Supervised Average Consensus in Anonymous Dynamic Networks

Dariusz R. Kowalski
Augusta Univ.

Miguel A. Mosteiro
Pace Univ.

SPAA 2021

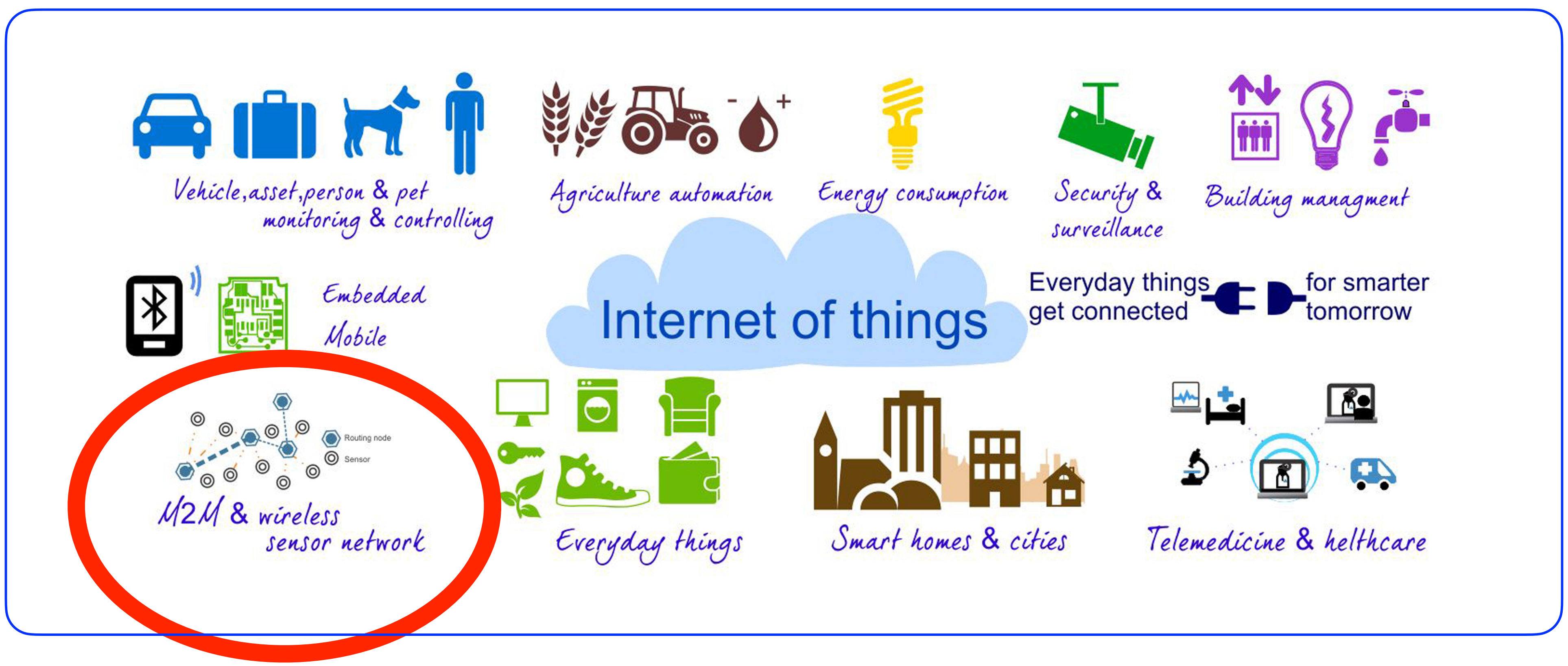
Anonymous Dynamic Networks

No node identifiers.

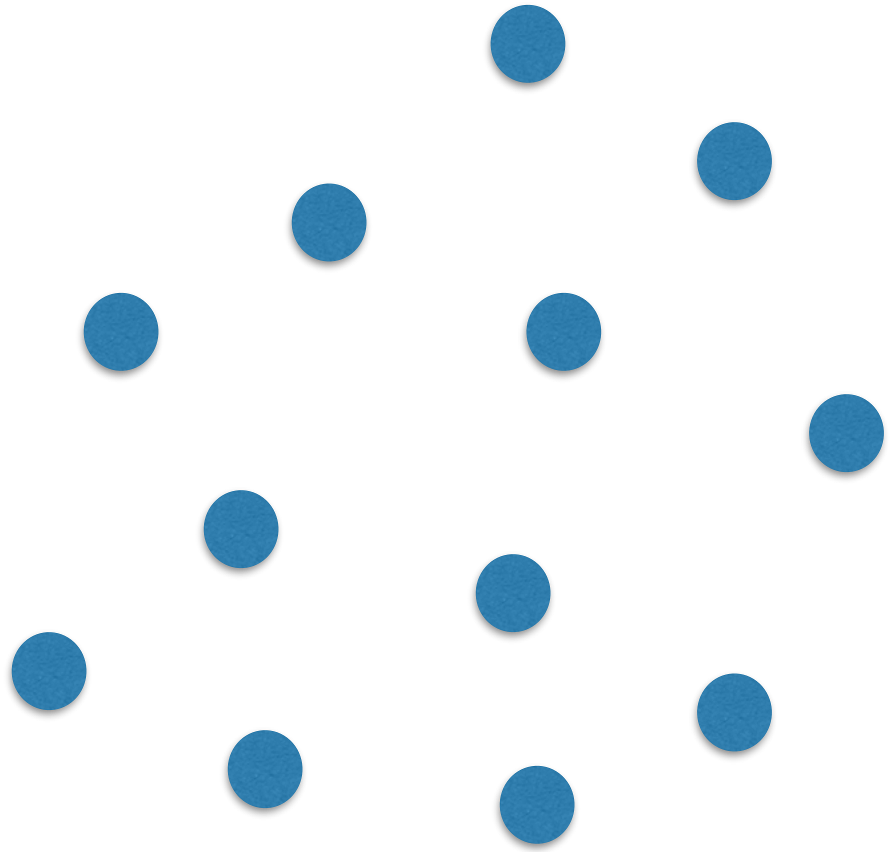
Due to massive number of nodes, low cost, etc.

Communication links change.

Due to mobility, failures, etc.

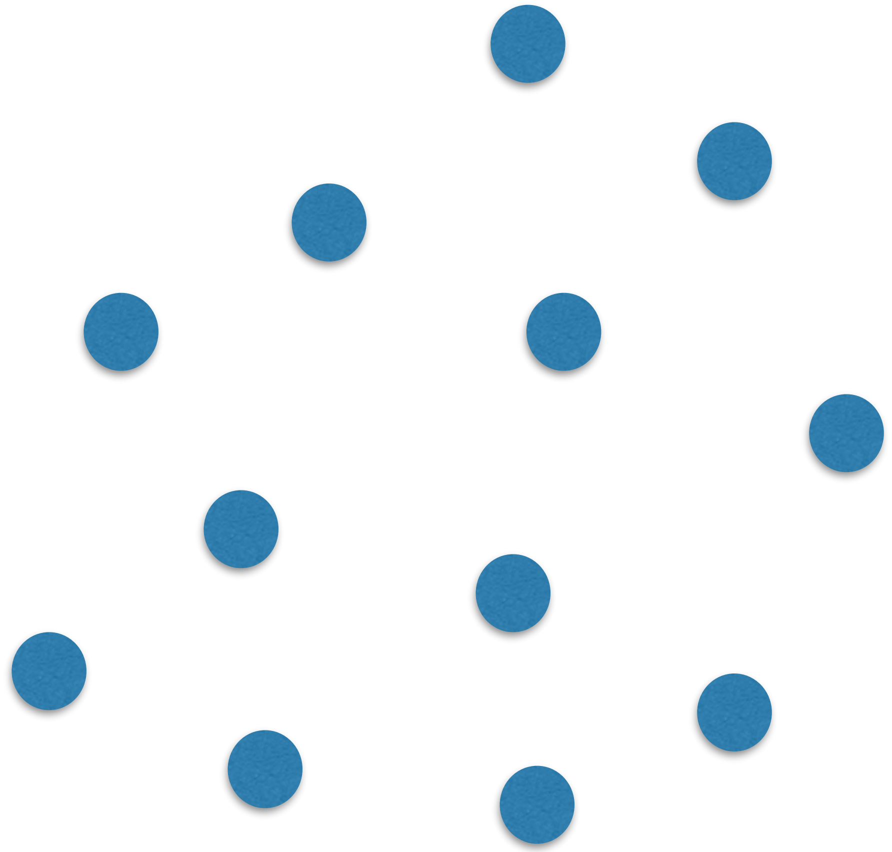


Anonymous Dynamic Networks



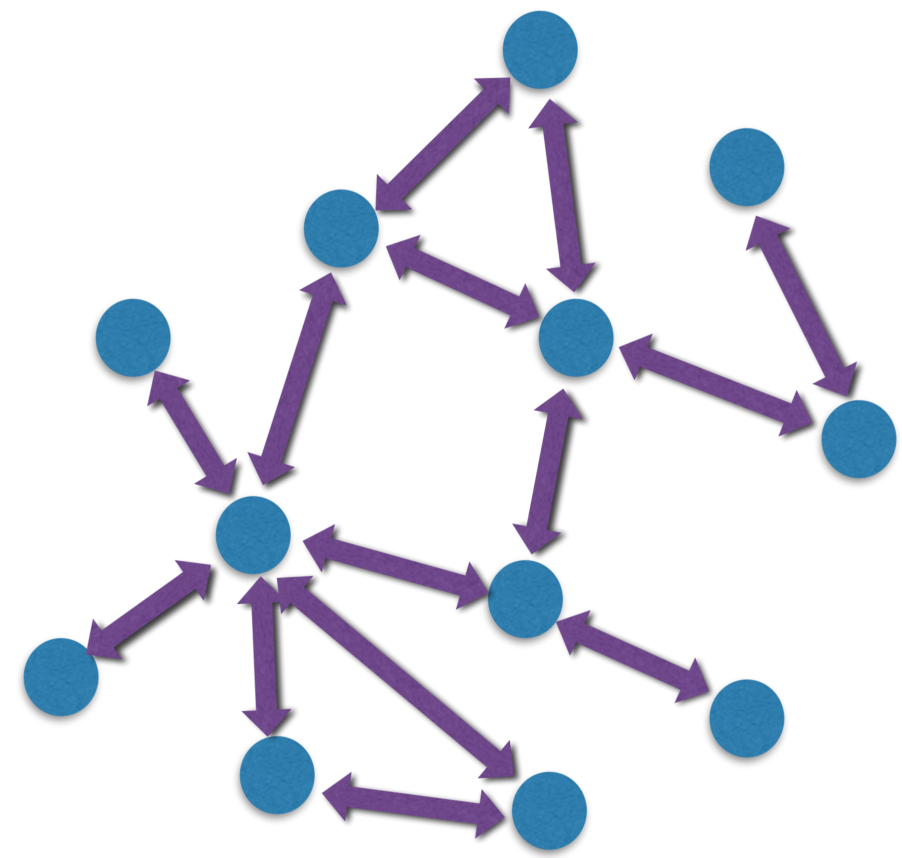
Anonymous Dynamic Networks

- **Fixed set of n nodes**
  - no identifiers or labels

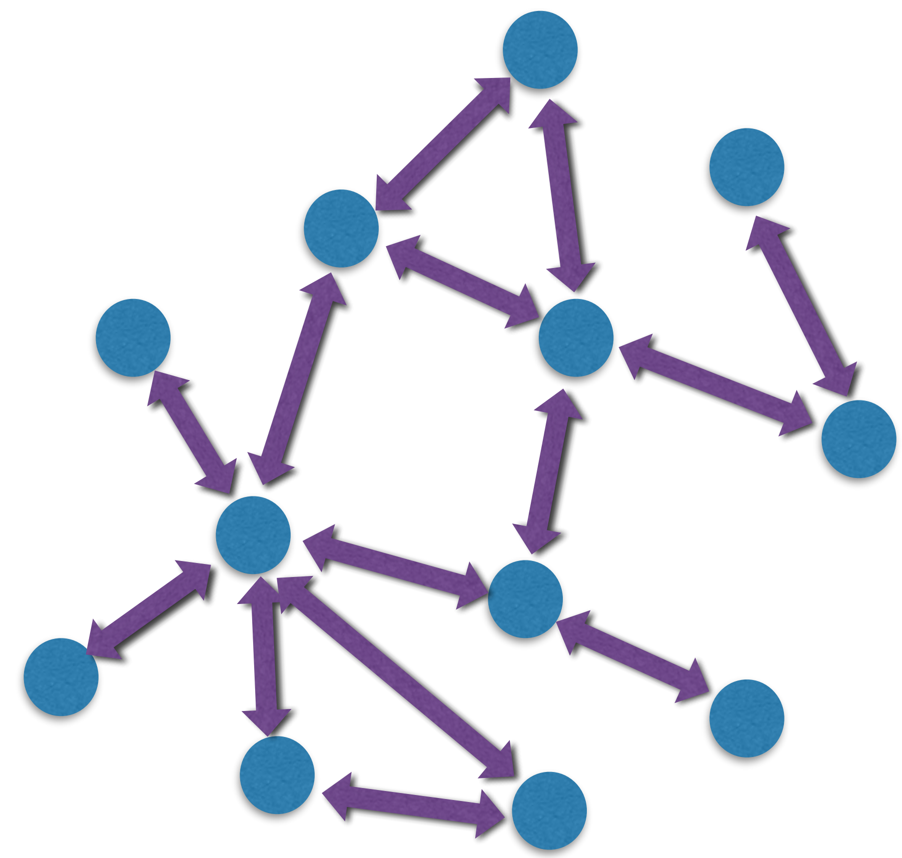


Anonymous Dynamic Networks

- **Fixed set of n nodes**
  - no identifiers or labels
- **Synchronous communication:** At each round
  - a node broadcasts a message to its neighbors
  - receives the messages of its neighbors
  - executes some local computation



Anonymous Dynamic Networks



- **Fixed set of n nodes**
  - no identifiers or labels
- **Synchronous communication:** At each round
  - a node broadcasts a message to its neighbors
  - receives the messages of its neighbors
  - executes some local computation
- **Topology:**
  - at each round the network is connected
  - dynamicity:
    - » <u>standard ADN</u>: links change arbitrarily → too pessimistic
    - » <u>in practice</u>: good expansion is the norm rather than the exception!!
    - » <u>for this work</u>: known lower bound on

Isoperimetric numbers: $i(G) = \min_{\substack{X : X \subset V, \\ |X| \leq |V|/2}} \frac{|\partial X|}{|X|}$,

$G = (V, E)$ each network-topology graph,
$\partial X \subseteq E$ set of links between X and $V \backslash X$.

Network Average Consensus

Fault-tolerant Consensus:

« Given a distributed system of n processors,

all agree on a value and stop »

Profusely studied in Distributed Computing.

Network Consensus:

« Given a network of n nodes, each holding an input value x_i,

every node obtains same $f(x_1, x_2, \ldots, x_n)$ and stop »

Profusely studied in Systems and Control Theory.

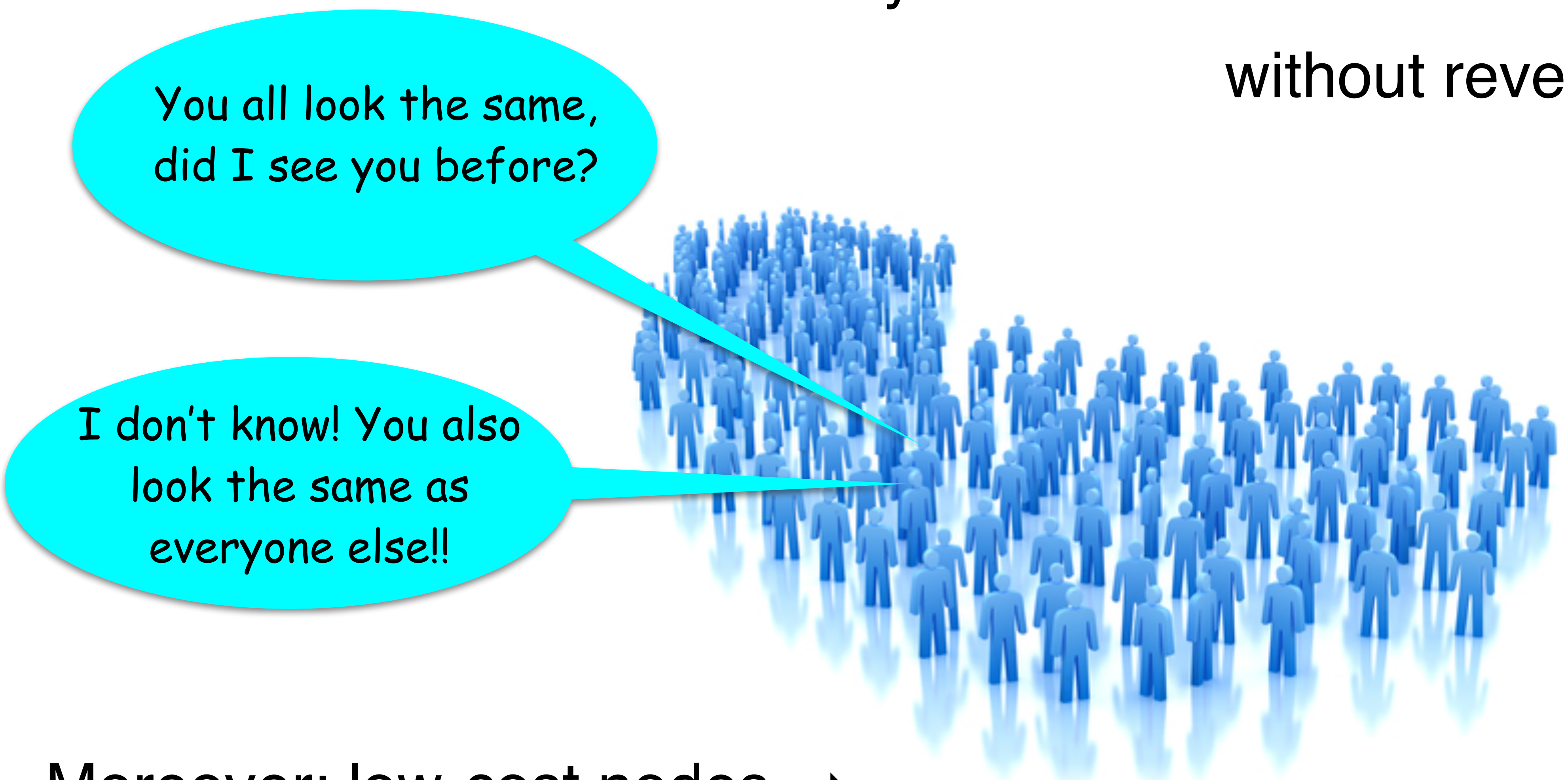
Popular functions: average, sum, maximum, etc .

Average Consensus in ADNs

How to reach consensus

in a dynamic crowd

without revealing identity?



Moreover: low-cost nodes →

start-up and late failures may occur →

n may be unknown!

Contributions

We study:

Network Average Consensus in Anonymous Dynamic Networks

- » unknown number of nodes
- » known (lower bound on) isoperimetric numbers

We show:

Contributions

We study:

Network Average Consensus in Anonymous Dynamic Networks

- » unknown number of nodes
- » known (lower bound on) isoperimetric numbers

We show:

- Randomized Network Average Consensus not possible without known number $\ell > 0$ of distinguished nodes, we call them *supervisors*

Given that: same applies to Deterministic Counting ≡ Average (prev. known), the claim is true for all algorithms.

Contributions

We study:

Network Average Consensus in Anonymous Dynamic Networks

» unknown number of nodes

» known (lower bound on) isoperimetric numbers

We show:

- Randomized Network Average Consensus not possible without known number $\ell > 0$ of distinguished nodes, we call them *supervisors*

Given that: same applies to Deterministic Counting ≡ Average (prev. known), the claim is true for all algorithms.

- Network Average Consensus Algorithm with $\ell > 0$ supervisors:

  isoperimetric Scalable Coordinated Anonymous Local Aggregation (iSCALA)

  - based on Methodical multi-Counting (prev. known) but
  - designed to use known isoperimetric dynamicity to improve time complexity
  - MMC (and others) inefficient for (practical) good expansion networks
  - iSCALA intrinsically adapts to changes

Contributions

<u>We study</u>:

Network Average Consensus in Anonymous Dynamic Networks

» unknown number of nodes

» known (lower bound on) isoperimetric numbers

<u>We show</u>:

- Randomized Network Average Consensus not possible without known number $\ell > 0$ of distinguished nodes, we call them *supervisors*

Given that: same applies to Deterministic Counting ≡ Average (prev. known), the claim is true for all algorithms.

- Network Average Consensus Algorithm with $\ell > 0$ supervisors:

  isoperimetric Scalable Coordinated Anonymous Local Aggregation (iSCALA)

- Analysis for adversarial and various stochastic topologies
- Thorough simulations

Impossibility

Theorem 4.1. *For any constant* $0 < c < 1$*, there exists an ADN with* $\ell = 0$ *such that there is no randomized algorithm that, with probability at least* c*, solves the Network Average Consensus Problem, even knowing a lower estimate of the isoperimetric number.*

Corollary 4.2. *For any constant* $0 < c < 1$ *and any* $\ell > 0$*, there exists an ADN with* ℓ *supervisor nodes such that, if* ℓ *is unknown to the network nodes, there is no randomized algorithm that, with probability at least* c*, solves the Network Average Consensus Problem, even knowing a lower estimate of the isoperimetric number.*

Proved showing a carefully designed network that has constant isoperimetric number globally, and also locally.

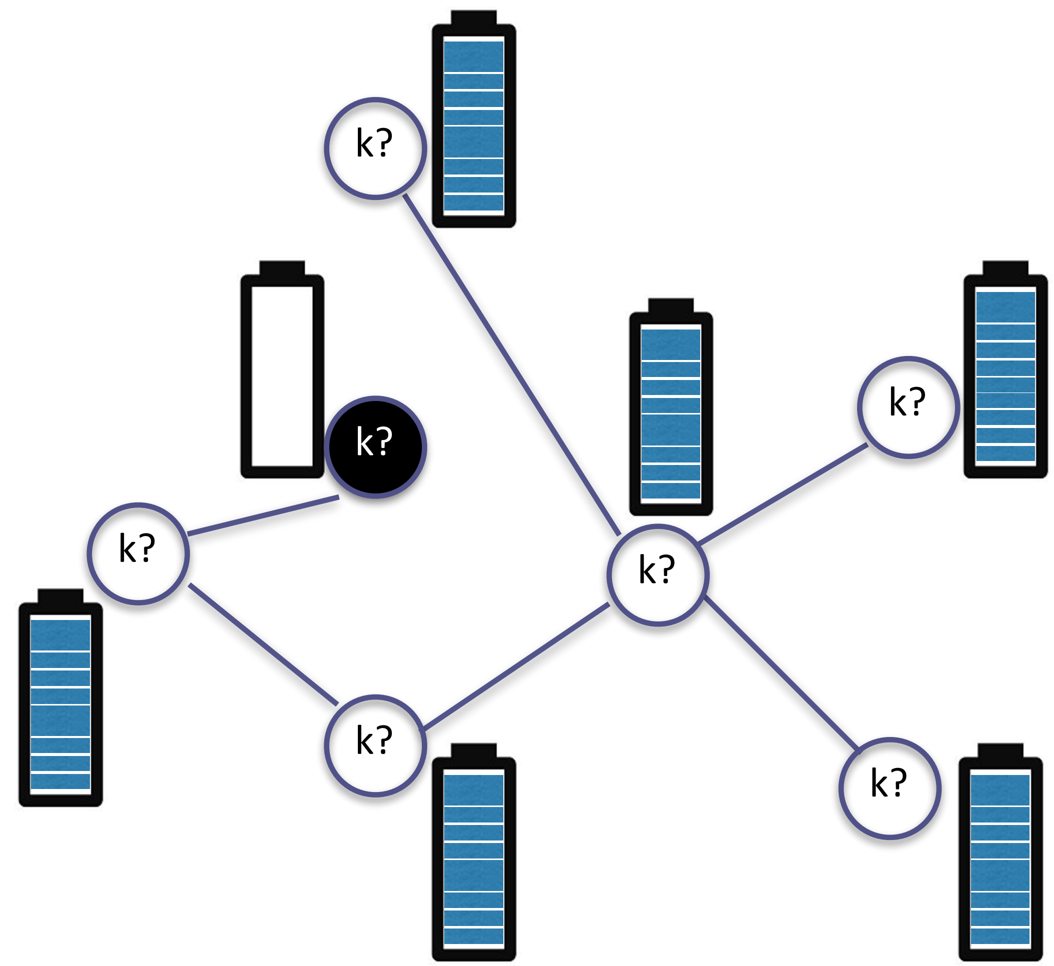
Then showing that, with constant probability, any algorithm reaches a termination configuration locally before receiving global information.

iSCALA structure

epochs:

- one for each estimate $k = \ell + 1, 2(\ell + 1), 4(\ell + 1), \ldots$
- initially, "potential" value: $\Phi_{supervised} = \ell$, $\Phi_{supervisor} = 0$

Example with $\ell = 1$ supervisor and $n - \ell = 6$ supervised nodes



iSCALA structure

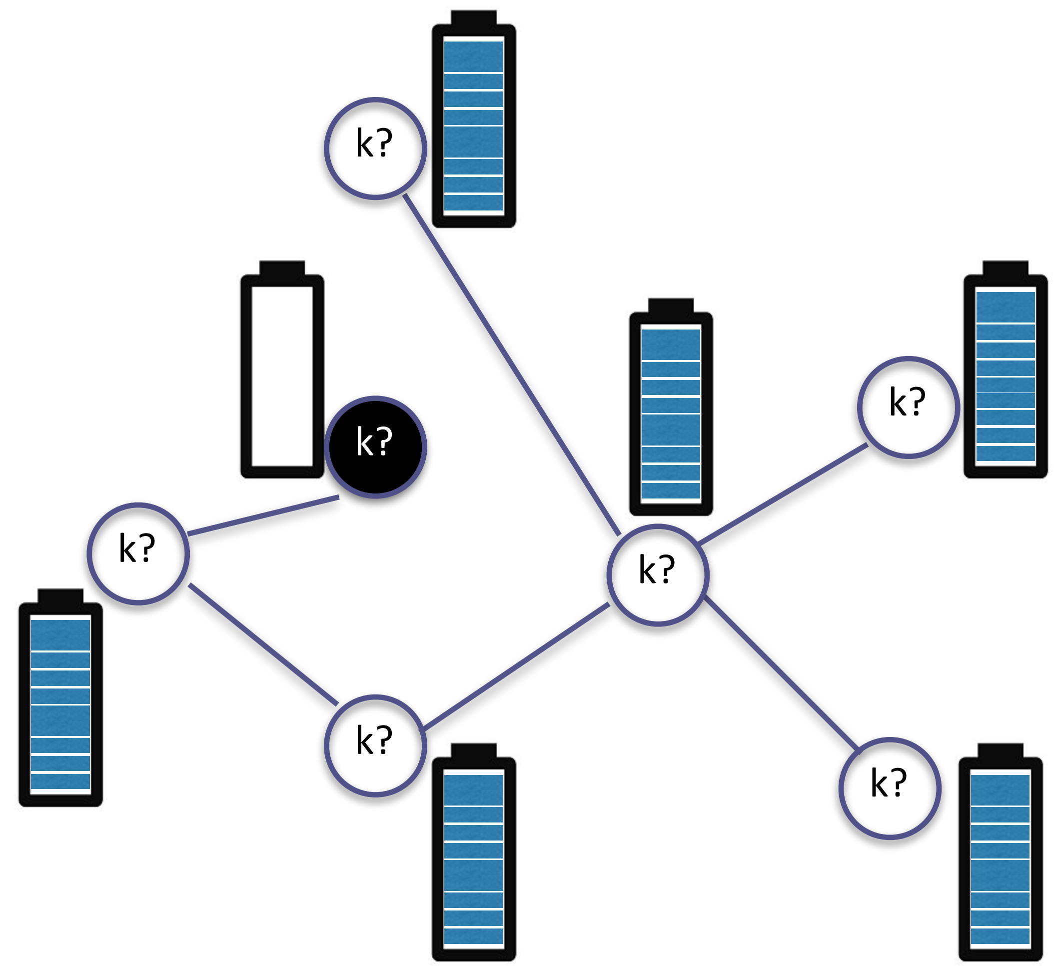
epochs:

- one for each estimate $k = \ell + 1, 2(\ell + 1), 4(\ell + 1), \ldots$
- initially, "potential" value: $\Phi_{supervised} = \ell$, $\Phi_{supervisor} = 0$

$p(k, i(G_T))$ phases:

(to let supervisors remove "enough" potential ρ)

Example with $\ell = 1$ supervisor and $n - \ell = 6$ supervised nodes



iSCALA structure

epochs:

- one for each estimate $k = \ell + 1, 2(\ell + 1), 4(\ell + 1), \ldots$
- initially, "potential" value: $\Phi_{supervised} = \ell$, $\Phi_{supervisor} = 0$

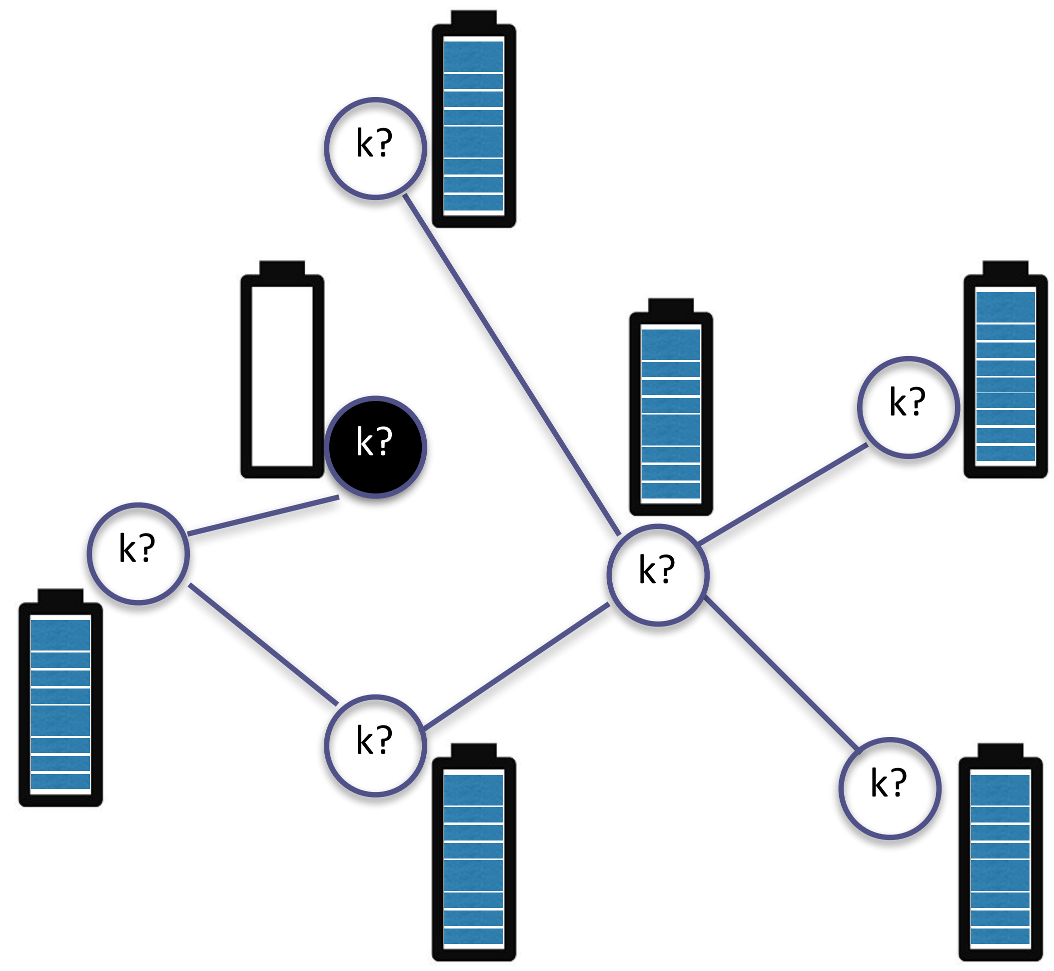
$p(k, i(G_T))$ phases:

(to let supervisors remove "enough" potential ρ)

$r(k, i(G_T))$ rounds:

(to "average" the current potentials Φ)

Example with $\ell = 1$ supervisor and $n - \ell = 6$ supervised nodes



epochs:

- one for each estimate $k = \ell + 1, 2(\ell + 1), 4(\ell + 1), \ldots$
- initially, "potential" value: $\Phi_{supervised} = \ell$, $\Phi_{supervisor} = 0$

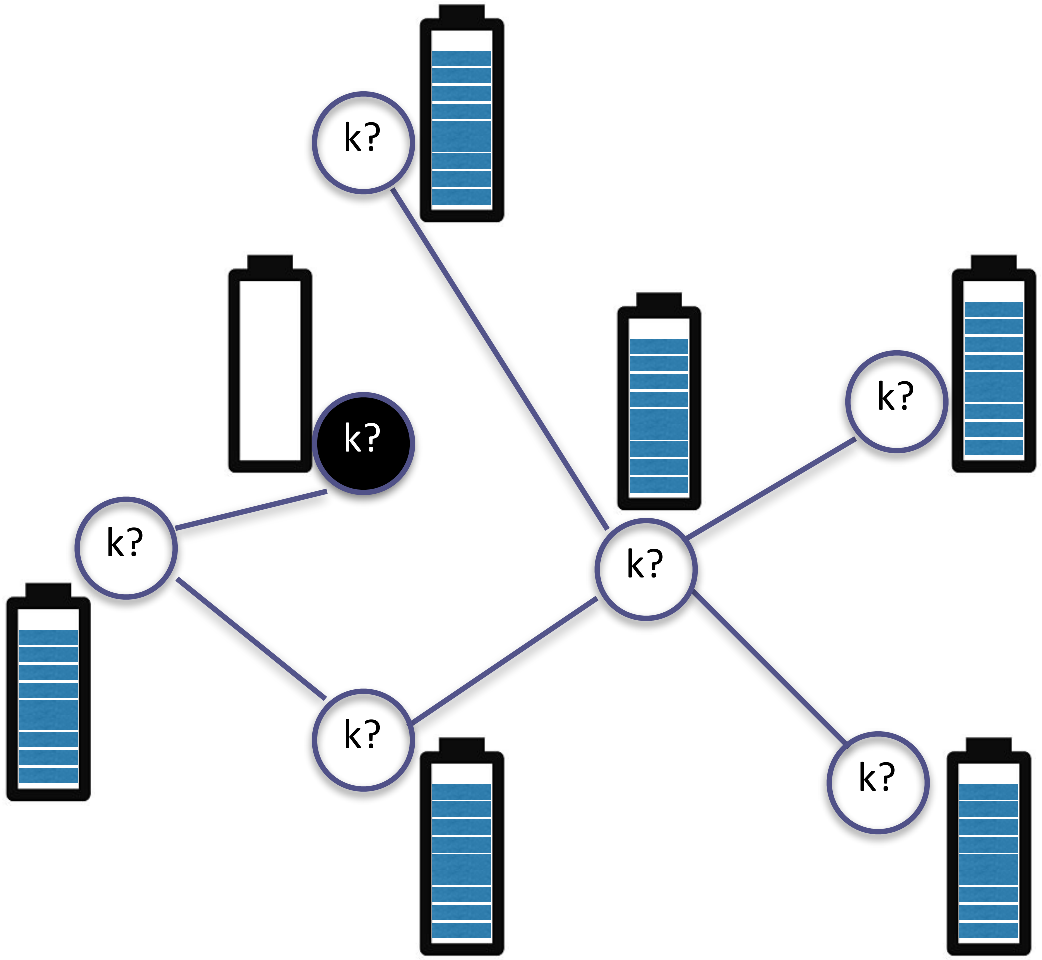
$p(k, i(G_T))$ phases:

(to let supervisors remove "enough" potential ρ)

$r(k, i(G_T))$ rounds:

(to "average" the current potentials Φ)

Example with $\ell = 1$ supervisor and $n - \ell = 6$ supervised nodes



$\rho =$

iSCALA Epoch Example

epochs:

- one for each estimate $k = \ell + 1, 2(\ell + 1), 4(\ell + 1), \ldots$
- initially, "potential" value:
  $\Phi_{supervised} = \ell,\ \Phi_{supervisor} = 0$

$p(k, i(G_T))$ phases:

(to let supervisors remove "enough" potential ρ)

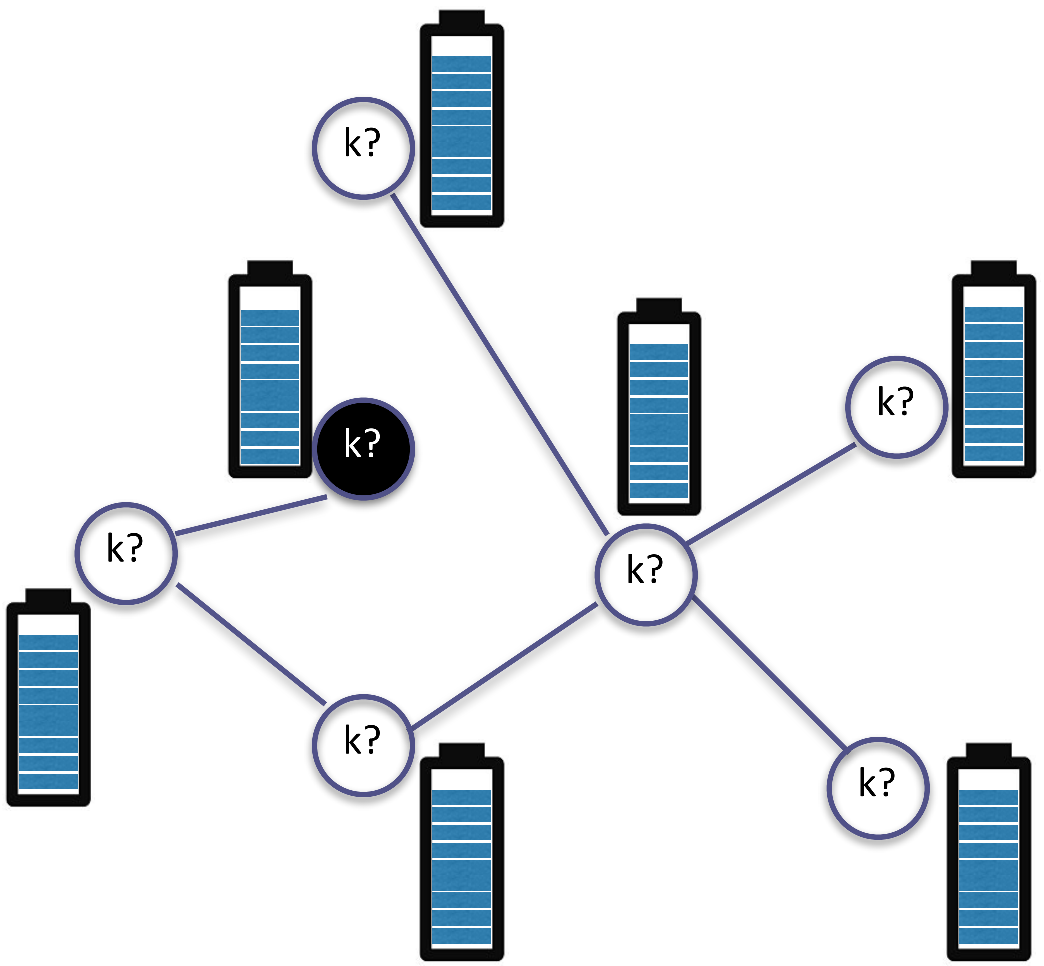
$r(k, i(G_T))$ rounds:

(to "average" the current potentials Φ)

mass distribution:

- broadcast Φ and receive neighbors' Φ_i
- $\Phi = \Phi + \sum_{i \in N} \frac{\Phi_i}{d(k)} - |N| \frac{\Phi}{d(k)}$

Example with $\ell = 1$ supervisor and $n - \ell = 6$ supervised nodes



$\rho =$

iSCALA Epoch Example

epochs:

- one for each estimate $k = \ell + 1, 2(\ell + 1), 4(\ell + 1), \ldots$
- initially, "potential" value: $\Phi_{supervised} = \ell,\ \Phi_{supervisor} = 0$

$p(k, i(G_T))$ phases:

(to let supervisors remove "enough" potential ρ)

$r(k, i(G_T))$ rounds:

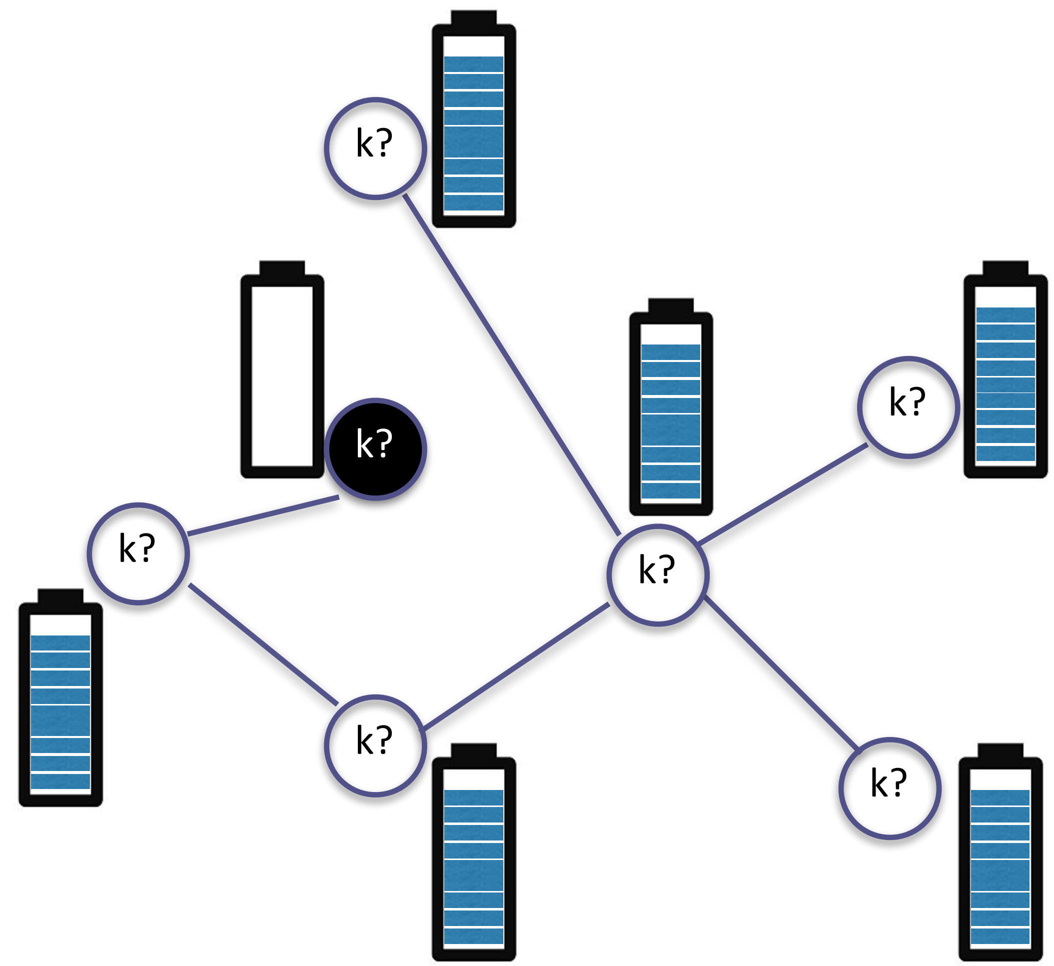
(to "average" the current potentials Φ)

mass distribution:

- broadcast Φ and receive neighbors' Φ_i
- $\Phi = \Phi + \sum_{i \in N} \frac{\Phi_i}{d(k)} - |N| \frac{\Phi}{d(k)}$

- supervisors "remove" their potential: $\rho = \rho + \Phi, \Phi = 0$

Example with $\ell = 1$ supervisor and $n - \ell = 6$ supervised nodes



$\rho =$

iSCALA Epoch Example

epochs:

- one for each estimate $k = \ell + 1, 2(\ell + 1), 4(\ell + 1), \ldots$
- initially, "potential" value: $\Phi_{supervised} = \ell$, $\Phi_{supervisor} = 0$

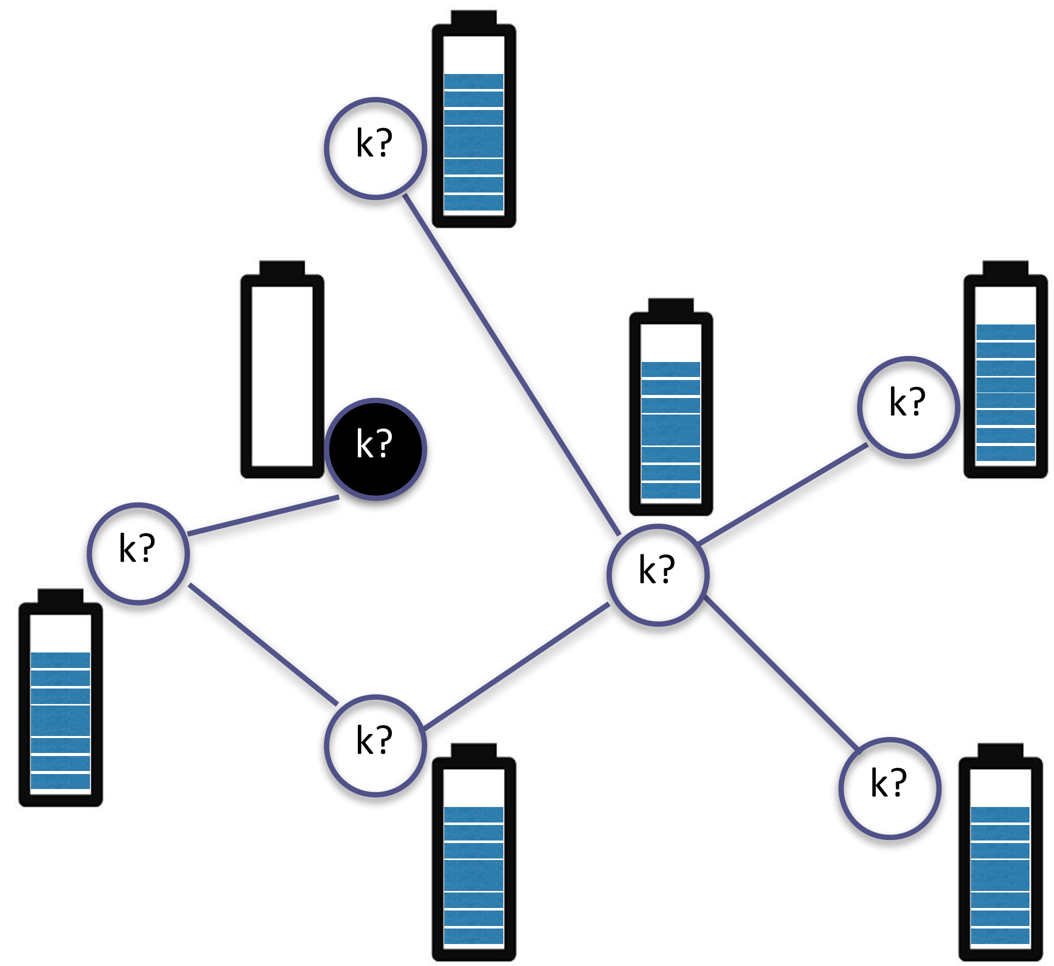
$p(k, i(G_T))$ phases:

(to let supervisors remove "enough" potential ρ)

$r(k, i(G_T))$ rounds:

(to "average" the current potentials Φ)

Example with $\ell = 1$ supervisor and $n - \ell = 6$ supervised nodes



$\rho =$

iSCALA Epoch Example

epochs:

- one for each estimate $k = \ell + 1, 2(\ell + 1), 4(\ell + 1), \ldots$
- initially, "potential" value: $\Phi_{supervised} = \ell$, $\Phi_{supervisor} = 0$

$p(k, i(G_T))$ phases:

(to let supervisors remove "enough" potential ρ)

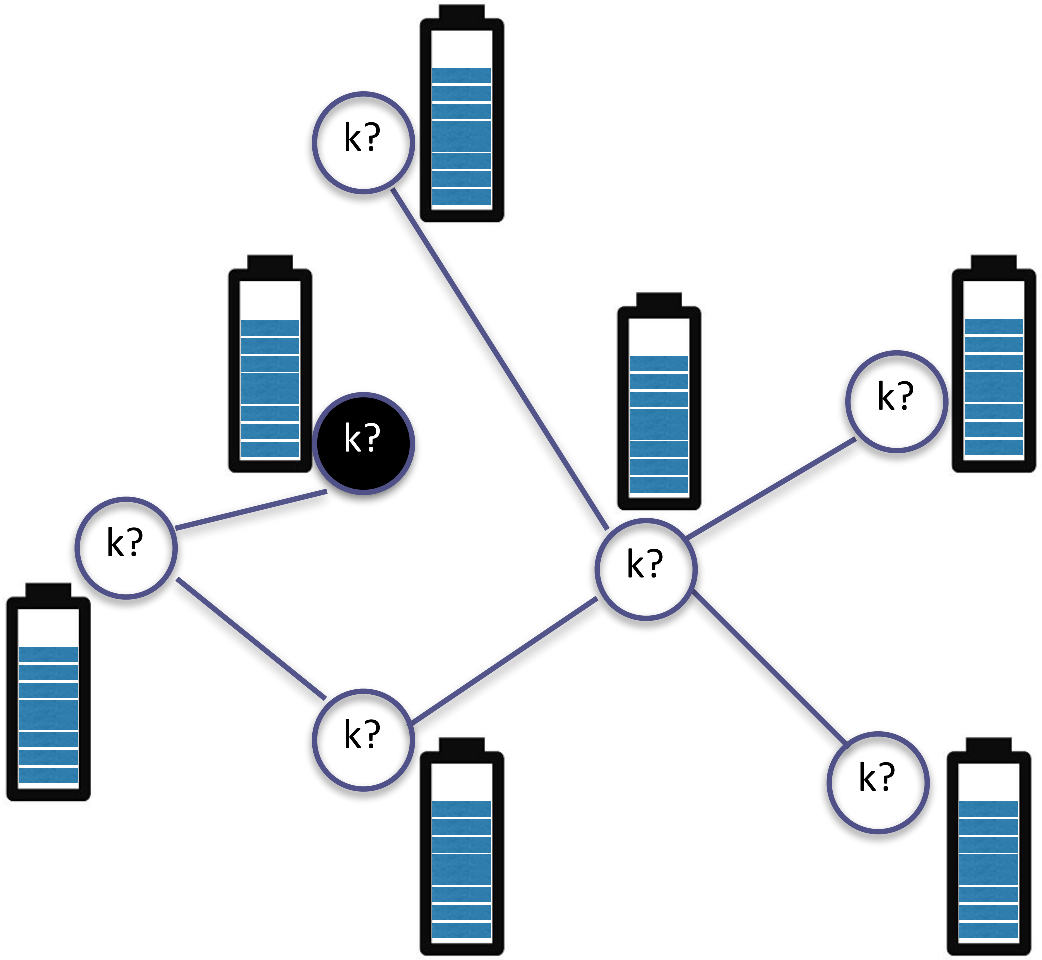
$r(k, i(G_T))$ rounds:

(to "average" the current potentials Φ)

mass distribution:

- broadcast Φ and receive neighbors' Φ_i
- $\Phi = \Phi + \sum_{i \in N} \frac{\Phi_i}{d(k)} - |N| \frac{\Phi}{d(k)}$

Example with $\ell = 1$ supervisor and $n - \ell = 6$ supervised nodes



$\rho =$

iSCALA Epoch Example

epochs:

- one for each estimate $k = \ell + 1, 2(\ell + 1), 4(\ell + 1), \ldots$
- initially, "potential" value: $\Phi_{supervised} = \ell,\ \Phi_{supervisor} = 0$

$p(k, i(G_T))$ phases:

(to let supervisors remove "enough" potential ρ)

$r(k, i(G_T))$ rounds:

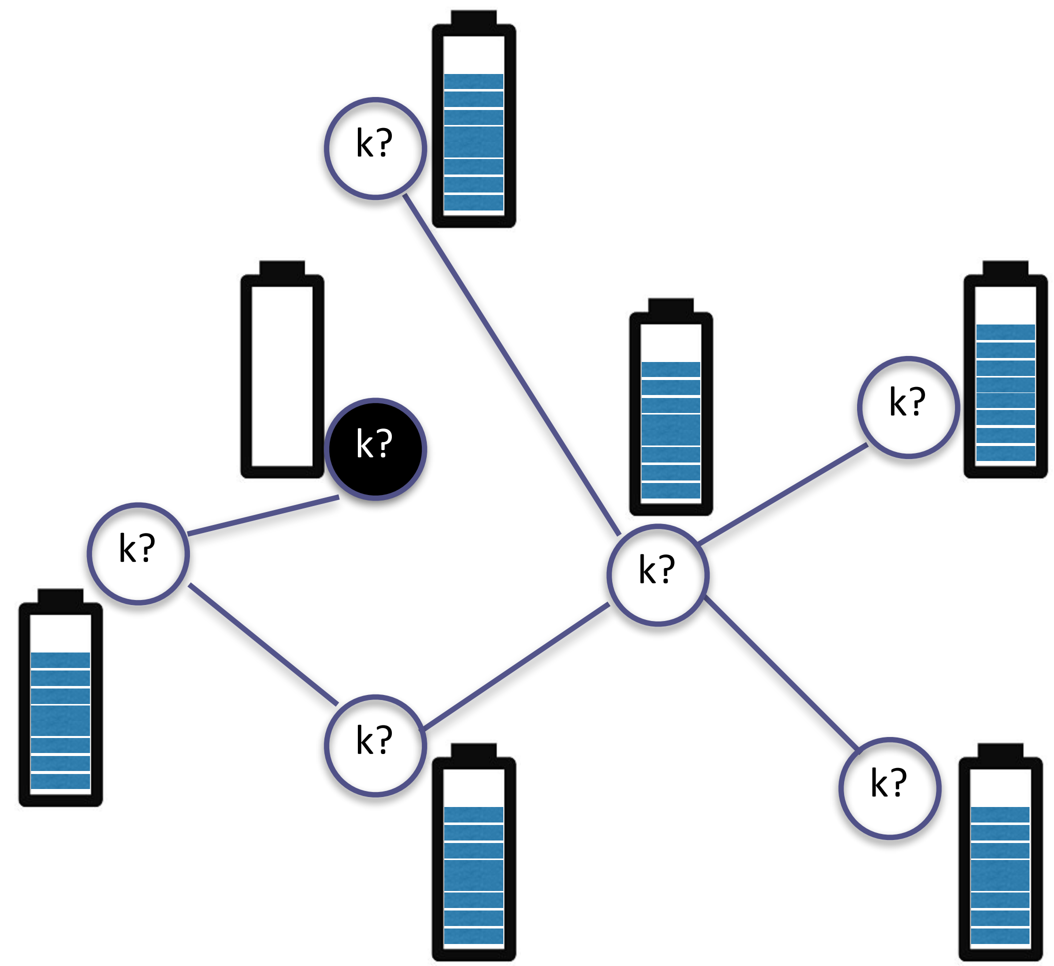
(to "average" the current potentials Φ)

mass distribution:

- broadcast Φ and receive neighbors' Φ_i
- $\Phi = \Phi + \sum_{i \in N} \frac{\Phi_i}{d(k)} - |N| \frac{\Phi}{d(k)}$

- supervisors "remove" their potential: $\rho = \rho + \Phi, \Phi = 0$

Example with $\ell = 1$ supervisor and $n - \ell = 6$ supervised nodes



$\rho =$

iSCALA Epoch Example

epochs:

- one for each estimate $k = \ell + 1, 2(\ell + 1), 4(\ell + 1), \ldots$
- initially, "potential" value:
  $\Phi_{supervised} = \ell, \; \Phi_{supervisor} = 0$

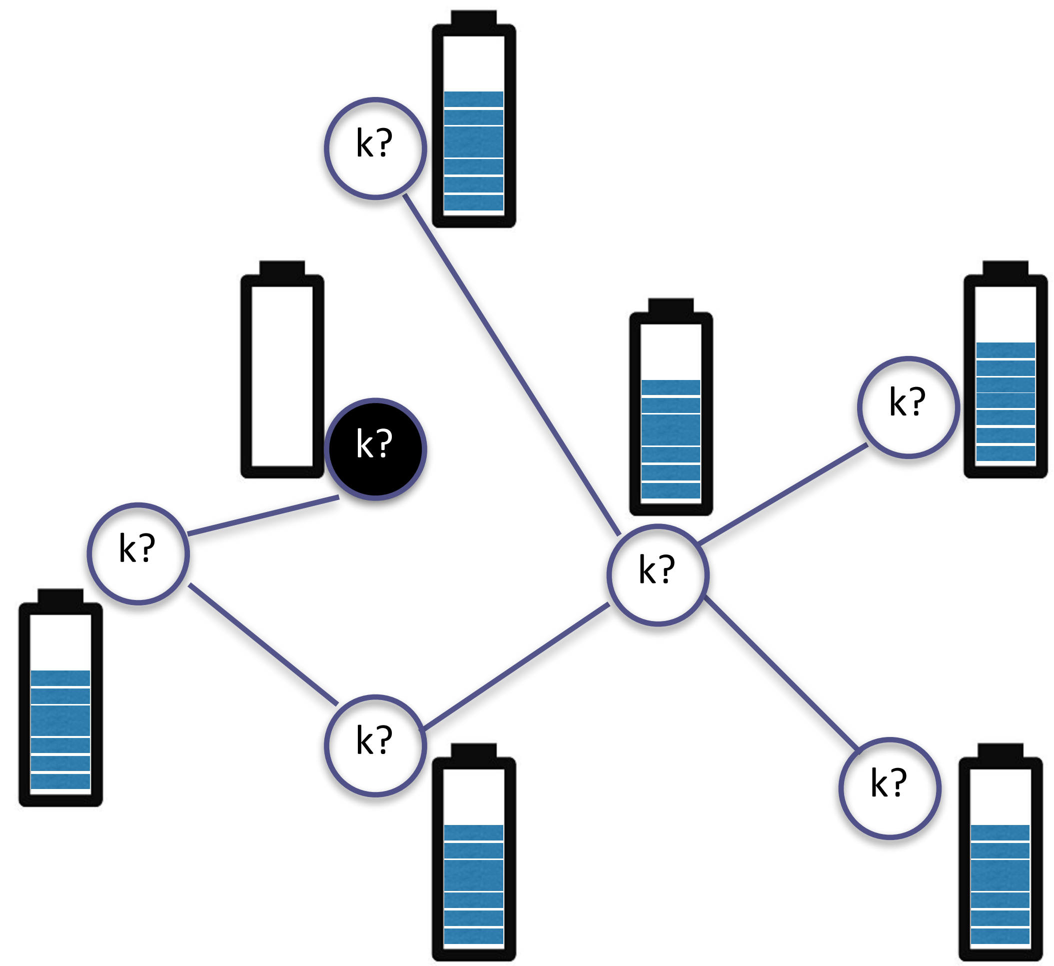
$p(k, i(G_T))$ phases:

(to let supervisors remove "enough" potential ρ)

$r(k, i(G_T))$ rounds:

(to "average" the current potentials Φ)

Example with $\ell = 1$ supervisor and $n - \ell = 6$ supervised nodes



$\rho =$

iSCALA Epoch Example

epochs:

- one for each estimate $k = \ell + 1, 2(\ell + 1), 4(\ell + 1), \ldots$
- initially, "potential" value: $\Phi_{supervised} = \ell$, $\Phi_{supervisor} = 0$

$p(k, i(G_T))$ phases:

(to let supervisors remove "enough" potential ρ)

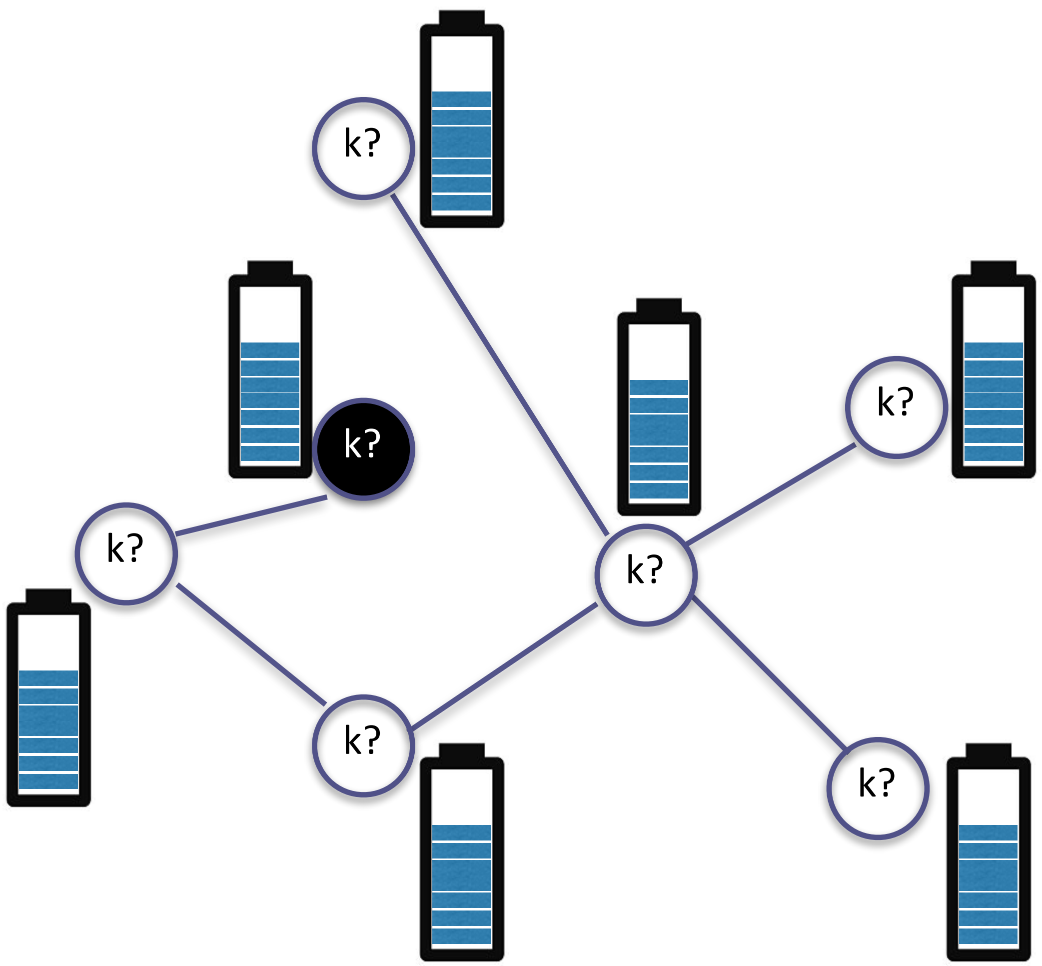
$r(k, i(G_T))$ rounds:

(to "average" the current potentials Φ)

mass distribution:

- broadcast Φ and receive neighbors' Φ_i
- $$\Phi = \Phi + \sum_{i \in N} \frac{\Phi_i}{d(k)} - |N| \frac{\Phi}{d(k)}$$

Example with $\ell = 1$ supervisor and $n - \ell = 6$ supervised nodes



$\rho =$

iSCALA Epoch Example

epochs:

- one for each estimate $k = \ell + 1, 2(\ell + 1), 4(\ell + 1), \ldots$
- initially, "potential" value: $\Phi_{supervised} = \ell,\ \Phi_{supervisor} = 0$

$p(k, i(G_T))$ phases:

(to let supervisors remove "enough" potential ρ)

$r(k, i(G_T))$ rounds:

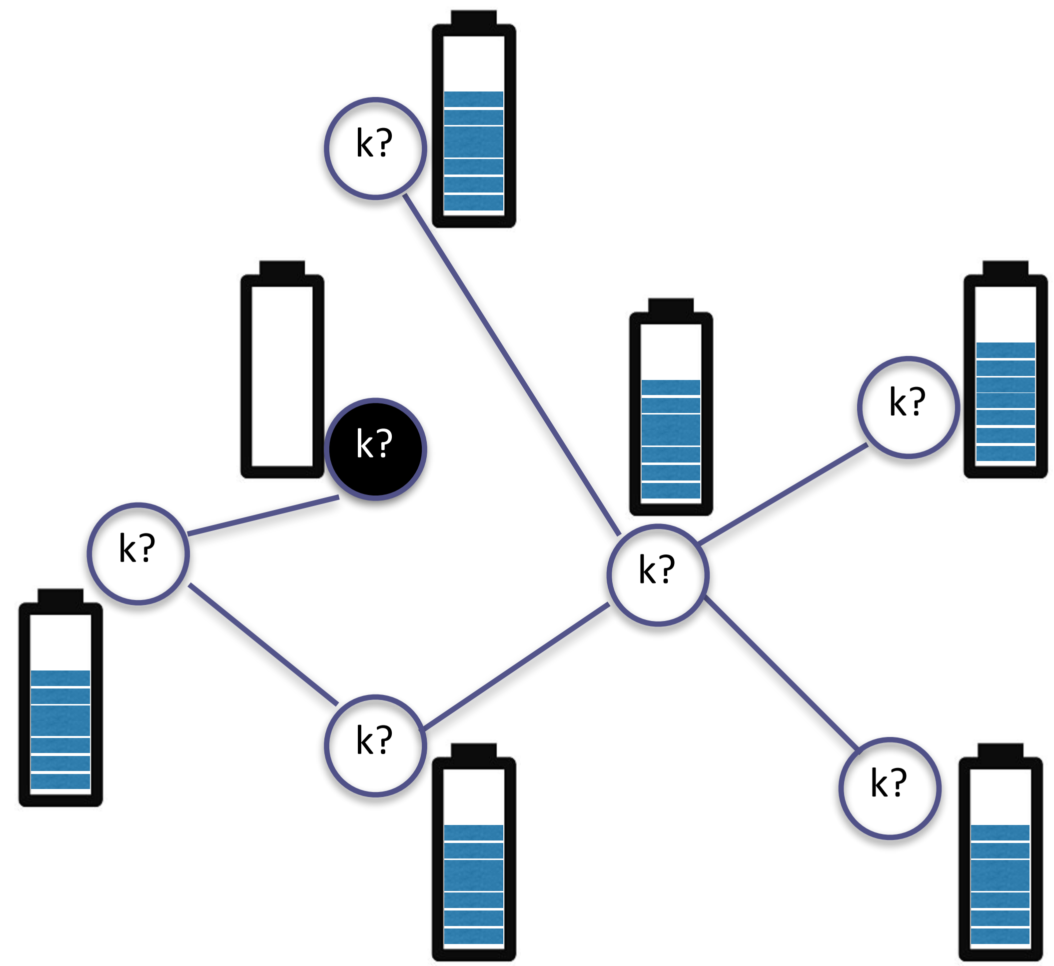
(to "average" the current potentials Φ)

mass distribution:

- broadcast Φ and receive neighbors' Φ_i
- $\Phi = \Phi + \sum_{i \in N} \frac{\Phi_i}{d(k)} - |N| \frac{\Phi}{d(k)}$

- supervisors "remove" their potential: $\rho = \rho + \Phi, \Phi = 0$

Example with $\ell = 1$ supervisor and $n - \ell = 6$ supervised nodes



$\rho =$

iSCALA Epoch Example

epochs:

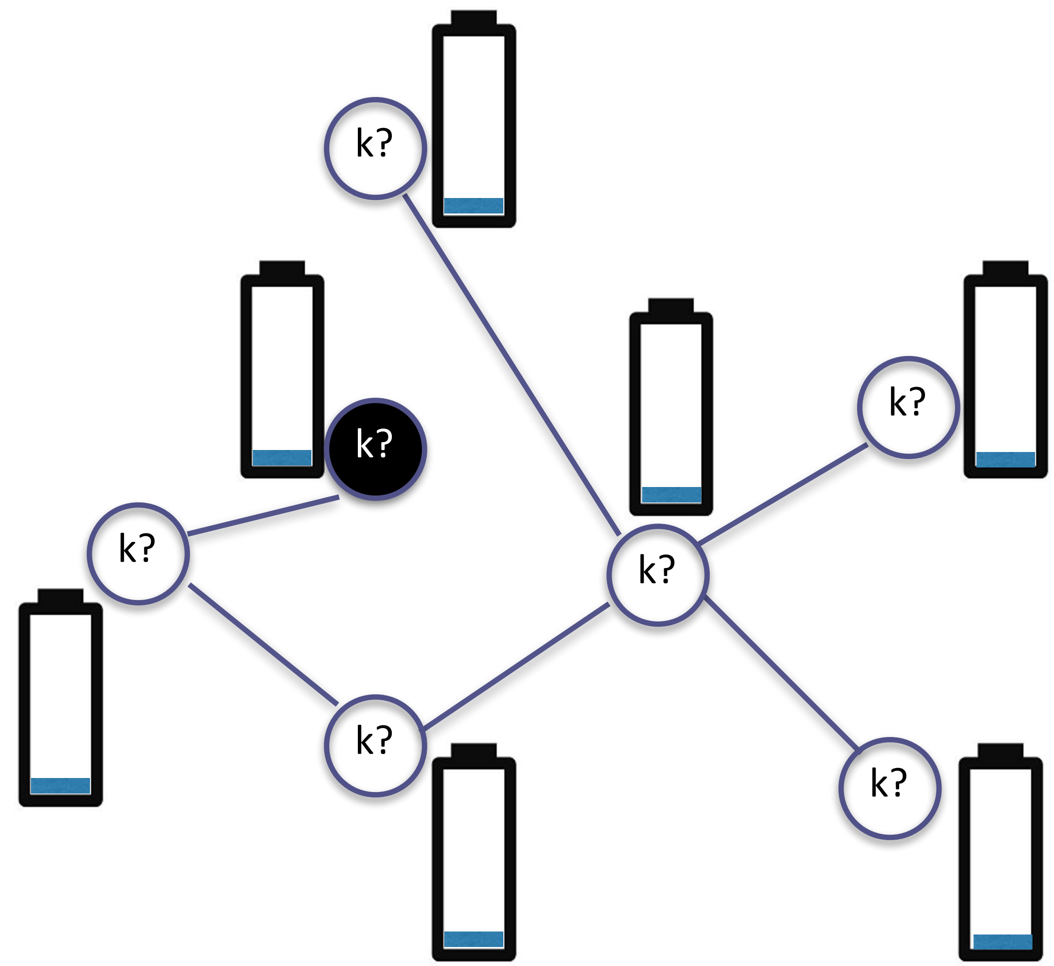
- one for each estimate $k = \ell + 1, 2(\ell + 1), 4(\ell + 1), \ldots$
- initially, "potential" value: $\Phi_{supervised} = \ell,\ \Phi_{supervisor} = 0$

$p(k, i(G_T))$ phases:

(to let supervisors remove "enough" potential ρ)

$r(k, i(G_T))$ rounds:

(to "average" the current potentials Φ)



$\rho =$

iSCALA Epoch Example

epochs:

- one for each estimate $k = \ell + 1, 2(\ell + 1), 4(\ell + 1), \ldots$
- initially, "potential" value:
  $\Phi_{supervised} = \ell,\ \Phi_{supervisor} = 0$

$p(k, i(G_T))$ phases:

(to let supervisors remove "enough" potential ρ)

$r(k, i(G_T))$ rounds:

(to "average" the current potentials Φ)

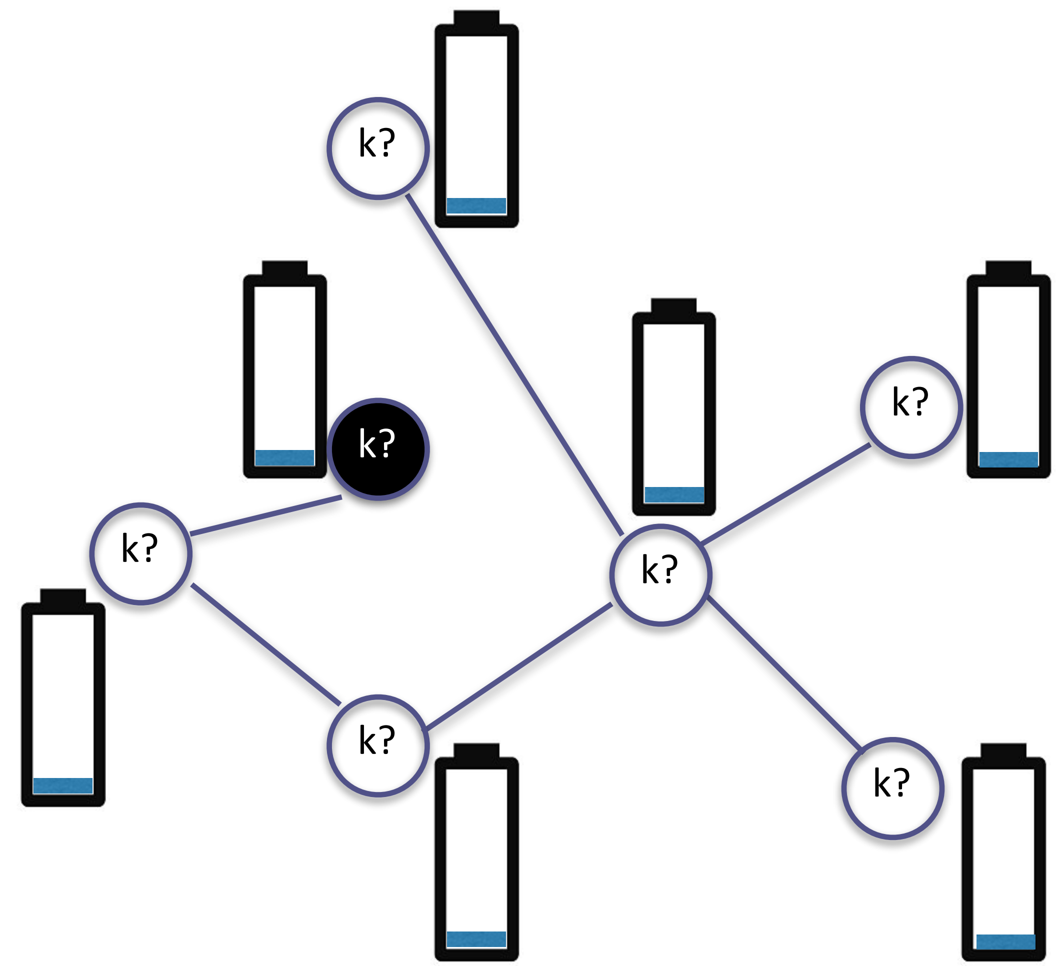
mass distribution:

- broadcast Φ and receive neighbors' Φ_i
- $$\Phi = \Phi + \sum_{i \in N} \frac{\Phi_i}{d(k)} - |N| \frac{\Phi}{d(k)}$$

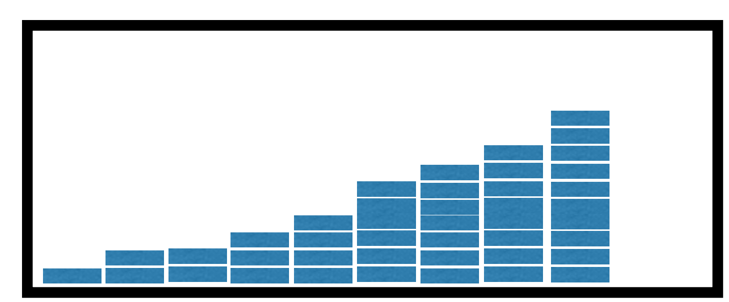
- supervisors "remove" their potential: $\rho = \rho + \Phi, \Phi = 0$

- supervisors decide according to ρ
- supervisors notify if $k \geq n$
- try next k if needed

After $p(k, i(G_T))$ phases…



$\rho =$



iSCALA Upper Bounds

	topology	communication rounds	
	adversarial $i_{\min}$	$O\left(\frac{n^{3+\epsilon}}{\ell i_{\min}^2}\log^3 n\right)$	Cor 6.6
stochastic	$\tilde{i}_{\min}$ whp	$O\left(\frac{n^{3+\epsilon}}{\ell \tilde{i}_{\min}^2}\log^3 n\right)$ whp	Thm 7.1
stochastic	Erdős-Rényi - Gilbert $p_{\min}$	$O\left(\frac{n^{1+\epsilon}}{\ell p_{\min}^2}\log^5 n\right)$ whp	Cor 7.2
stochastic	RGG $r_{\min} > 2\sqrt{\frac{\log n}{n}}$	$O\left(\frac{n^{3+\epsilon}}{\ell r_{\min}^2}\log^3 n\right)$ whp	Cor 7.3
stochastic	Watts-Strogatz $K_{\min}, \beta_{\min}$	$O\left(\frac{n^{3+\epsilon}}{\ell (K_{\min}\beta_{\min})^2}\frac{\log^9 n}{(\log\log n)^2}\right)$ whp	Cor 7.4
stochastic	Barabàsi-Albert, $m_{0,\min} \leq m_{\min}$	$O\left(\frac{n^{2.75+\epsilon}}{\ell m_{\min}}\log^3 n\right)$ whp	Cor 7.5

iSCALA Upper Bounds

Improves over MMC's $O(n^{4+\epsilon}/\ell \log^3 n)$ for isop. number $i(G) \in \Omega(1/\sqrt{n})$, even if only a lower bound $i_{\min}$ is given.

	topology	communication rounds	
adversarial	$i_{\min}$	$O\left(\frac{n^{3+\epsilon}}{\ell i_{\min}^2} \log^3 n\right)$	Cor 6.6
stochastic	$\tilde{i}_{\min}$ whp	$O\left(\frac{n^{3+\epsilon}}{\ell \tilde{i}_{\min}^2} \log^3 n\right)$ whp	Thm 7.1
stochastic	Erdős-Rényi - Gilbert $p_{\min}$	$O\left(\frac{n^{1+\epsilon}}{\ell p_{\min}^2} \log^5 n\right)$ whp	Cor 7.2
stochastic	RGG $r_{\min} > 2\sqrt{\frac{\log n}{n}}$	$O\left(\frac{n^{3+\epsilon}}{\ell r_{\min}^2} \log^3 n\right)$ whp	Cor 7.3
stochastic	Watts-Strogatz $K_{\min}, \beta_{\min}$	$O\left(\frac{n^{3+\epsilon}}{\ell (K_{\min}\beta_{\min})^2} \frac{\log^9 n}{(\log \log n)^2}\right)$ whp	Cor 7.4
stochastic	Barabàsi-Albert, $m_{0,\min} \leq m_{\min}$	$O\left(\frac{n^{2.75+\epsilon}}{\ell m_{\min}} \log^3 n\right)$ whp	Cor 7.5

Upper bounds for a variety of stochastic network models.

Simulations

- Goal: Evaluation of a hypothetical early-stopping centralized version of iSCALA against the upper bounds in the analysis.
- Inputs:

random graphs	other
Watts-Strogatz (WS)	Trees
Barabàsi-Albert (BA)	Stars
Erdős-Rényi,Gilbert (ER)	Paths

Supervisor nodes located at random. All topologies T-stable connected.

- Parameters:

$n = 6,9,12,\ldots,48$ ER: $p = 0.5$

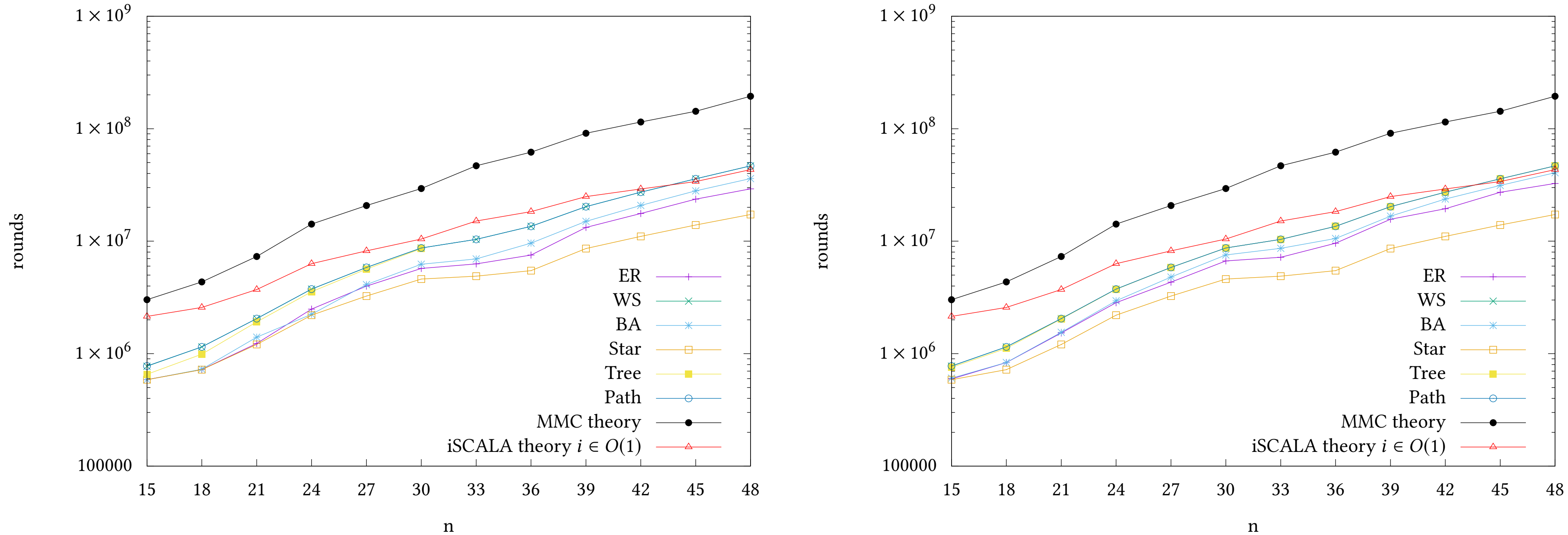
$\ell = 1,\ldots,n/2$ in various steps WS: $\beta = 0.1,0.2,0.4$ and $K = 2,4$

$T = 1,100$ BA: $m = m_0 = 2,4$

$i_{\min}$ lower bound

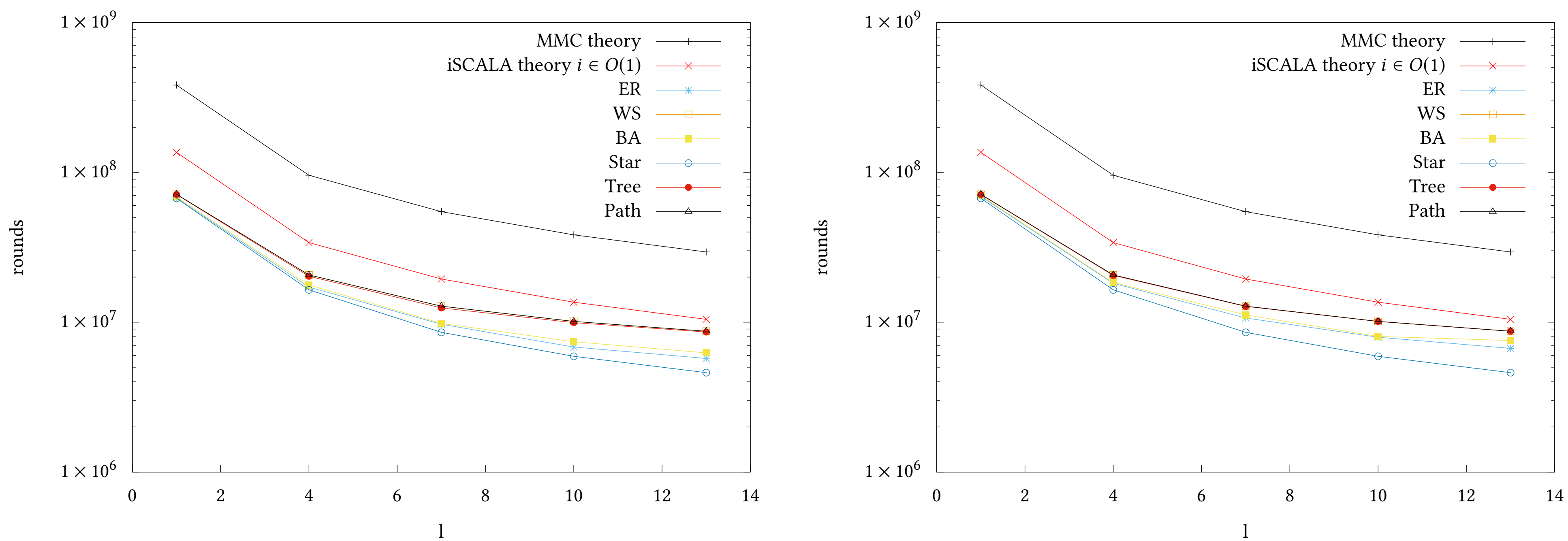
- Average behavior over multiple executions of the simulator.

Simulations Results Examples



(a) 1**-stable networks.** **(b)** 100**-stable networks.**

Figure 2: Simulation results for some $\ell \in [n/3, n/2]$.



(a) 1**-stable networks.** **(b)** 100**-stable networks.**

Figure 3: Simulation results for $n = 30$.

Questions?