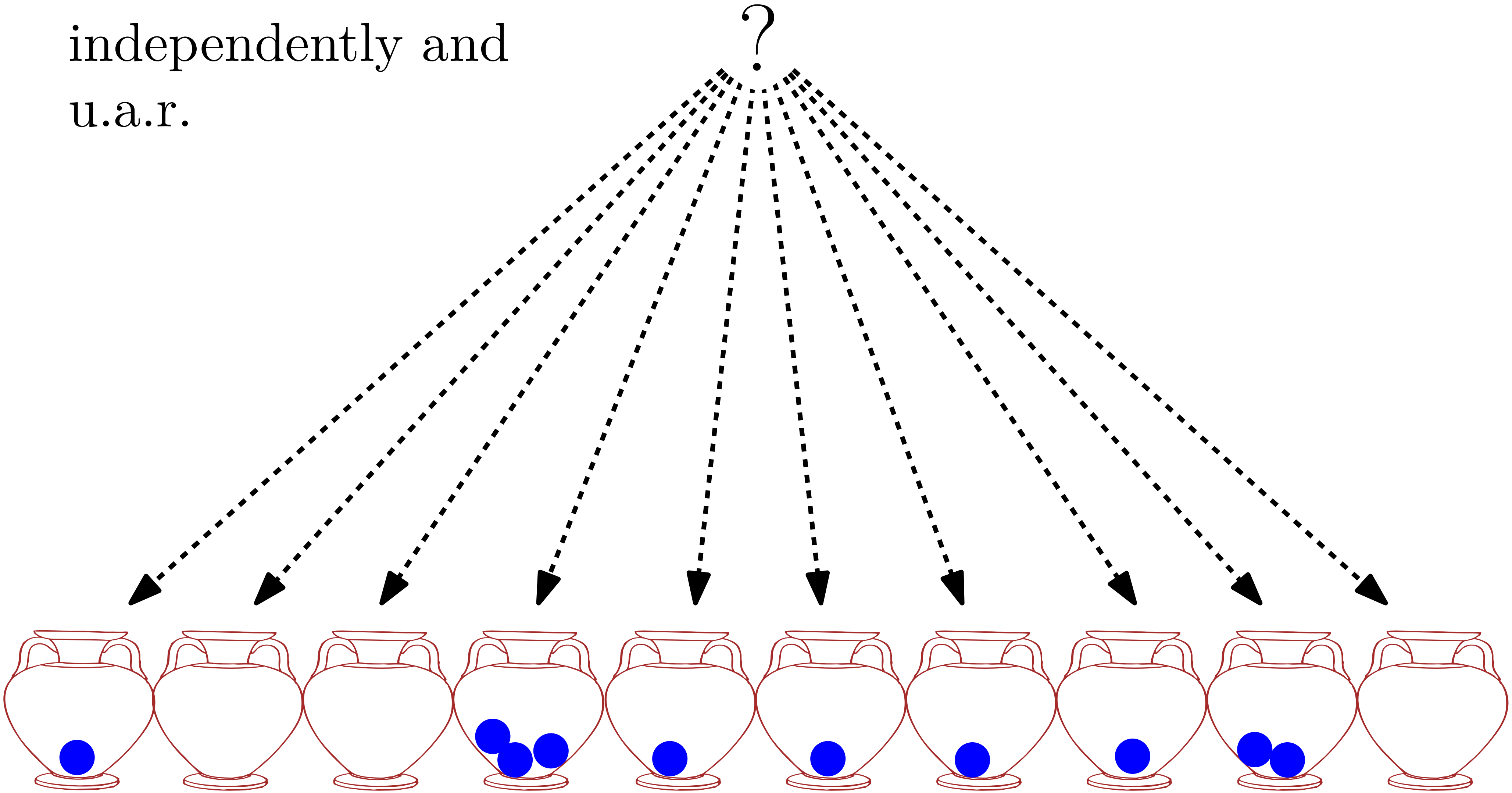
Seemingly Off Topic: Balls-into-Bins

Each ball is thrown in one bin chosen independently and u.a.r.



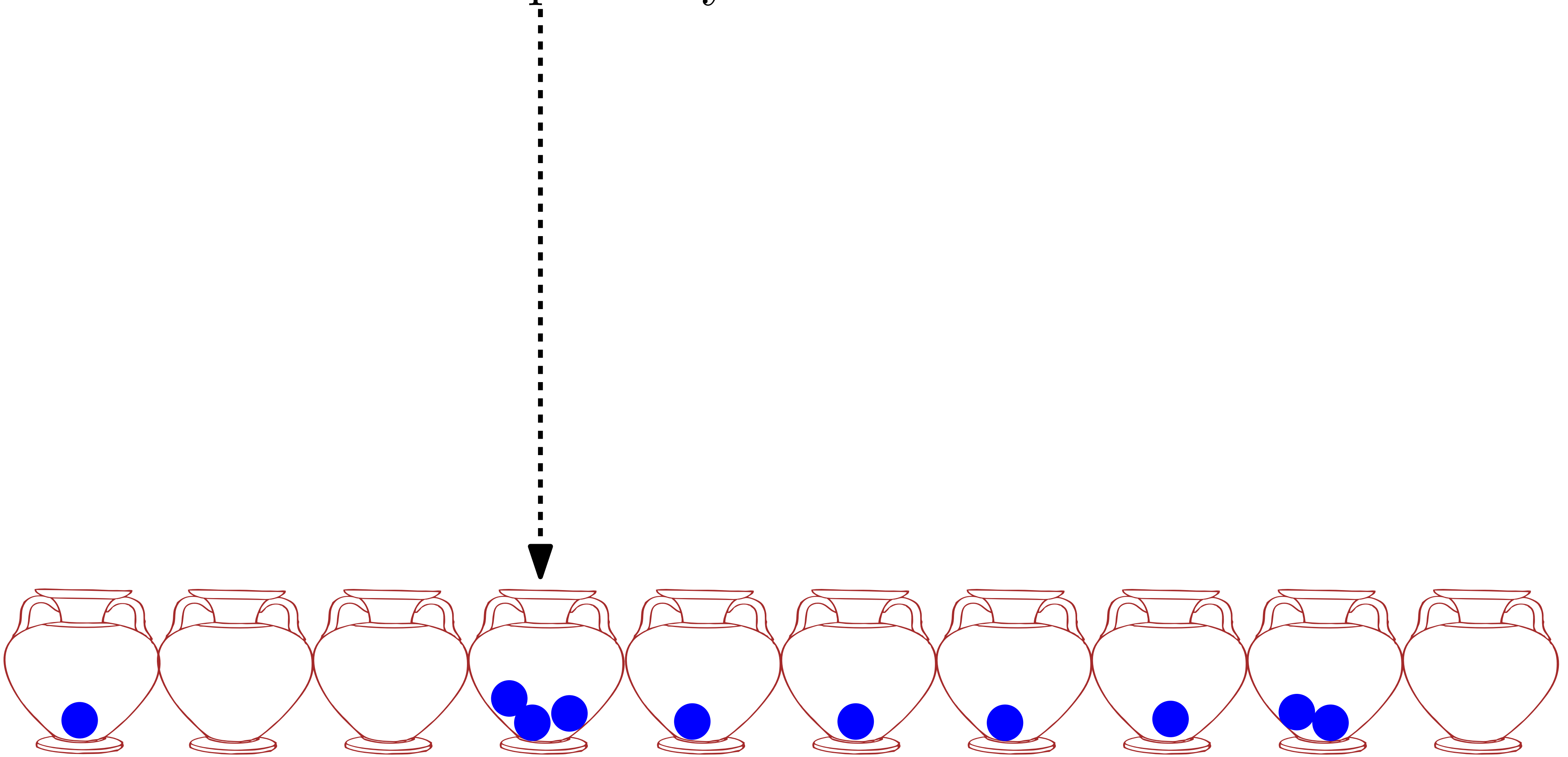
Seemingly Off Topic: Balls-into-Bins

Each ball is thrown in one bin chosen independently and u.a.r.



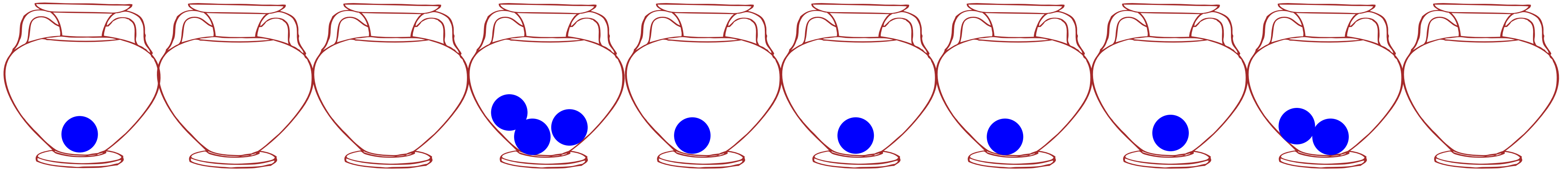
Seemingly Off Topic: Balls-into-Bins

Maximum load: maximum number of balls that end up in any bin.



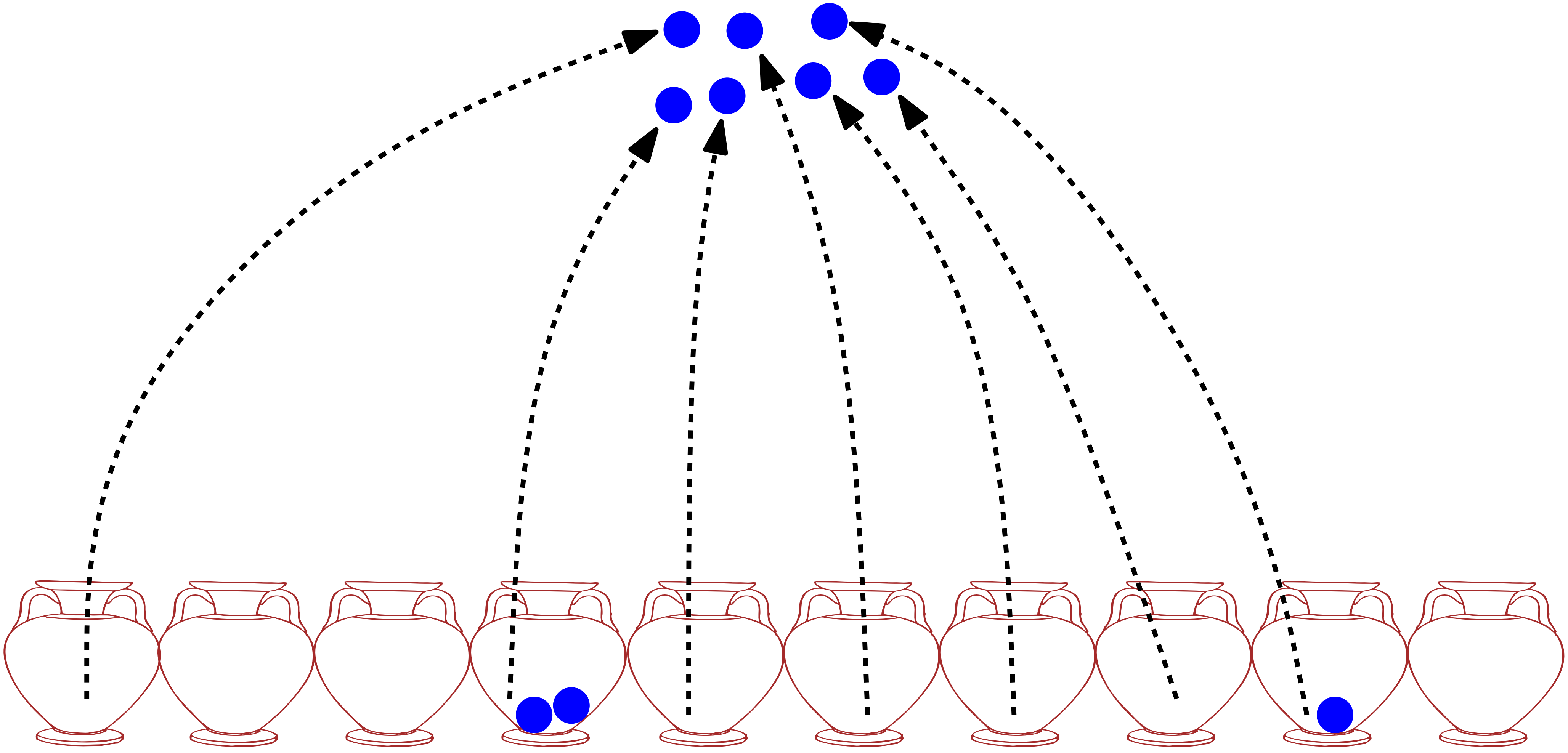
Repeated Balls-into-Bins & $\mathcal{GOSSIP}$ R. W.s

At each round, pick one ball from each non-empty bin...



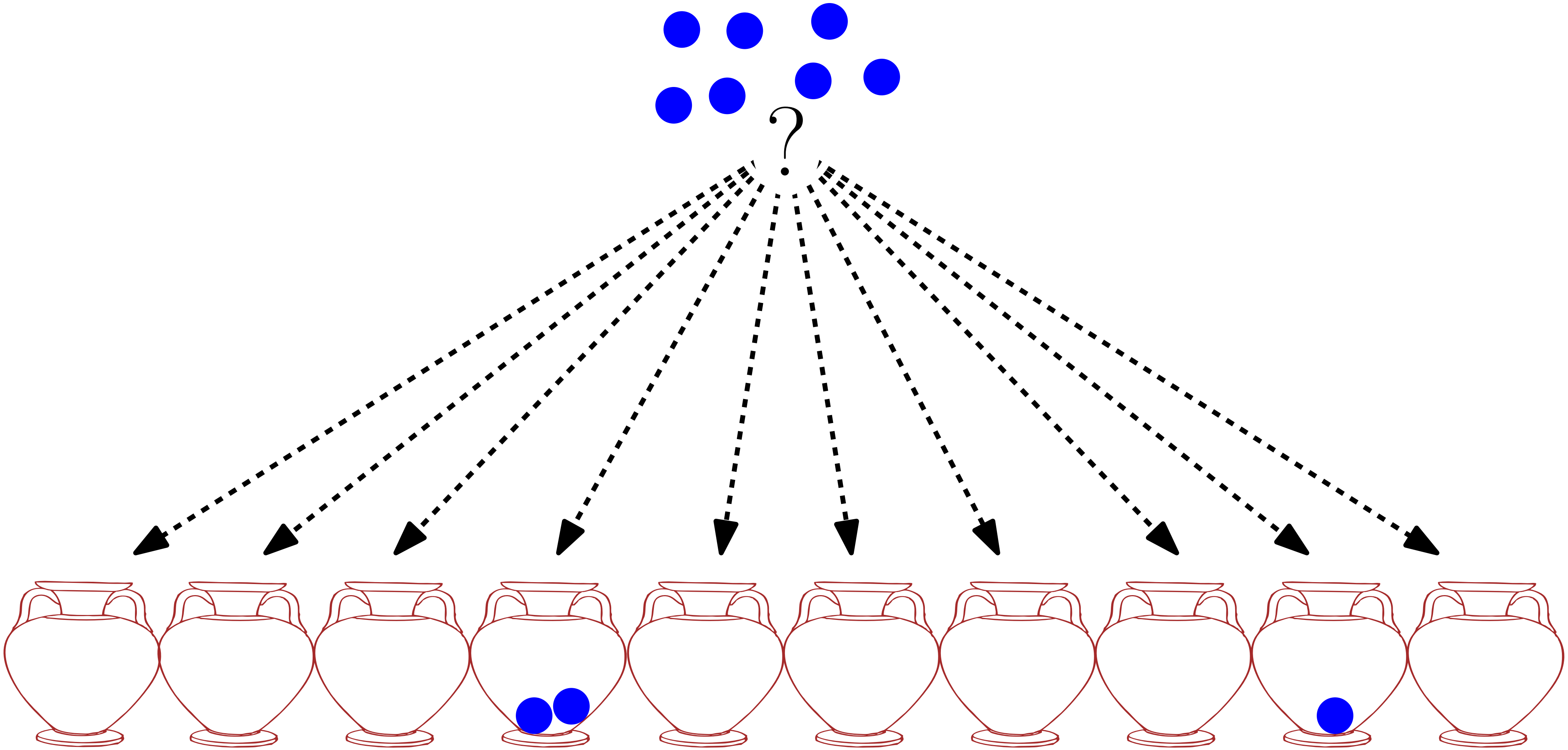
Repeated Balls-into-Bins & $\mathcal{GOSSIP}$ R. W.s

At each round, pick one ball from each non-empty bin...



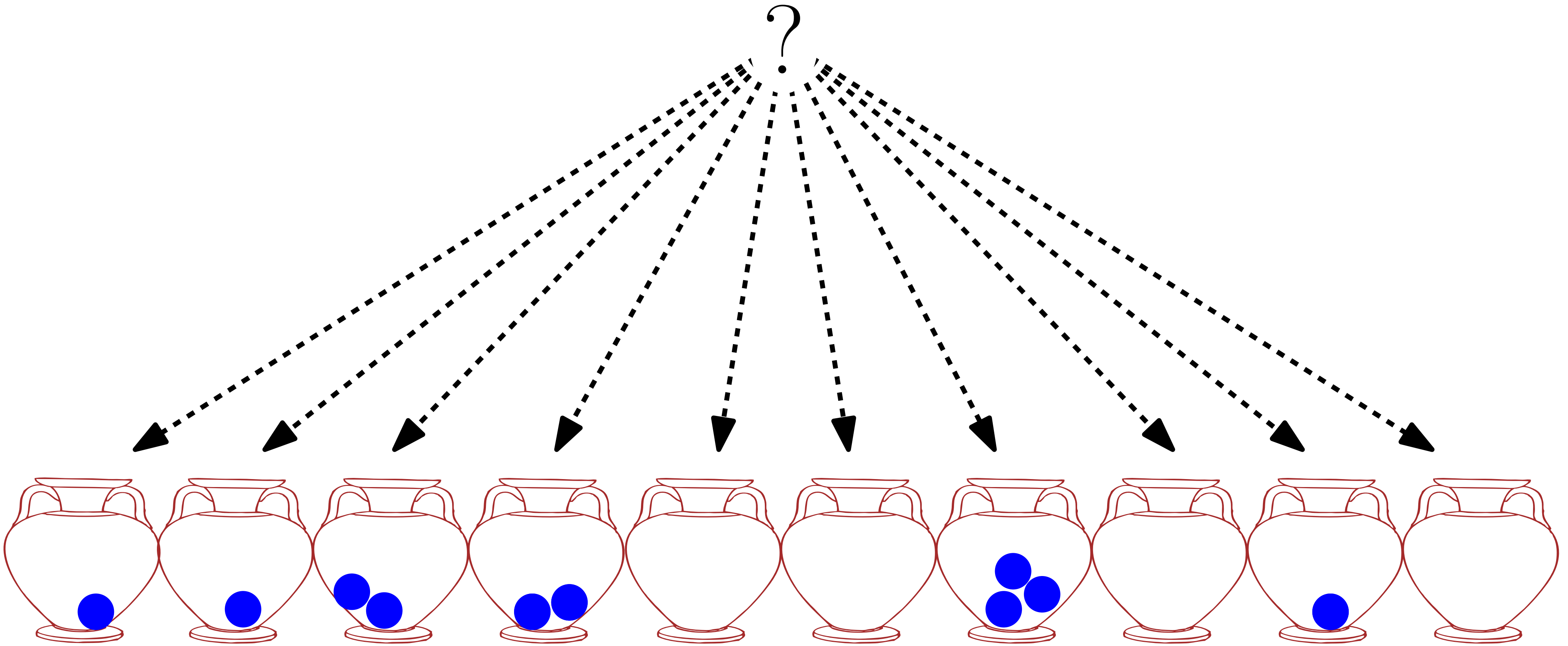
Repeated Balls-into-Bins & *GOSSIP* R. W.s

At each round, pick one ball from each non-empty bin...
...and throw them again u.a.r.



Repeated Balls-into-Bins & $\mathcal{GOSSIP}$ R. W.s

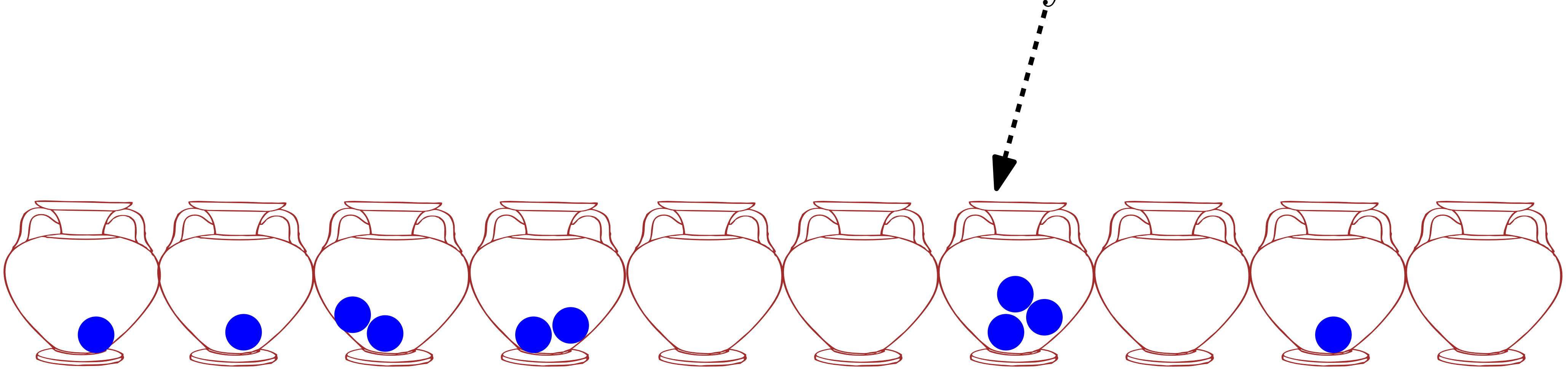
At each round, pick one ball from each non-empty bin...
...and throw them again u.a.r.



Repeated Balls-into-Bins & $\mathcal{GOSSIP}$ R. W.s

At each round, pick one ball from each non-empty bin...
...and throw them again u.a.r.

Max load: max. number of balls in any bin.



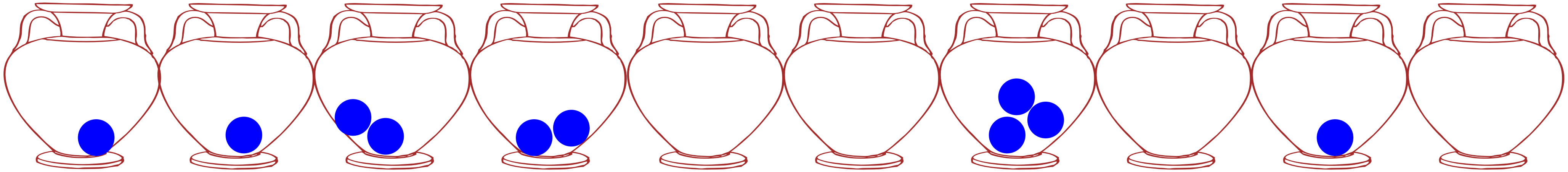
Repeated Balls-into-Bins & $\mathcal{GOSSIP}$ R. W.s

At each round, pick one ball from each non-empty bin...
...and throw them again u.a.r.

Repeated n balls in n bins

=

n $\mathcal{GOSSIP}$ r.w.s on n-node complete graph (with loops)



Analyzing Repeated Balls-into-Bins

The infamous stochastic dependence:
negative association,
Poisson approximation...

Analyzing Repeated Balls-into-Bins

The infamous stochastic dependence:
negative association,
Poisson approximation...

Stochastic dependence in repeated balls-into-bins:

How to handle *time dependence*?

Analyzing Repeated Balls-into-Bins

The infamous stochastic dependence:
negative association,
Poisson approximation...

Stochastic dependence in repeated balls-into-bins:

How to handle *time dependence*?

A *coupling w.h.p.*: the tetris process

Analyzing Repeated Balls-into-Bins

The infamous stochastic dependence:
negative association,
Poisson approximation...

Stochastic dependence in repeated balls-into-bins:

How to handle *time dependence*?

A *coupling w.h.p.*: the tetris process

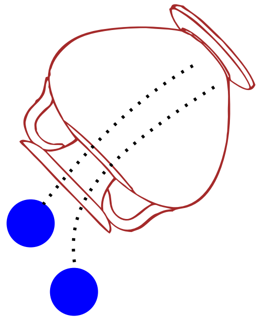
$M_t^{(RBB)} :=$ time t max. load in repeated b.i.b.
$M_t^{(T)} :=$ time t max. load in tetris proc.

$$\Pr(M_t^{(RBB)} \geq k) \leq \Pr(M_t^{(T)} \geq k) + t \cdot e^{-\Theta(n)}$$

Analysis of b.i.b.

Lemma (empty bins).

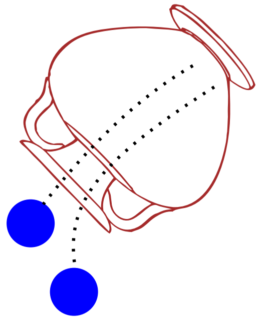
At the next round $|\{\text{empty bins}\}| \geq \frac{n}{4}$ w.h.p.



Analysis of b.i.b.

Lemma (empty bins).

At the next round $|\{\text{empty bins}\}| \geq \frac{n}{4}$ w.h.p.



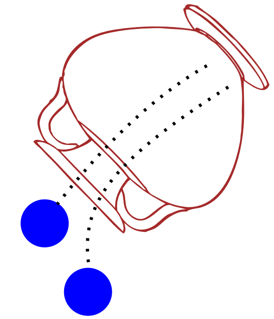
Corollary

At the next round $|\{\text{thrown balls}\}| \leq \frac{3n}{4}$ w.h.p.

Analysis of b.i.b.

Lemma (empty bins).

At the next round $|\{\text{empty bins}\}| \geq \frac{n}{4}$ w.h.p.



Corollary

At the next round $|\{\text{thrown balls}\}| \leq \frac{3n}{4}$ w.h.p.

Proof

$a := |\{\text{empty bins}\}|$, $b := |\{\text{bins with 1 ball}\}|$,
$X := |\{\text{new empty bins}\}|$

1. $\mathbb{E}[X] = (a+b)(1-1/n)^{n-a}$
2. $n-(a+b) \leq a \implies \mathbb{E}[X] \geq (1+\epsilon)\frac{n}{4}$
3. Chernoff bound (negative association)

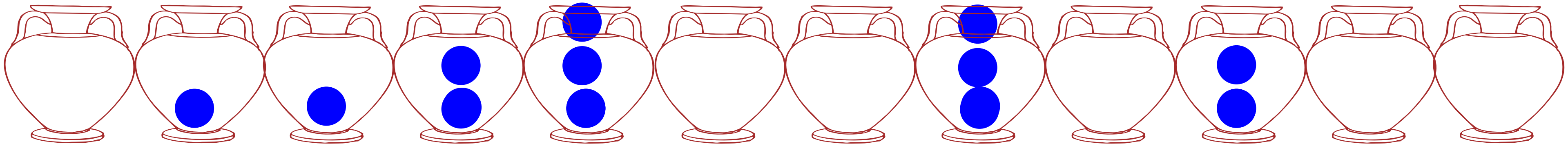
□

Analysis of b.i.b.

Tetris Process

1- Throw away a ball from each non-empty bin

2- Throw $3n/4$ balls in the bins u.a.r.



Analysis of b.i.b.

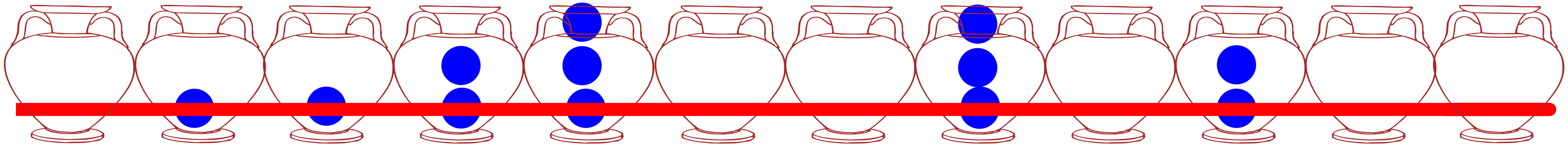
Tetris Process

1- Throw away a ball from each non-empty bin

2- Throw $3n/4$ balls in the bins u.a.r.

Coupling

Step 1: As rep. b.i.b., take one ball from each bin



Analysis of b.i.b.

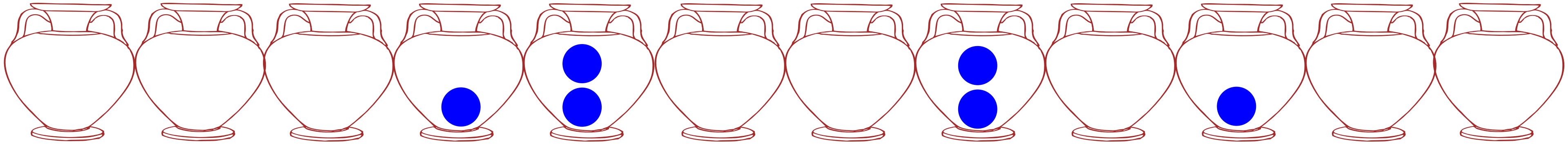
Tetris Process

1- Throw away a ball from each non-empty bin

2- Throw $3n/4$ balls in the bins u.a.r.

Coupling

Step 1: As rep. b.i.b., take one ball from each bin



Analysis of b.i.b.

Tetris Process

1- Throw away a ball from each non-empty bin

2- Throw $3n/4$ balls in the bins u.a.r.

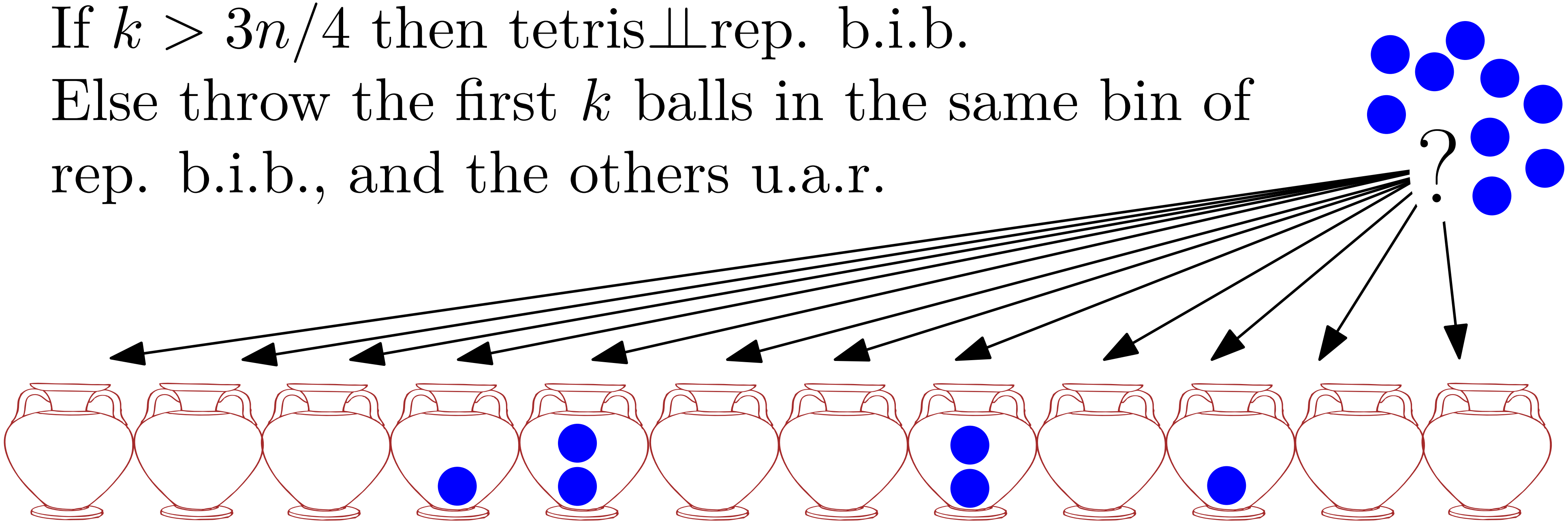
Coupling

Step 1: As rep. b.i.b., take one ball from each bin

Step 2: Let $k :=$ non-empty bins in rep. b.i.b.

If $k > 3n/4$ then tetris$\perp\!\!\!\perp$rep. b.i.b.

Else throw the first k balls in the same bin of rep. b.i.b., and the others u.a.r.



Analysis of b.i.b.

Tetris Process

1- Throw away a ball from each non-empty bin

2- Throw $3n/4$ balls in the bins u.a.r.

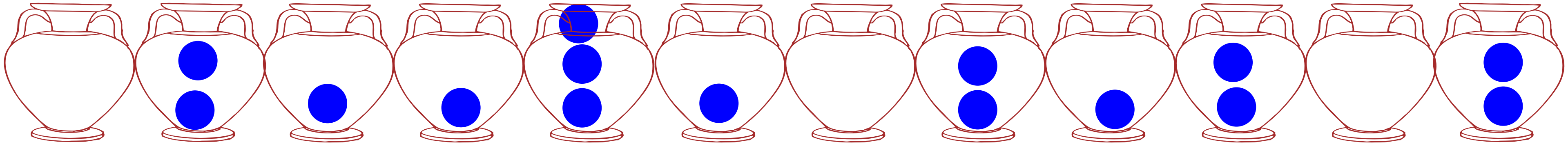
Coupling

Step 1: As rep. b.i.b., take one ball from each bin

Step 2: Let $k :=$ non-empty bins in rep. b.i.b.

If $k > 3n/4$ then tetris$\perp\!\!\!\perp$rep. b.i.b.

Else throw the first k balls in the same bin of rep. b.i.b., and the others u.a.r.



Analysis of b.i.b.

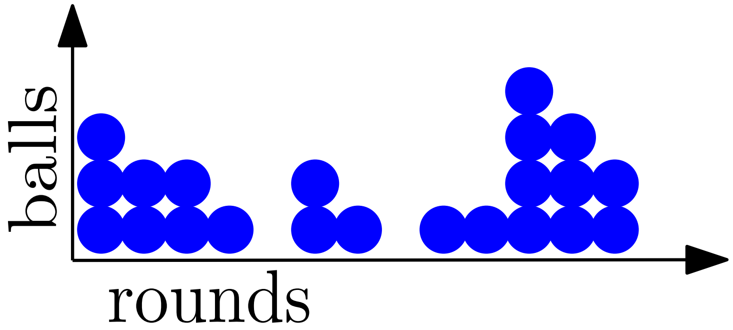
Theorem

The max. load of the tetris process is $\mathcal{O}(\log n)$ for poly(n) rounds w.h.p.

Analysis of b.i.b.

Theorem
The max. load of the tetris process is $\mathcal{O}(\log n)$ for poly(n) rounds w.h.p.

Proof

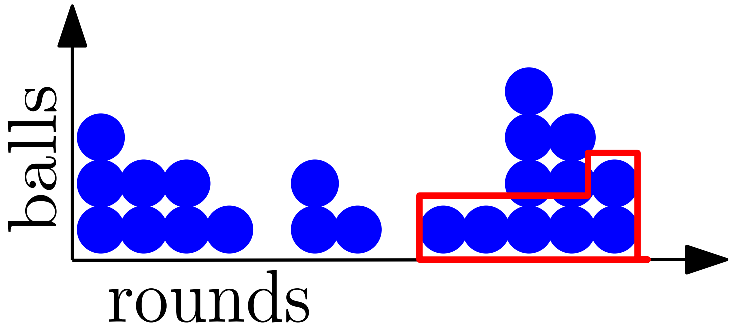


$\mathbb{E}[\text{incoming balls in t rounds}] = \frac{3t}{4}$

Analysis of b.i.b.

Theorem
The max. load of the tetris process is $\mathcal{O}(\log n)$ for poly(n) rounds w.h.p.

Proof



$\mathbb{E}[\text{incoming balls in t rounds}] = \frac{3t}{4}$

$T :=$ # rounds from last time the bin was empty

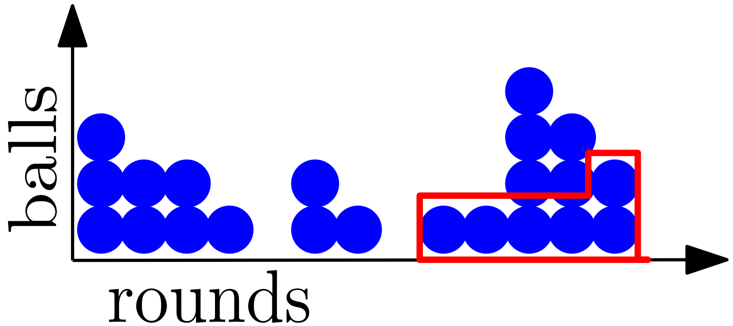
For each bin: load k at round $t \implies$ received $k + T$ balls

$\square$

Analysis of b.i.b.

Theorem
The max. load of the tetris process is $\mathcal{O}(\log n)$ for poly(n) rounds w.h.p.

Proof



$\mathbb{E}[\text{incoming balls in t rounds}] = \frac{3t}{4}$

$T :=$ # rounds from last time the bin was empty

For each bin: load k at round $t \implies$ received $k + T$ balls

□

Lemma
From any configuration, every bin in the tetris proc. is empty at least once every $5n$ rounds w.h.p.

Our Contribution [SPAA '15]

From any configuration, in $\mathcal{O}(n)$ rounds the repeated balls-into-bins process reaches a conf. with max load $\mathcal{O}(\log n)$ w.h.p. and, from any conf. with max load $\mathcal{O}(\log n)$, the max load keeps $\mathcal{O}(\log n)$ for poly(n) rounds w.h.p.

Our Contribution [SPAA '15]

From any configuration, in $\mathcal{O}(n)$ rounds the repeated balls-into-bins process reaches a conf. with max load $\mathcal{O}(\log n)$ w.h.p. and, from any conf. with max load $\mathcal{O}(\log n)$, the max load keeps $\mathcal{O}(\log n)$ for poly(n) rounds w.h.p.

Theorem
After at most $\mathcal{O}(n)$ rounds the max. load of n $\mathcal{GOSSIP}$ r.w.s on n-node complete graph is $\mathcal{O}(\log n)$ w.h.p., and keeps $\mathcal{O}(\log n)$ for poly(n) rounds.

$\mathcal{GOSSIP}$ R.W.s on non-Complete Graphs

The analysis for the complete graph can still be applied *locally* provided that the minimum degree is αn for some constant $\alpha > 0$ (G. Scornavacca's MSc thesis).

$\mathcal{GOSSIP}$ R.W.s on non-Complete Graphs

The analysis for the complete graph can still be applied *locally* provided that the minimum degree is αn for some constant $\alpha > 0$ (G. Scornavacca's MSc thesis).

On other topologies the technique fails because we don't know how to locate the empty nodes!

$\mathcal{GOSSIP}$ R.W.s on non-Complete Graphs

The analysis for the complete graph can still be applied *locally* provided that the minimum degree is αn for some constant $\alpha > 0$ (G. Scornavacca's MSc thesis).

On other topologies the technique fails because we don't know how to locate the empty nodes!

Open Problems:
Maximum load on regular graphs?
Maximum load on the ring?

Part 3: Stabilizing Almost-Consensus

1. Majority Consensus
   (a) 3-Majority (take I)
   (b) Undecided-State
2. Congestion of $\mathcal{GOSSIP}$ random walks
3. Stabilizing Consensus
   (a) 3-Majority (take II)

Stabilizing Almost-Consensus

A *stabilizing almost-consensus* protocol guarantees, for some $\gamma < 1$
From any initial conf., in finite number of rounds, w.h.p. the system reaches a family of conf.s where $n - \mathcal{O}(n^{\gamma})$ nodes hold the same opinion (*almost agreement*), which was held in the initial conf. (*almost validity*), and the convergence hold w.h.p. for any polynomial number of rounds (*almost stability*).

Stabilizing Almost-Consensus

A *stabilizing almost-consensus* protocol guarantees, for some $\gamma < 1$
From any initial conf., in finite number of rounds, w.h.p. the system reaches a family of conf.s where $n - \mathcal{O}(n^\gamma)$ nodes hold the same opinion (*almost agreement*), which was held in the initial conf. (*almost validity*), and the convergence hold w.h.p. for any polynomial number of rounds (*almost stability*).

No termination!

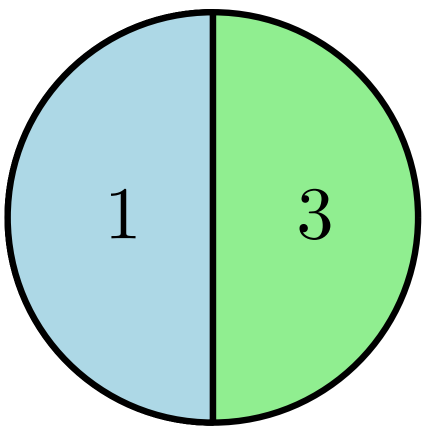
Failed Attempts: 3-Median

Theorem (Doerr et al. SPAA '11). For any $\sqrt{n}$-bounded adversary,in $\mathcal{O}(\log m \cdot \log\log n + \log n)$ time the 3-median rule computes w.h.p. an almost stable value between the $(n/2 - c\sqrt{nlogn})$-largest and the $(n/2 + c\sqrt{nlogn})$- largest of the initial values.

Failed Attempts: 3-Median

Theorem (Doerr et al. SPAA '11). For any $\sqrt{n}$-bounded adversary,in $\mathcal{O}(\log m \cdot \log\log n + \log n)$ time the 3-median rule computes w.h.p. an almost stable value between the $(n/2 - c\sqrt{nlogn})$-largest and the $(n/2 + c\sqrt{nlogn})$- largest of the initial values.

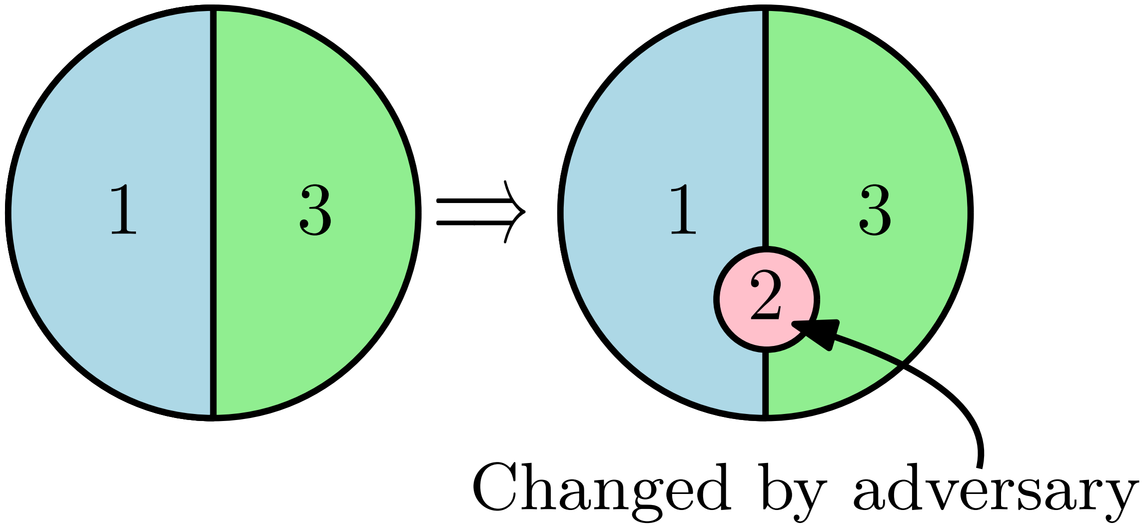
No almost validity



Failed Attempts: 3-Median

Theorem (Doerr et al. SPAA '11). For any $\sqrt{n}$-bounded adversary,in $\mathcal{O}(\log m \cdot \log\log n + \log n)$ time the 3-median rule computes w.h.p. an almost stable value between the $(n/2 - c\sqrt{nlogn})$-largest and the $(n/2 + c\sqrt{nlogn})$- largest of the initial values.

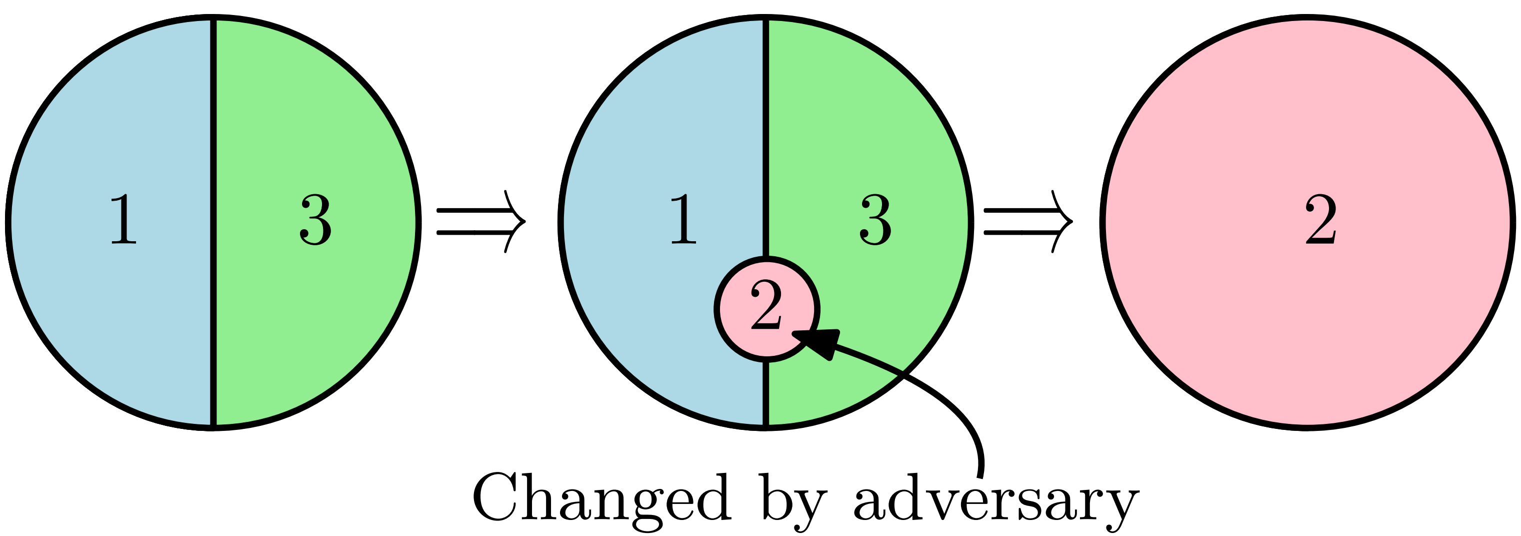
No almost validity



Failed Attempts: 3-Median

Theorem (Doerr et al. SPAA '11). For any $\sqrt{n}$-bounded adversary,in $\mathcal{O}(\log m \cdot \log\log n + \log n)$ time the 3-median rule computes w.h.p. an almost stable value between the $(n/2 - c\sqrt{nlogn})$-largest and the $(n/2 + c\sqrt{nlogn})$- largest of the initial values.

No almost validity



Part 2-a: 3-Majority (take II)

1. Majority Consensus
   (a) 3-Majority (take I)
   (b) Undecided-State
2. Congestion of $\mathcal{GOSSIP}$ random walks
3. Stabilizing Consensus
   (a) 3-Majority (take II)

3-Majority without Bias [SODA '16]

What if we start from any initial configuration, i.e. there may be no initial bias?

3-Majority without Bias [SODA '16]

What if we start from any initial configuration, i.e. there may be no initial bias?

Theorem. Let $k \leq n^\alpha$, for a suitable constant $\alpha < 1$, and $F = \beta\sqrt{n}/(k^{\frac{5}{2}} \log n)$ for some constant $\beta > 0$. The 3-majority dynamics is a stabilizing almost-consensus protocol in the presence of any F-dynamic adversary and its convergence time is $\mathcal{O}((k^2\sqrt{\log n} + k \log n)(k + \log n))$, w.h.p.

What's the Problem without Bias?

Lemma 2. Let 1 be the plurality opinion, then

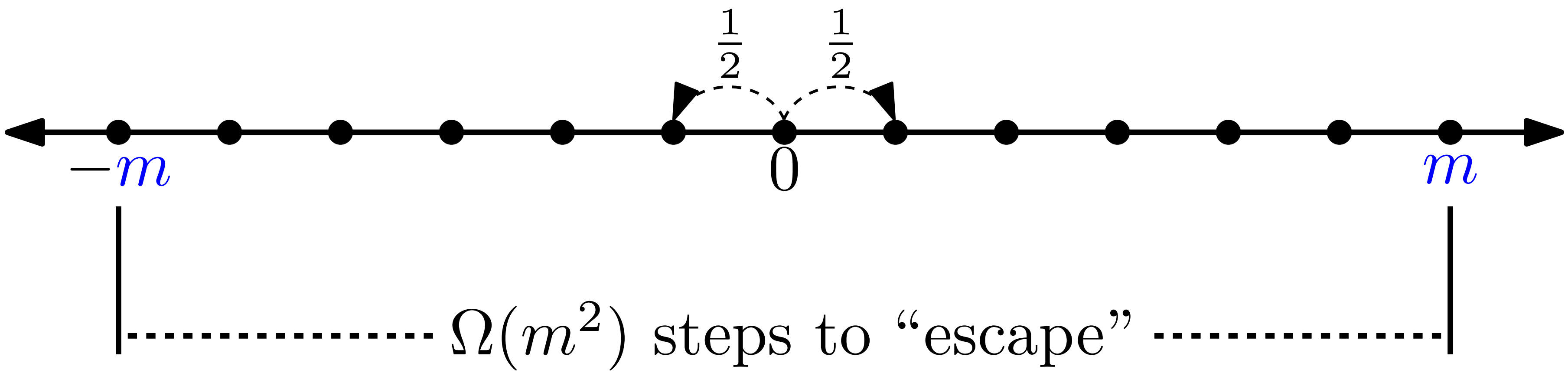
$$\mu_1 - \mu_j \geqslant s(\mathbf{c}) \left(1 + \frac{c_1}{n} \left(1 - \frac{c_1}{n}\right)\right).$$

Proof.

$$\begin{aligned}
\mu_1 - \mu_j \geqslant \mu_1 - \mu_2 &= (c_1 - c_2) + \frac{\left(c_1^2 - c_2^2\right)}{n} - \frac{c_1 - c_2}{n^2} \sum_{h \in k} c_h^2 \\
&= s(\mathbf{c}) \left(1 + \frac{c_1 + c_2}{n} - \boxed{\frac{1}{n^2} \sum_{h \in k} c_h^2}\right) \\
&\geqslant s(\mathbf{c}) \left(1 + \frac{c_1 + c_2}{n} - \boxed{\frac{c_1^2 + n c_2}{n^2}}\right) \\
&= s(\mathbf{c}) \left(1 + \frac{c_1}{n} \left(1 - \frac{c_1}{n}\right)\right).
\end{aligned}$$

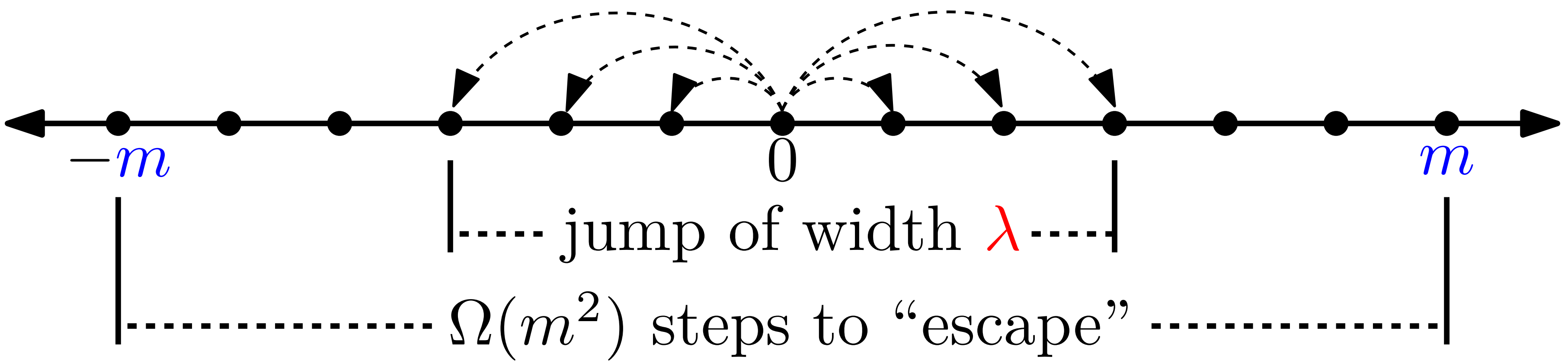
"Unbiased" Analysis

Symmetry Breaking



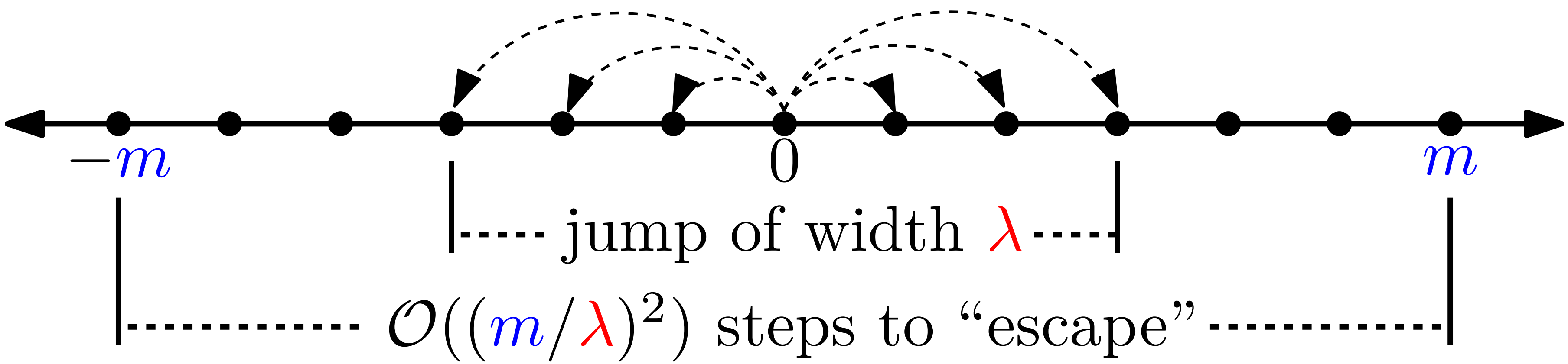
“Unbiased” Analysis

Symmetry Breaking



“Unbiased” Analysis

Symmetry Breaking



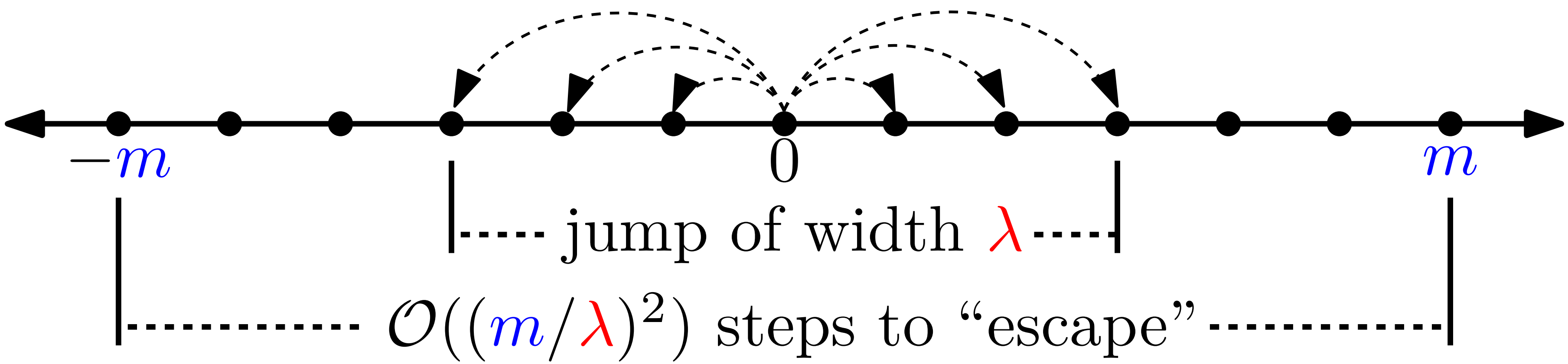
Lemma. $\{X_t\}_t$ a Markov chain with finite state space Ω, $f : \Omega \to \mathbf{N}$, $\{Y_t\}_t$ the stochastic process $Y_t = f(X_t)$, $m \in N$ a “target value” and $\tau = \inf\{t \in \mathbb{N} : Y_t \geq m\}$ the r.v. of the first time Y_t surpasses m. Assume that, $\forall x \in \Omega$ with $f(x) \leq m - 1$, it holds

1. (Positive drift). $\mathbf{E}[Y_{t+1} \mid X_t = x] \geqslant f(x) + \lambda$ for some $\lambda > 0$
2. (Bounded jumps). $\Pr Y_\tau \geq \alpha m \leq \alpha m/n$, for some $\alpha > 1$.

Then, $\forall x \in \Omega$, it holds $\mathbf{E}[\tau] \leq 2\alpha \frac{m}{\lambda}$.

"Unbiased" Analysis

Symmetry Breaking



Lemma. Let $\mathbf{c}$ be any configuration with j supported opinions. Within $t = \mathcal{O}\left(j^2 \log^{1/2} n\right)$ rounds it holds that

$$\Pr(\exists i \text{ such that } C_i^{(t)} \leq n/j - \sqrt{jn \log n}) \geq \frac{1}{2}$$

"Unbiased" Analysis

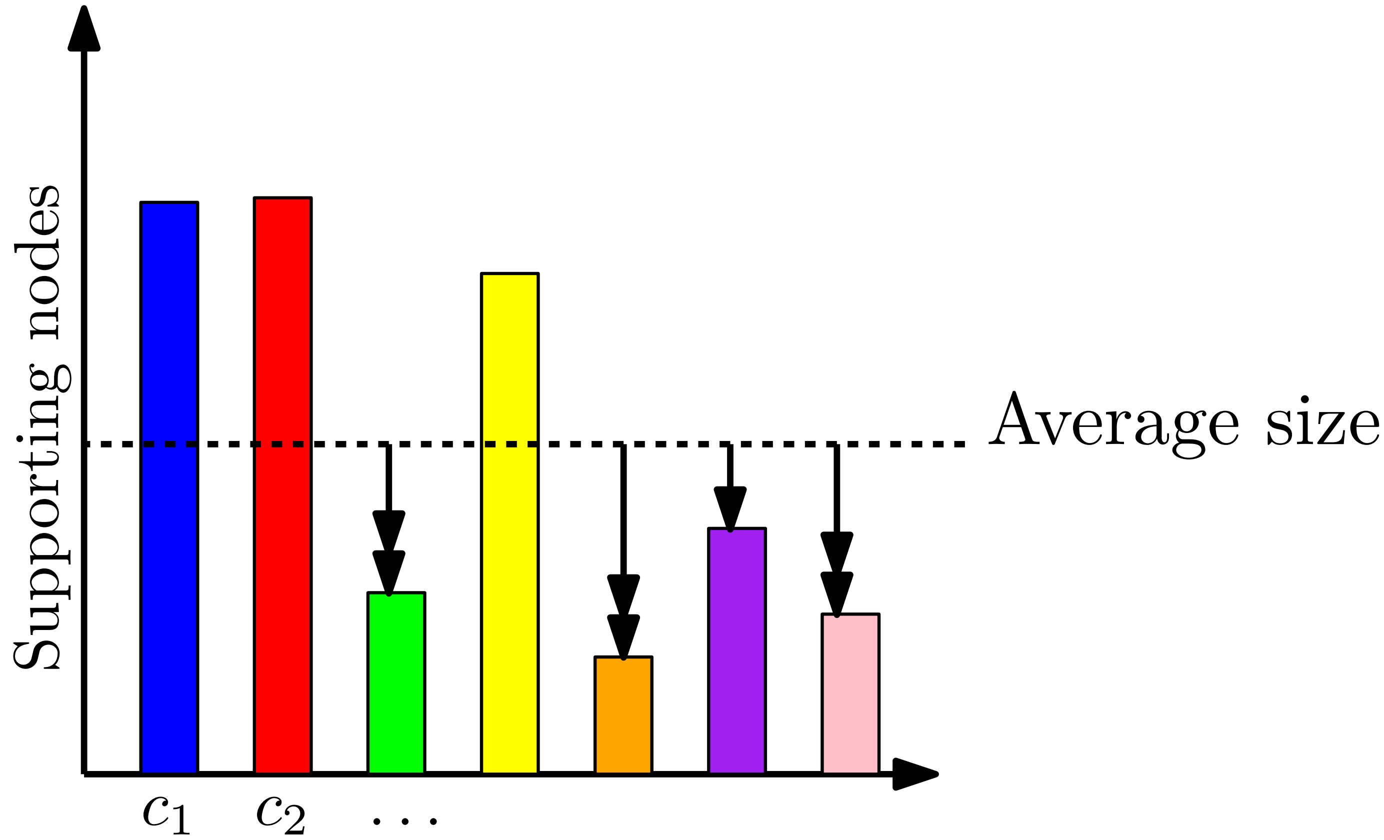
Lemma. Let $\mathbf{c}$ be the conf. at round t with j supported opinions. For any opinion i it holds,

$$\mathbf{E}[C_i^{(t+1)} \mid \mathbf{C}^{(t)} = \mathbf{c}] \leq c_i \left(1 + \frac{c_i}{n} - \frac{1}{j}\right).$$

"Unbiased" Analysis

Lemma. Let $\mathbf{c}$ be the conf. at round t with j supported opinions. For any opinion i it holds,

$$\mathbf{E}[C_i^{(t+1)} \mid \mathbf{C}^{(t)} = \mathbf{c}] \leq c_i \left(1 + \frac{c_i}{n} - \frac{1}{j}\right).$$



“Unbiased” Analysis

Lemma. Let $\mathbf{c}$ be any conf. with $j \leq n^{1/3-\varepsilon}$ supported opinions ($\forall \varepsilon > 0$ const), and such that an opinion i exists with $c_i \leq n/j - \sqrt{jn \log n}$. Within $t = \mathcal{O}(j \log n)$ rounds opinion i becomes $\mathcal{O}\left(j^2 \log n\right)$ w.h.p.

$$c_i \leq n/j - \sqrt{jn \log n} \xrightarrow[\text{w.h.p.}]{t = \mathcal{O}(j \log n)} c_i = \mathcal{O}(j^2 \log n)$$

“Unbiased” Analysis

Lemma. Let $\mathbf{c}$ be any conf. with $j \leq n^{1/3-\varepsilon}$ supported opinions ($\forall \varepsilon > 0$ const), and such that an opinion i exists with $c_i \leq n/j - \sqrt{jn \log n}$. Within $t = \mathcal{O}(j \log n)$ rounds opinion i becomes $\mathcal{O}\left(j^2 \log n\right)$ w.h.p.

$$c_i \leq n/j - \sqrt{jn \log n} \xrightarrow[\text{w.h.p.}]{t = \mathcal{O}(j \log n)} c_i = \mathcal{O}(j^2 \log n)$$

Lemma. Let $\mathbf{c}$ be any conf. with $j \leq n^{1/3-\varepsilon}$ supported opinions ($\forall \varepsilon > 0$ const), and such that an opinion i exists with $c_i \leq n/(2j)$. Within $t = \mathcal{O}(j \log n)$ rounds opinion i disappears with probability at least $1/2$.

$$c_i \leq n/(2j) \xrightarrow[\text{with prob. } \geq 1/2]{t = \mathcal{O}(j \log n)} c_i = 0$$

Stabilizing Consensus on not-Complete Graphs

Open Problems

Stabilizing consensus on random graphs?

Stabilizing consensus on expander graphs?

Stabilizing Consensus on not-Complete Graphs

Open Problems

Stabilizing consensus on random graphs?

Stabilizing consensus on expander graphs?

Theorem (Cooper et al. ICALP '14).
Let G be a random d-regular graph with initial opinions A and B. There is an absolute constant K (independent of d) such that, provided

$$\frac{|A-B|}{n} \geq K\sqrt{\frac{d}{n}+\frac{1}{d}},$$

two-sample voting is completed in $O(\log n)$ steps a.a.s., and the winner is the opinion with the initial majority.

Stabilizing Consensus on not-Complete Graphs

Open Problems

Stabilizing consensus on random graphs?

Stabilizing consensus on expander graphs?

Theorem (Cooper et al. ICALP '14).
Let G be a d-regular graph with initial opinions A and B, $1 = \lambda_1 \geq \lambda_2 \geq \cdots \lambda_n \geq -1$ be the eigenvalalues of the transition matrix of the r.w. on G, and $\lambda = \lambda_G = \max\{|\lambda_2|, |\lambda_n|\}$. For some const. K (indep. of d and λ_G), provided

$$|A - B|/n \;\geq\; K\lambda_G,$$

a.a.s. two-sample voting is completed in $O(\log n)$ steps and winner is the initial majority.

Stabilizing Consensus on not-Complete Graphs

Open Problems

Stabilizing consensus on random graphs?

Stabilizing consensus on expander graphs?

Expander Mixing Lemma (Alon, Chung).
Let $G = (V, E)$ be a d-regular n-vertex graph. Let $1 = \lambda_1 \geq \lambda_2 \geq \cdots \lambda_n \geq -1$ be the eigenvalues of the transition matrix of the random walk on G, and let $\lambda = \lambda_G = \max\{|\lambda_2|, |\lambda_n|\}$. Then for all $S, T \subseteq V$,

$$\left| E(S,T) - \frac{dST}{n} \right| \leq \lambda d \sqrt{ST}.$$

Thank you!