Neighborhood Gossip: Concurrent Averaging through Local Interference

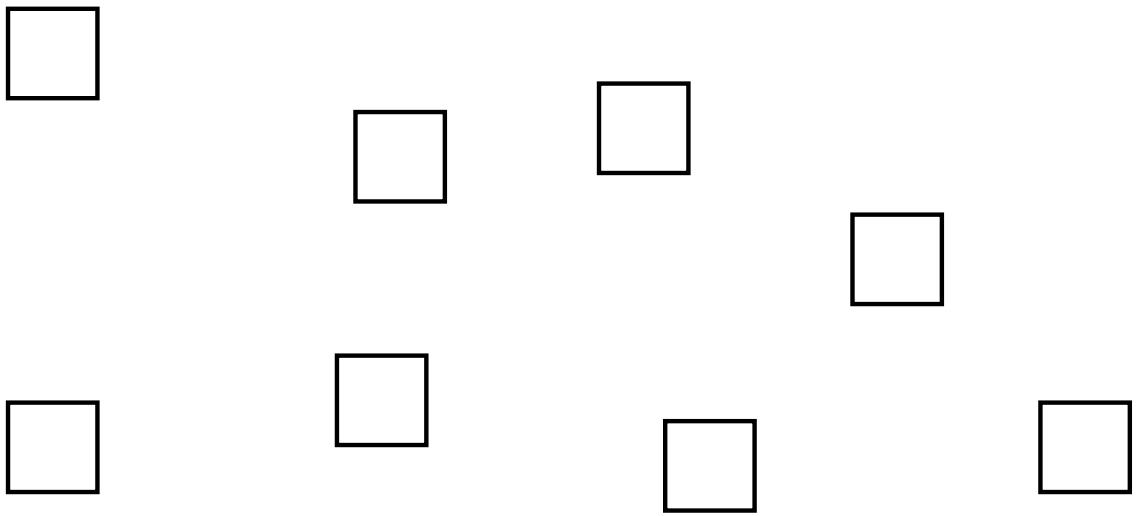
Bobak Nazer[†], Alexandros G. Dimakis[◦], and Michael Gastpar[†]

† Wireless Foundations Center
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley

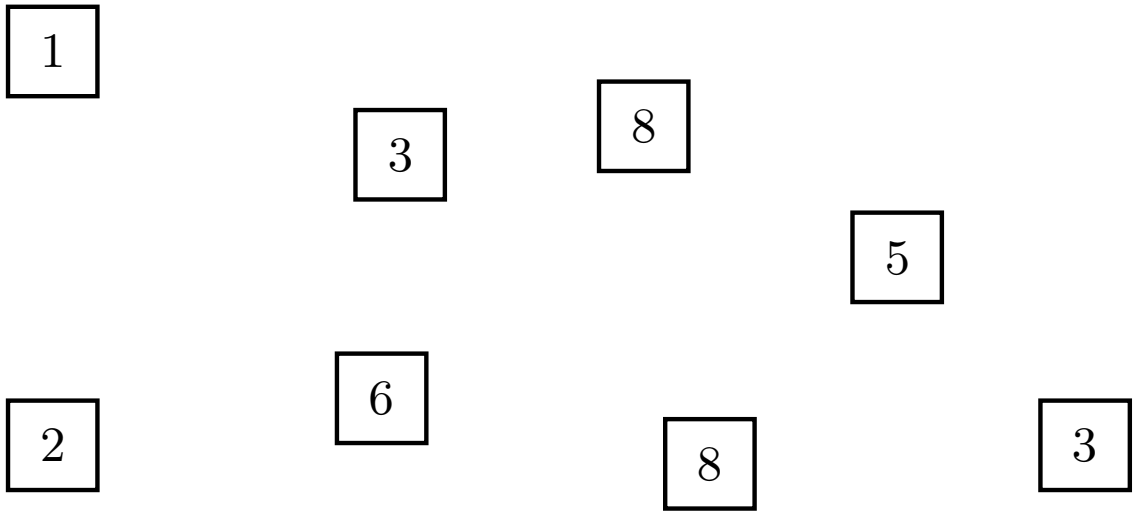
◦ Center for the Mathematics of Information
California Institute of Technology

ICASSP 2009

Wireless Sensor Network

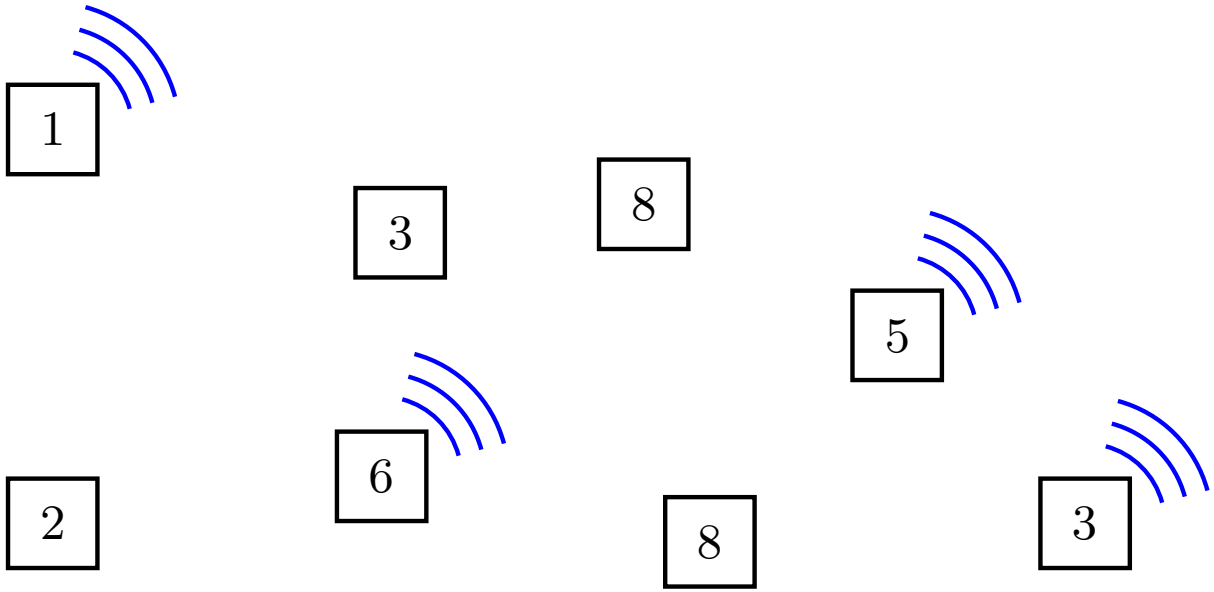


Wireless Sensor Network



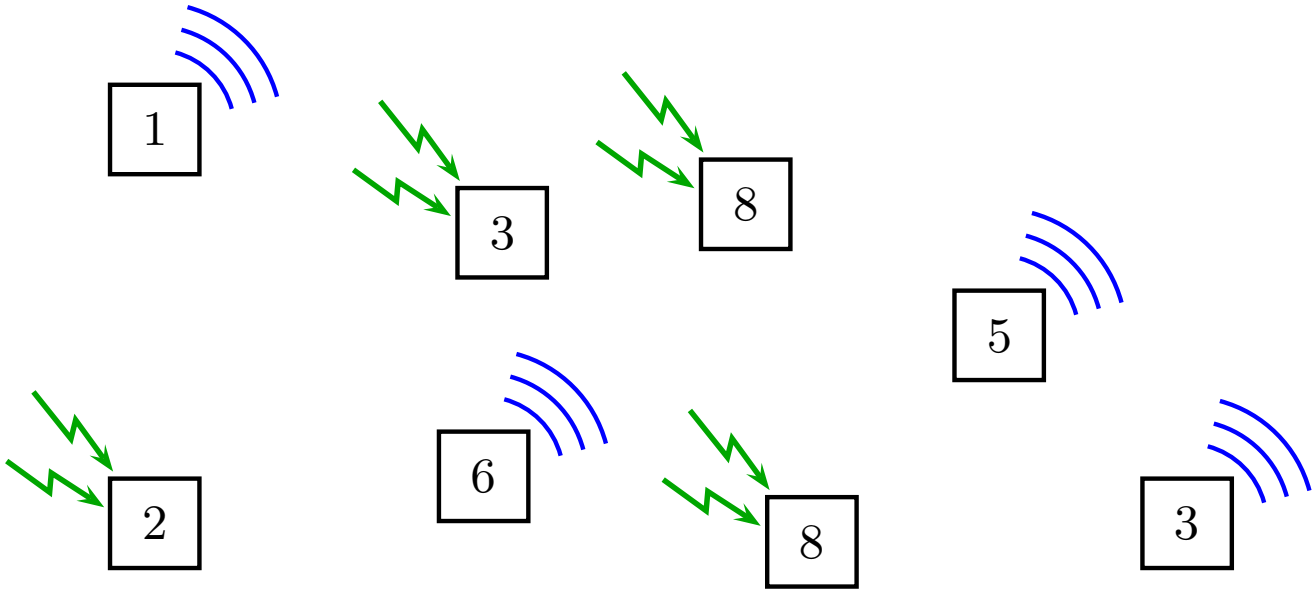
- Each node has a real-valued measurement.

Wireless Sensor Network



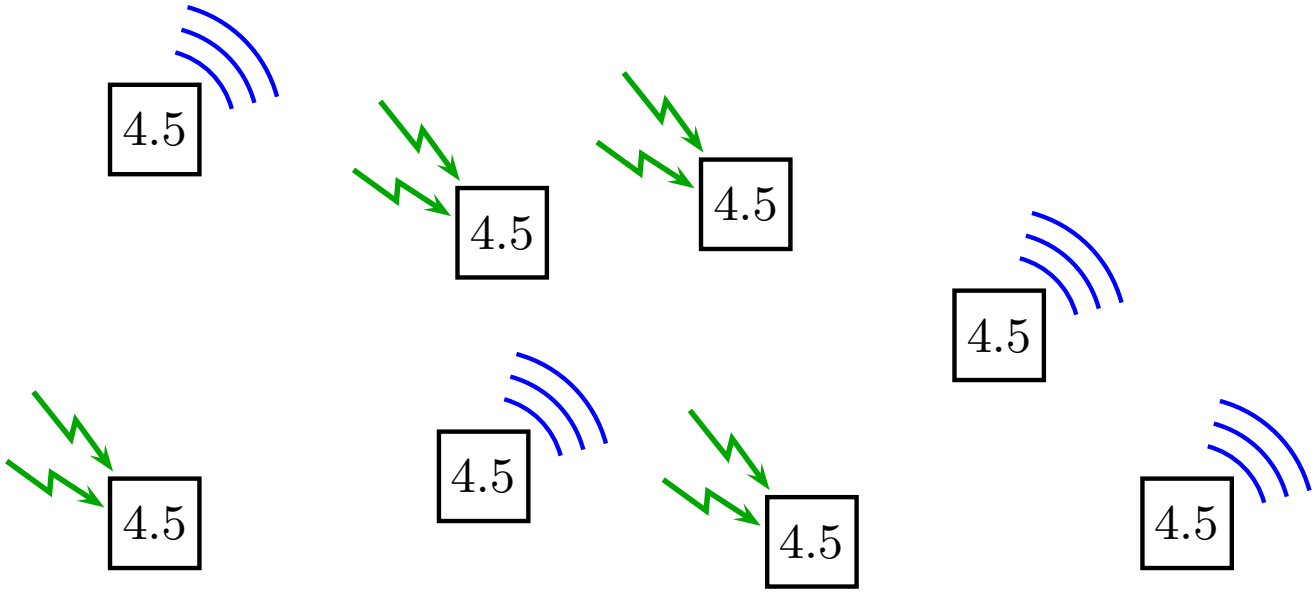
- Each node has a real-valued measurement.
- Nodes communicate over the wireless channel.

Wireless Sensor Network



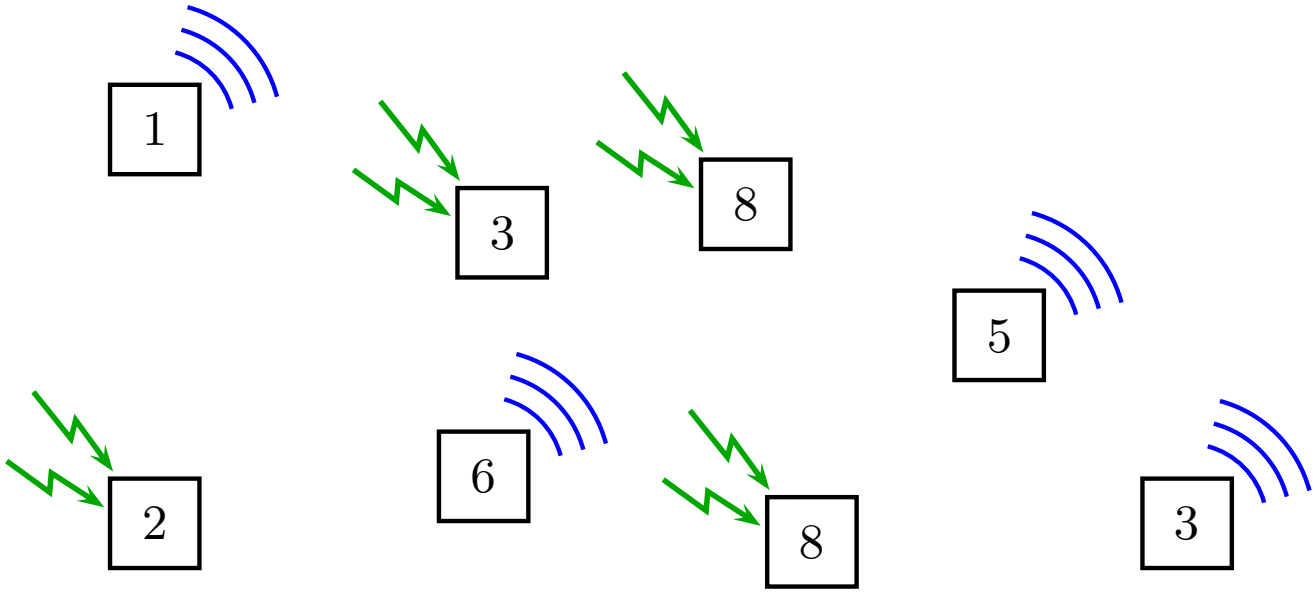
- Each node has a real-valued measurement.
- Nodes communicate over the wireless channel.
- Want distributed communication algorithm for each node to learn the global average.

Wireless Sensor Network



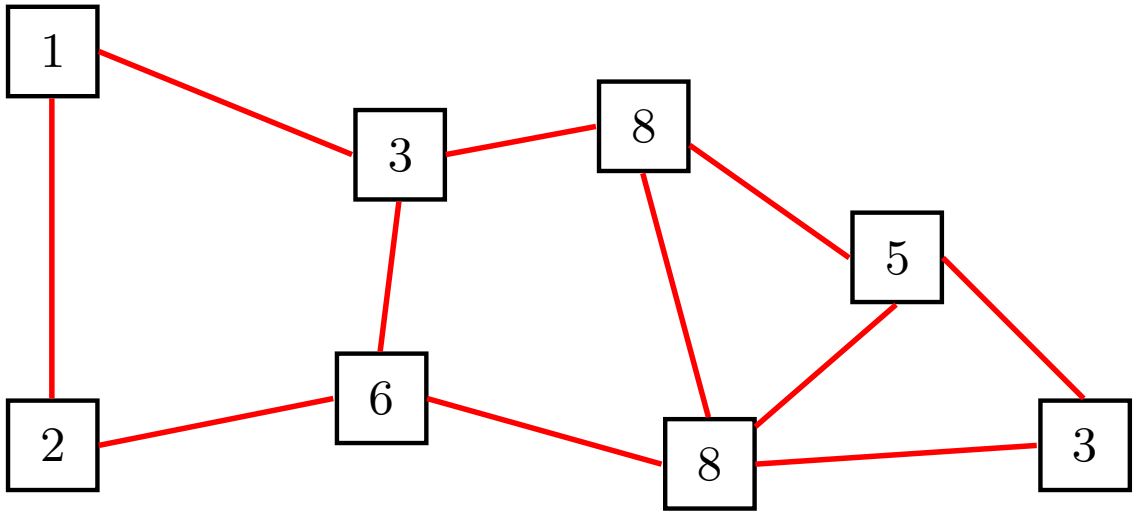
- Each node has a real-valued measurement.
- Nodes communicate over the wireless channel.
- Want distributed communication algorithm for each node to learn the global average.

Communication Protocol: Bit Pipes



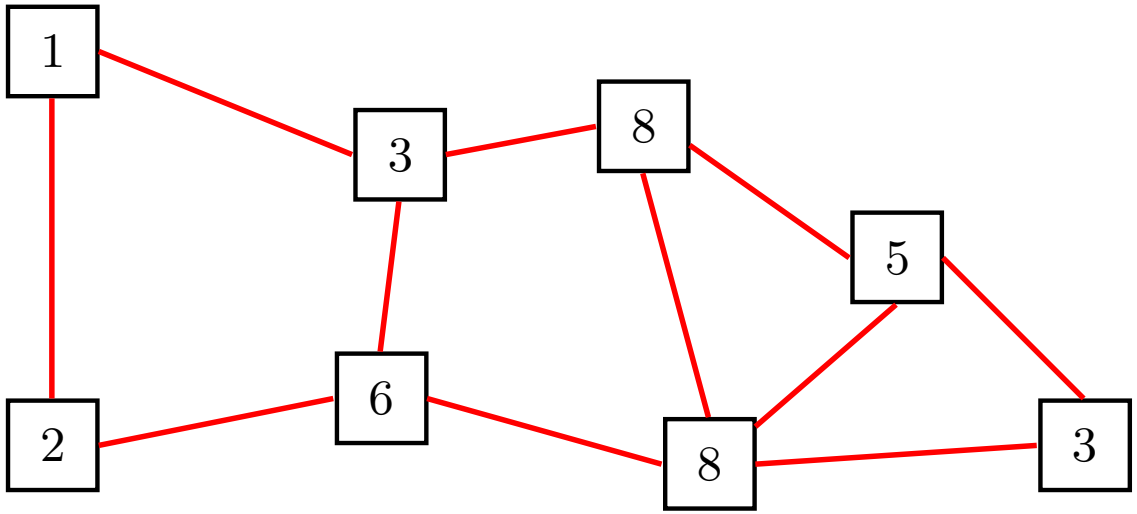
- Need a communication protocol to deal with the wireless channel (interference, noise, fading).

Communication Protocol: Bit Pipes



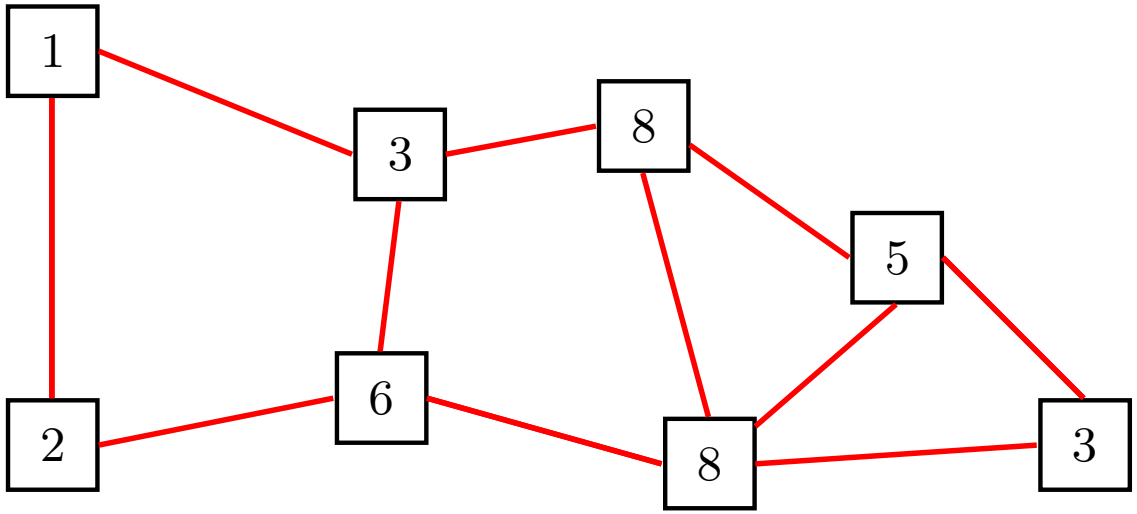
- Need a communication protocol to deal with the wireless channel (interference, noise, fading).
- Usual solution: nodes send messages to their neighbors.

Communication Protocol: Bit Pipes



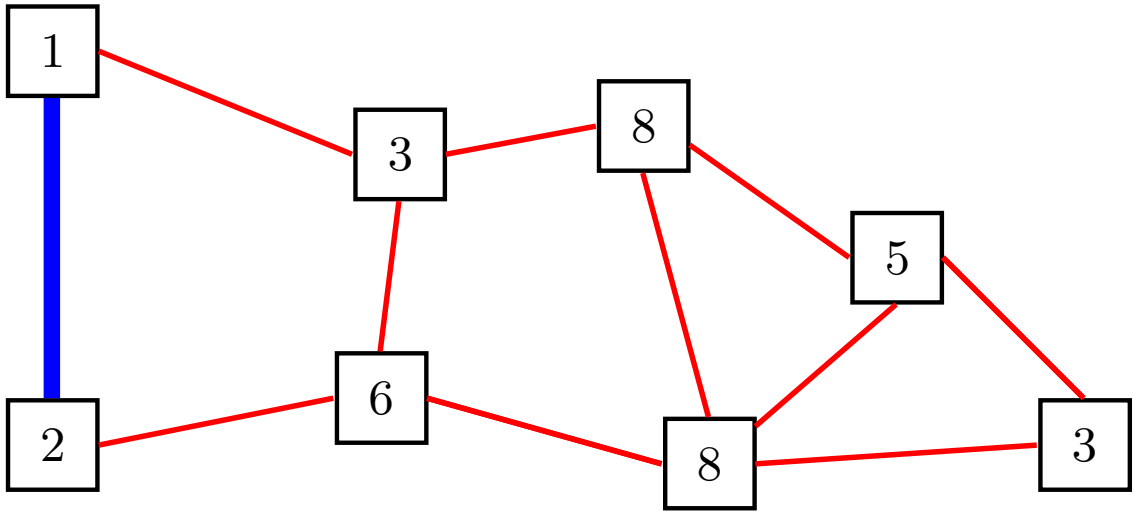
- Need a communication protocol to deal with the wireless channel (interference, noise, fading).
- Usual solution: nodes send messages to their neighbors.
- Wireless network becomes a graph of reliable bit pipes.

Gossip Algorithm



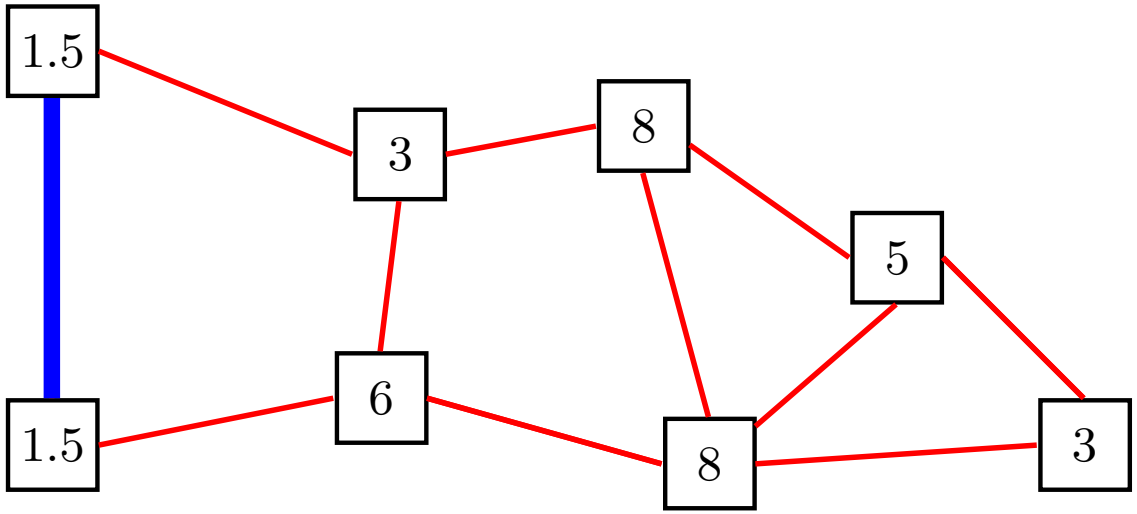
- A node randomly wakes up and computes a local average with a neighbor.

Gossip Algorithm



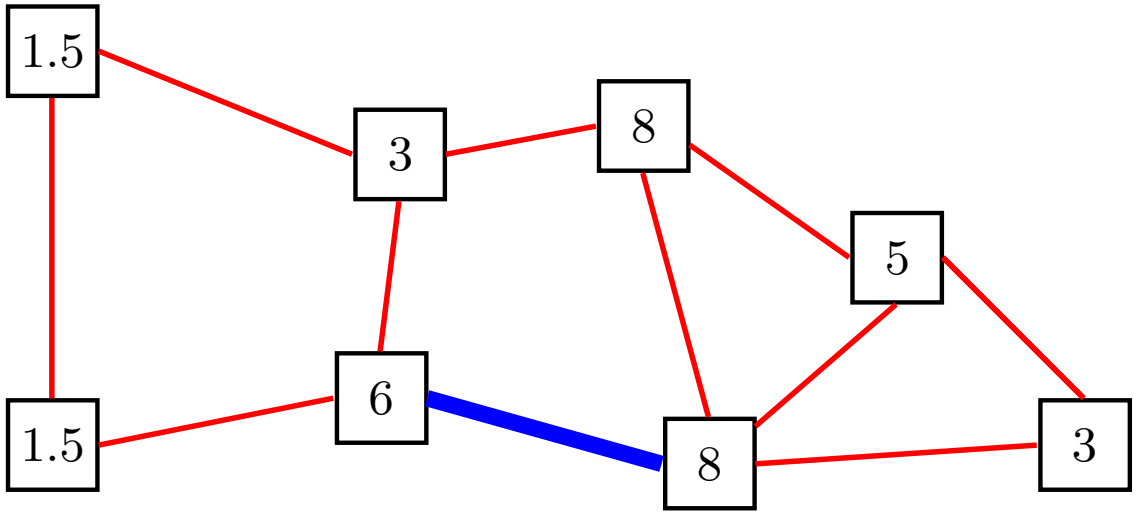
- A node randomly wakes up and computes a local average with a neighbor.
- After enough rounds, every node knows the global average.

Gossip Algorithm



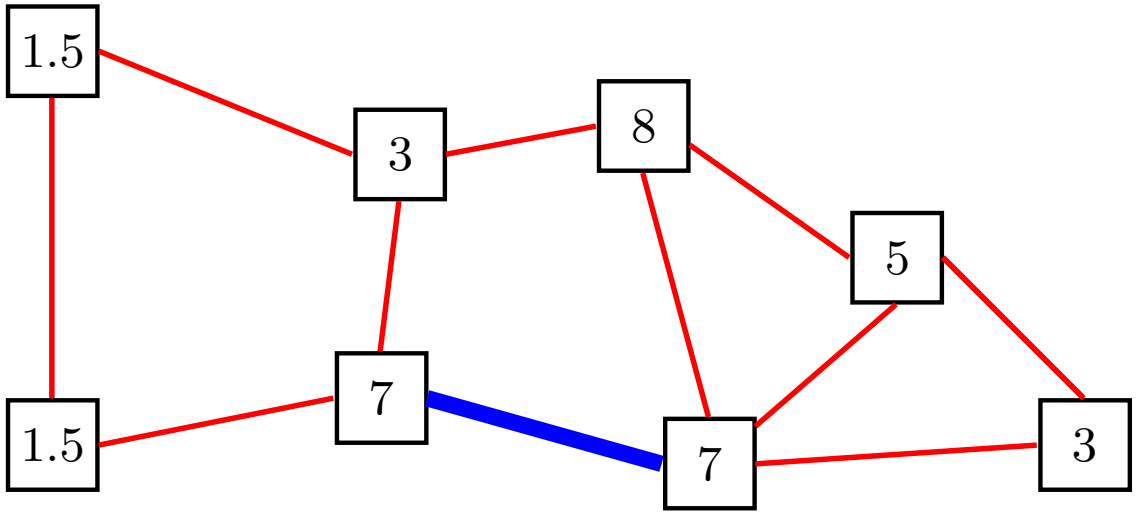
- A node randomly wakes up and computes a local average with a neighbor.
- After enough rounds, every node knows the global average.
- Pairwise gossip well-studied for general graphs by **Boyd-Ghosh-Prabhakar-Shah '06**.

Gossip Algorithm



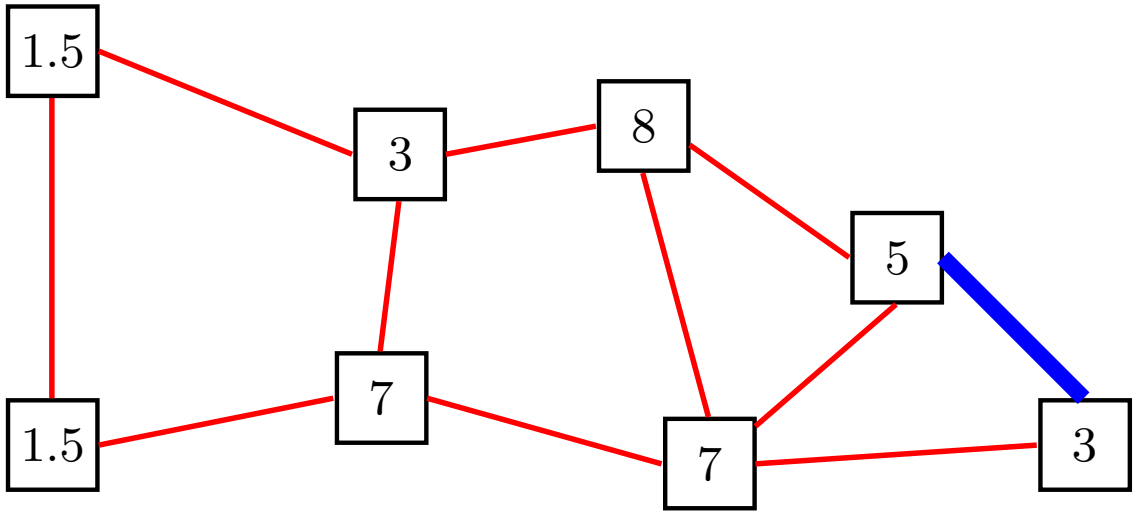
- A node randomly wakes up and computes a local average with a neighbor.
- After enough rounds, every node knows the global average.
- Pairwise gossip well-studied for general graphs by **Boyd-Ghosh-Prabhakar-Shah '06**.

Gossip Algorithm



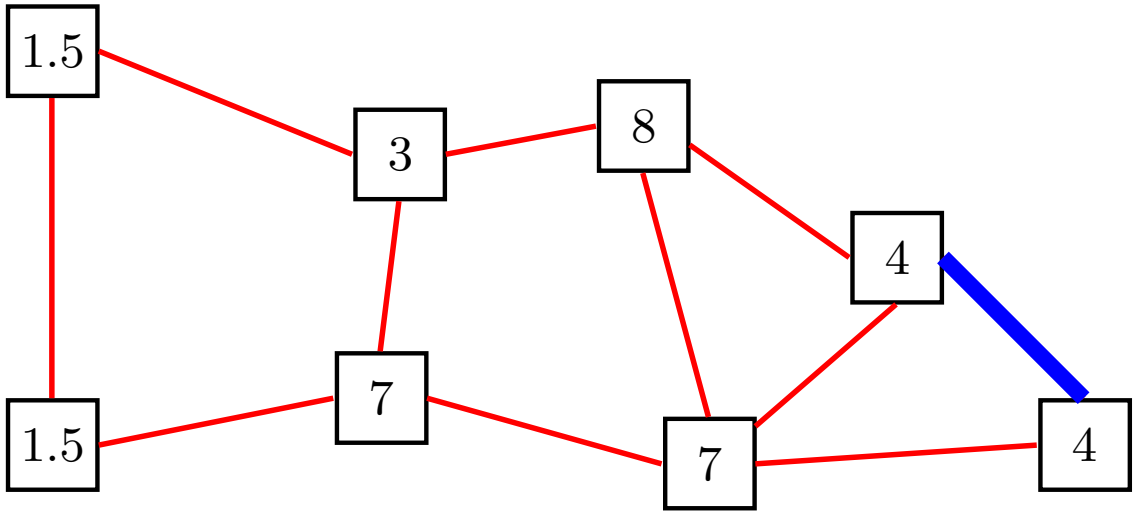
- A node randomly wakes up and computes a local average with a neighbor.
- After enough rounds, every node knows the global average.
- Pairwise gossip well-studied for general graphs by **Boyd-Ghosh-Prabhakar-Shah '06**.

Gossip Algorithm



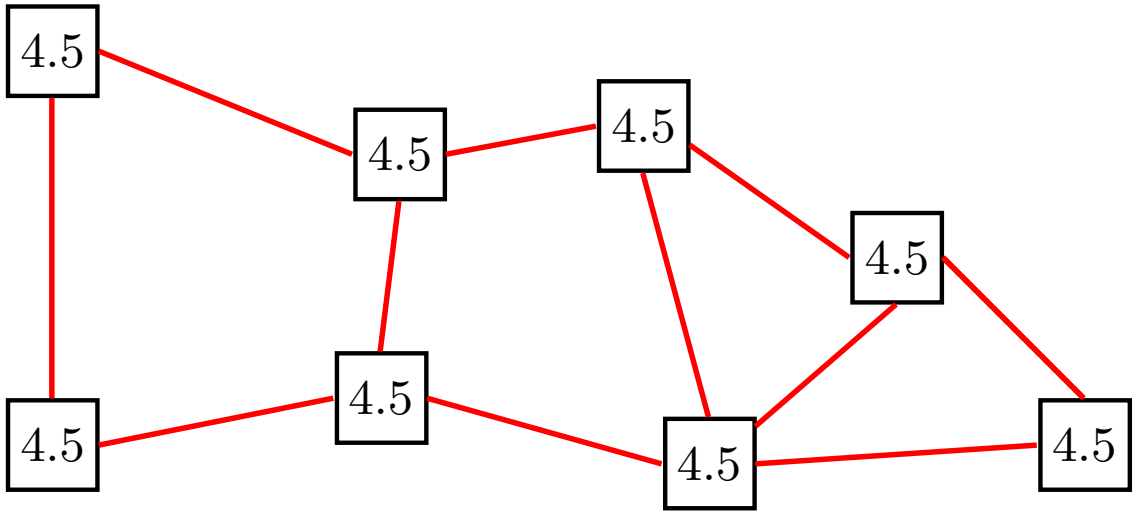
- A node randomly wakes up and computes a local average with a neighbor.
- After enough rounds, every node knows the global average.
- Pairwise gossip well-studied for general graphs by **Boyd-Ghosh-Prabhakar-Shah '06**.

Gossip Algorithm



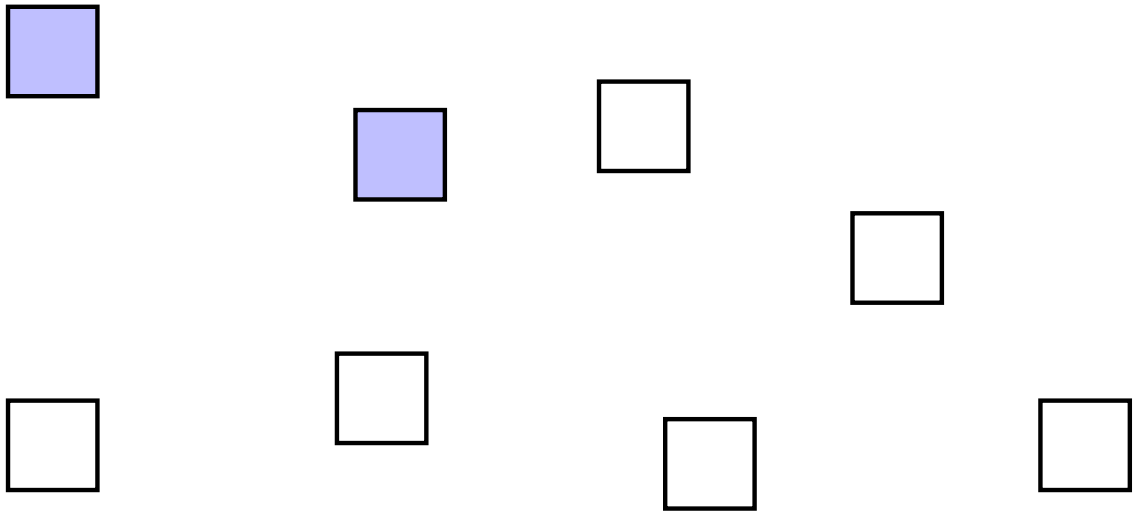
- A node randomly wakes up and computes a local average with a neighbor.
- After enough rounds, every node knows the global average.
- Pairwise gossip well-studied for general graphs by **Boyd-Ghosh-Prabhakar-Shah '06**.

Gossip Algorithm



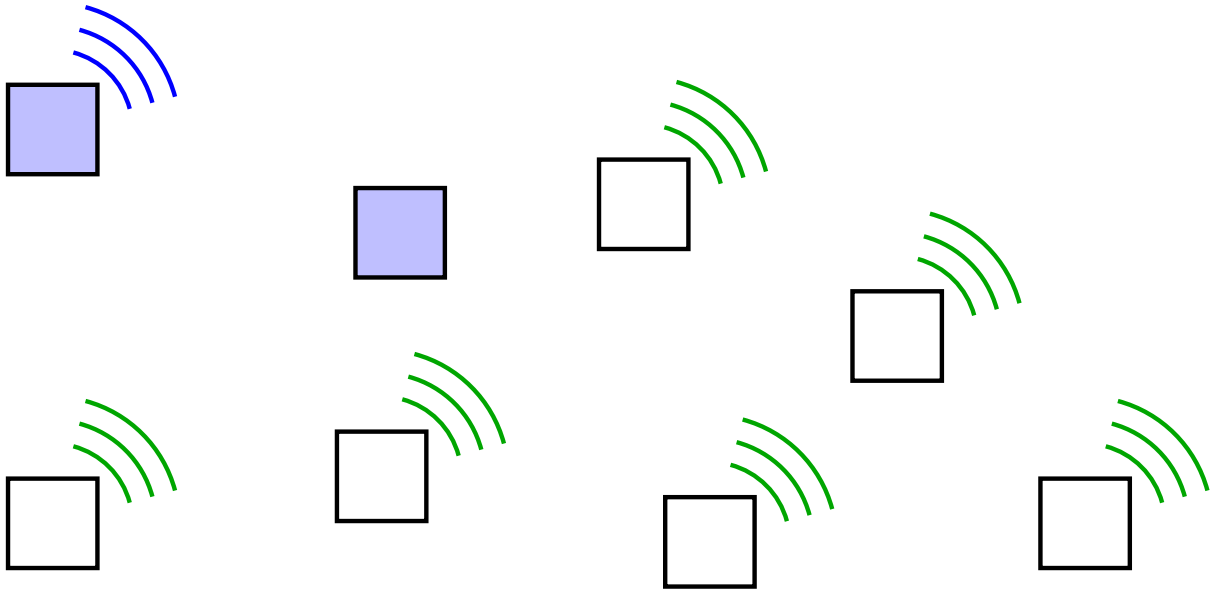
- A node randomly wakes up and computes a local average with a neighbor.
- After enough rounds, every node knows the global average.
- Pairwise gossip well-studied for general graphs by **Boyd-Ghosh-Prabhakar-Shah '06**.

Treating Interference as Noise



- Two neighbors communicating in the presence of interference.

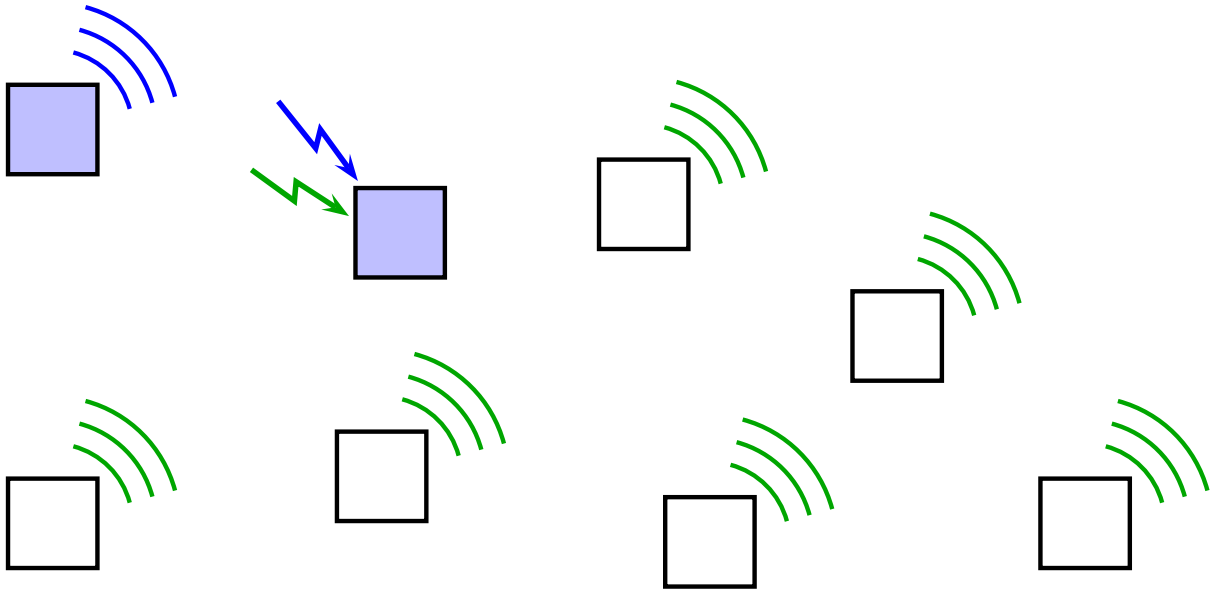
Treating Interference as Noise



- Two neighbors communicating in the presence of interference.
- Received signal:

$$\mathbf{y}_i = h_i \mathbf{x}_i + \sum_{j \neq i} h_j \mathbf{x}_j + \mathbf{z}_i$$

Treating Interference as Noise



- Two neighbors communicating in the presence of interference.
- Received signal:

$$\mathbf{y}_i = h_i \mathbf{x}_i + \sum_{j \neq i} h_j \mathbf{x}_j + \mathbf{z}_i$$

A New Approach

Usual approach: Treat interference as noise and extract desired message.

$$\mathbf{y}_i = h_i \mathbf{x}_i + \sum_{j \neq i} h_j \mathbf{x}_j + \mathbf{z}_i \xrightarrow{\text{Channel Code}} \mathbf{x}_i$$

A New Approach

Usual approach: Treat interference as noise and extract desired message.

$$\mathbf{y}_i = h_i \mathbf{x}_i + \sum_{j \neq i} h_j \mathbf{x}_j + \mathbf{z}_i \xrightarrow{\text{Channel Code}} \mathbf{x}_i$$

Problem: bit rates decrease as number of users increases.

A New Approach

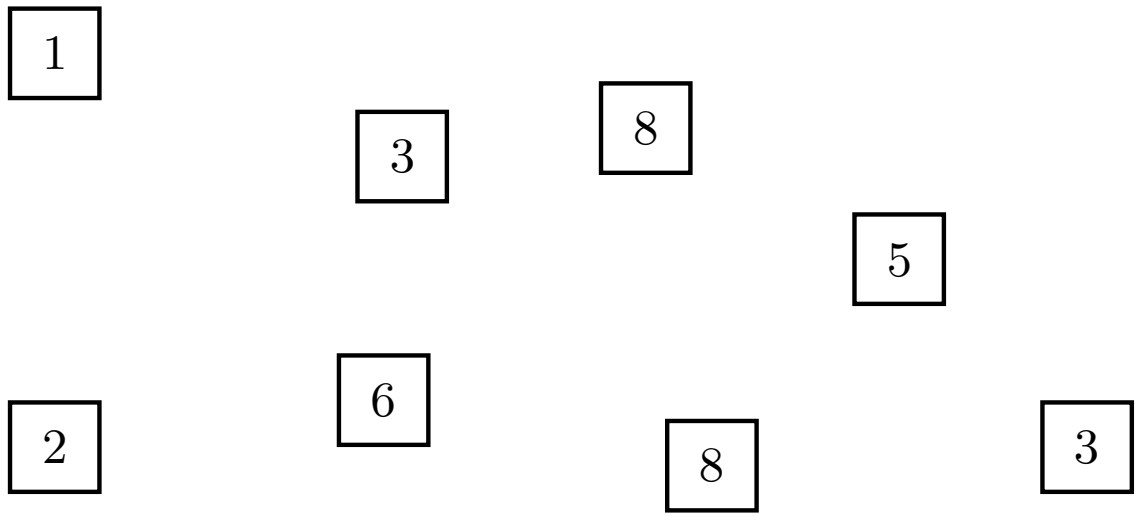
Usual approach: Treat interference as noise and extract desired message.

$$\mathbf{y}_i = h_i \mathbf{x}_i + \sum_{j \neq i} h_j \mathbf{x}_j + \mathbf{z}_i \xrightarrow{\text{Channel Code}} \mathbf{x}_i$$

Problem: bit rates decrease as number of users increases.

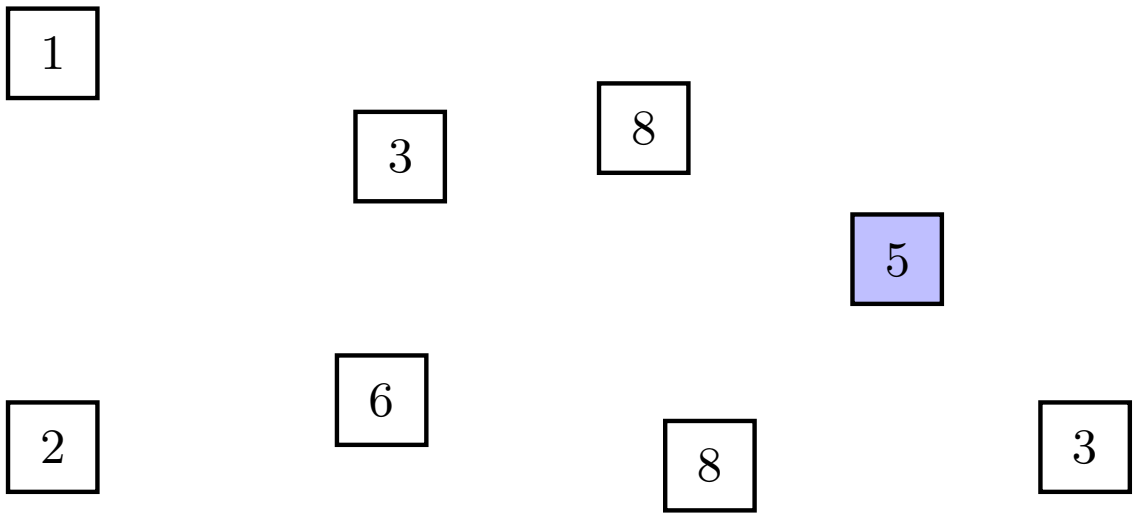
New approach: What if the interference could be used for averaging?

Neighborhood Gossip



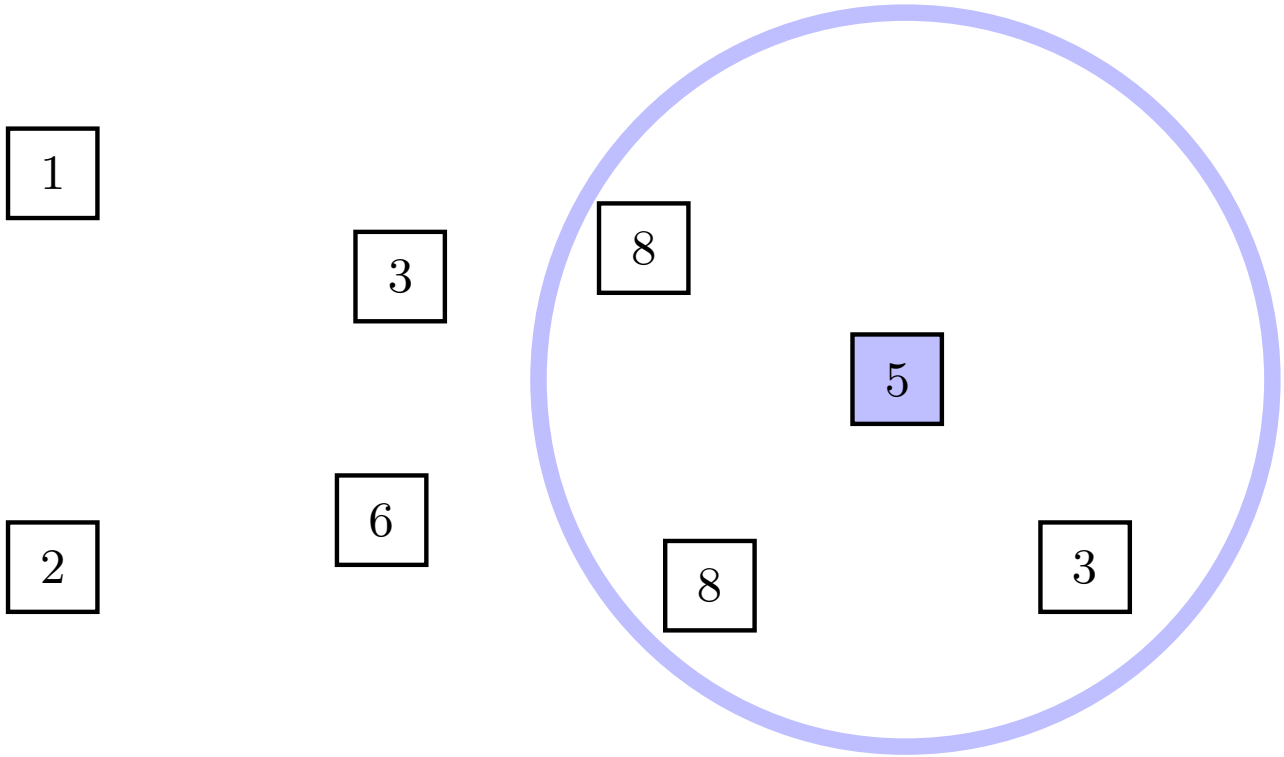
- What if instead of averaging with one neighbor at a time we could average over a neighborhood?
- How can we exploit the additive interference to do this?
- Does this save any time or energy?

Neighborhood Gossip



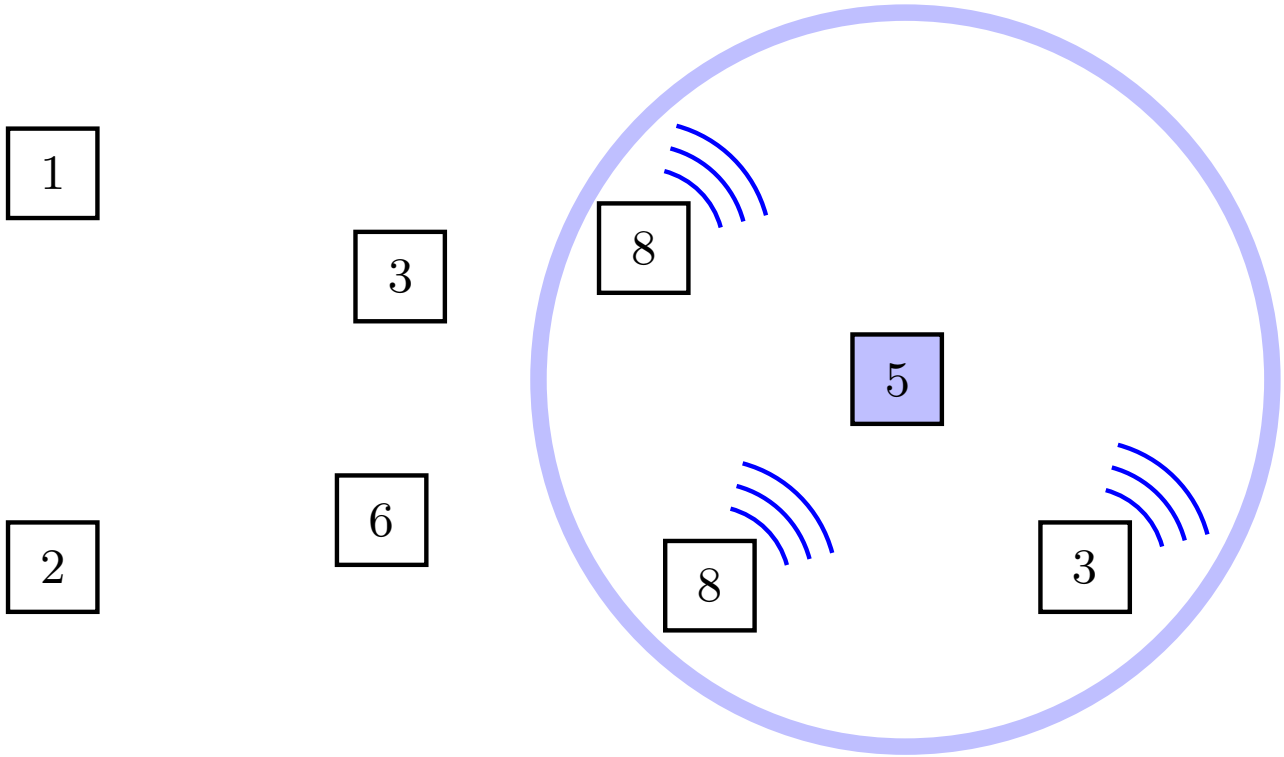
- What if instead of averaging with one neighbor at a time we could average over a neighborhood?
- How can we exploit the additive interference to do this?
- Does this save any time or energy?

Neighborhood Gossip



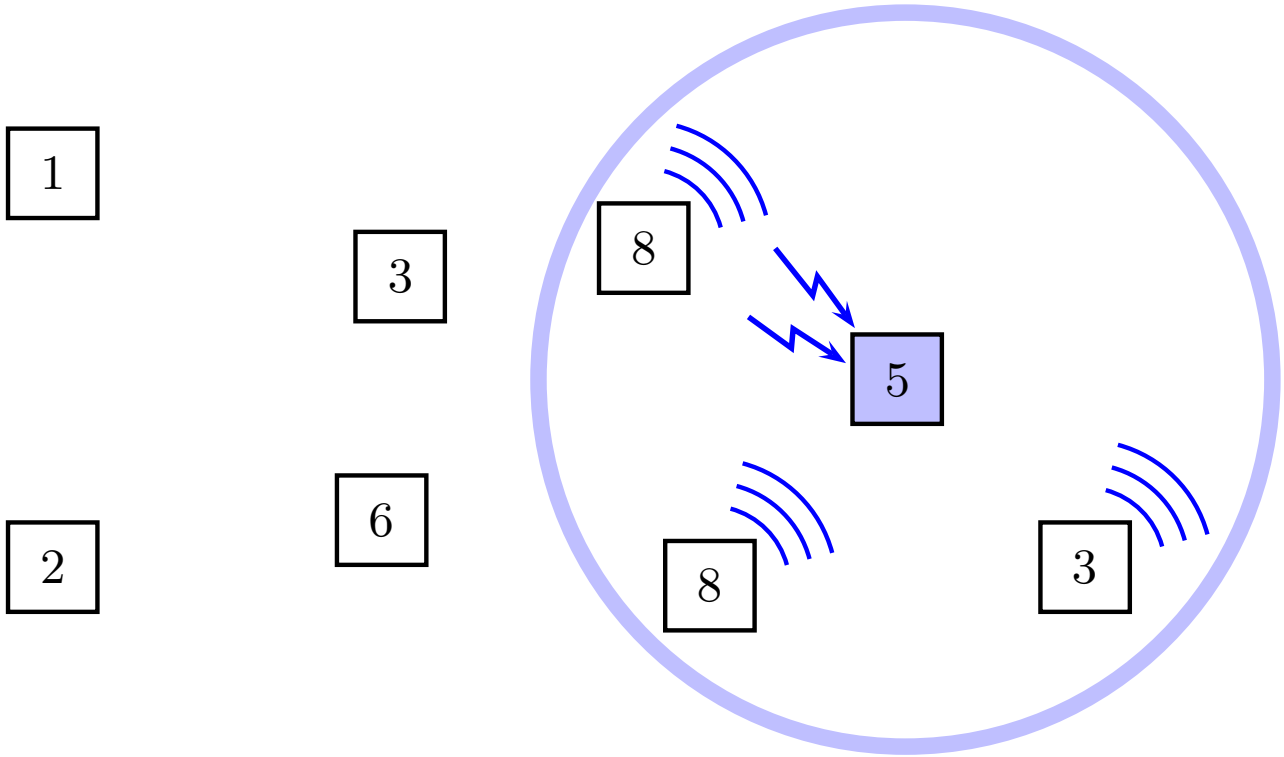
- What if instead of averaging with one neighbor at a time we could average over a neighborhood?
- How can we exploit the additive interference to do this?
- Does this save any time or energy?

Neighborhood Gossip



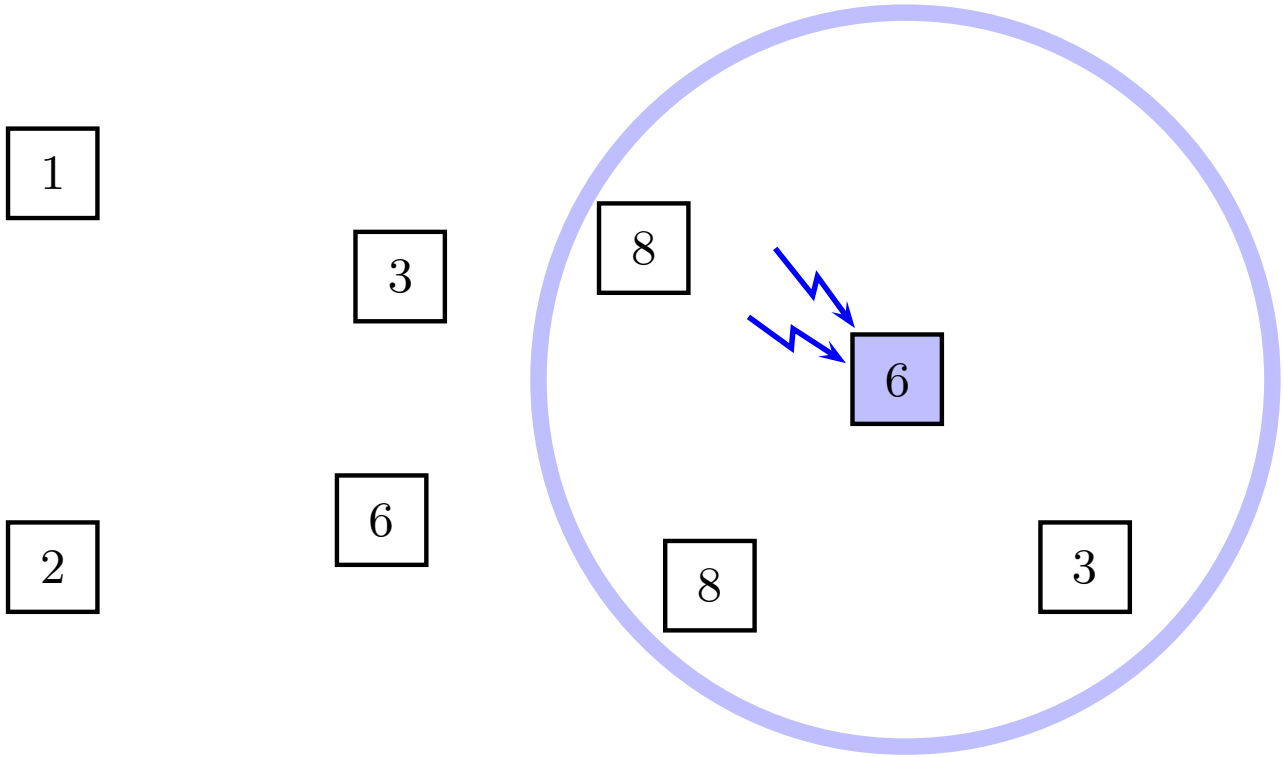
- What if instead of averaging with one neighbor at a time we could average over a neighborhood?
- How can we exploit the additive interference to do this?
- Does this save any time or energy?

Neighborhood Gossip



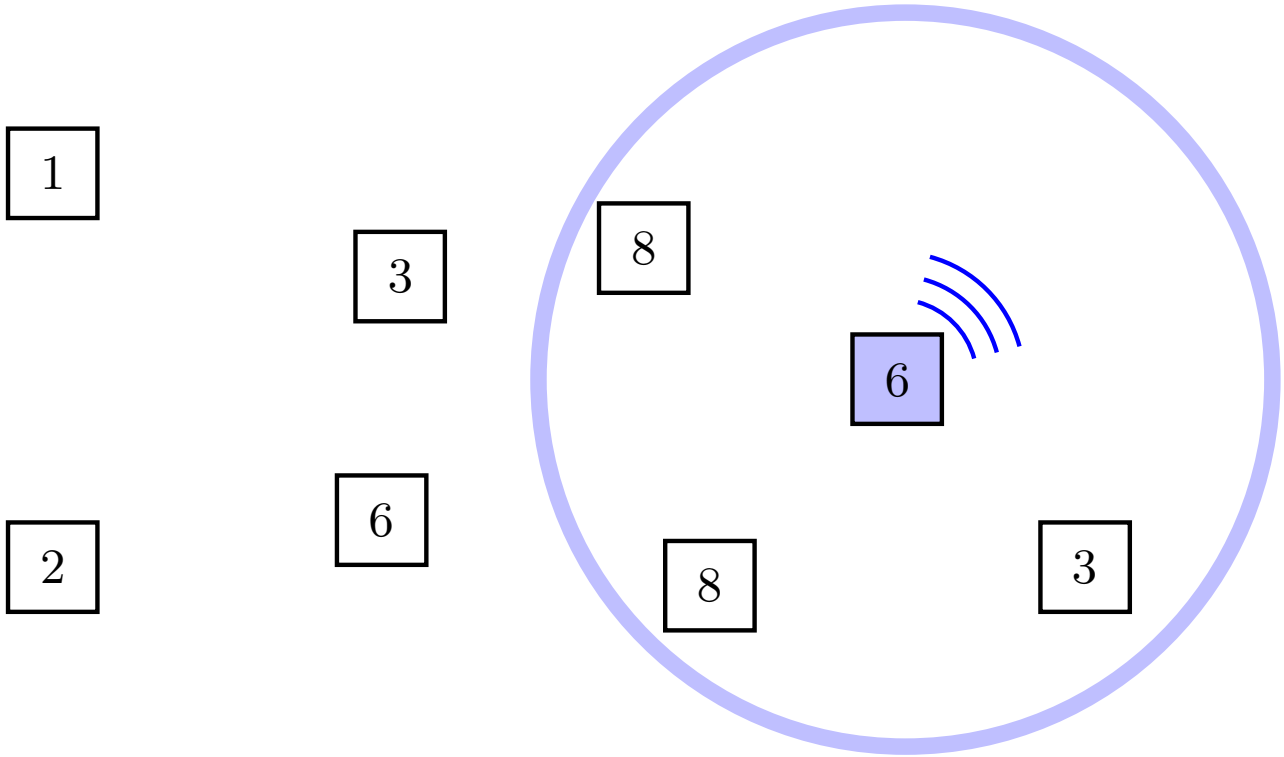
- What if instead of averaging with one neighbor at a time we could average over a neighborhood?
- How can we exploit the additive interference to do this?
- Does this save any time or energy?

Neighborhood Gossip



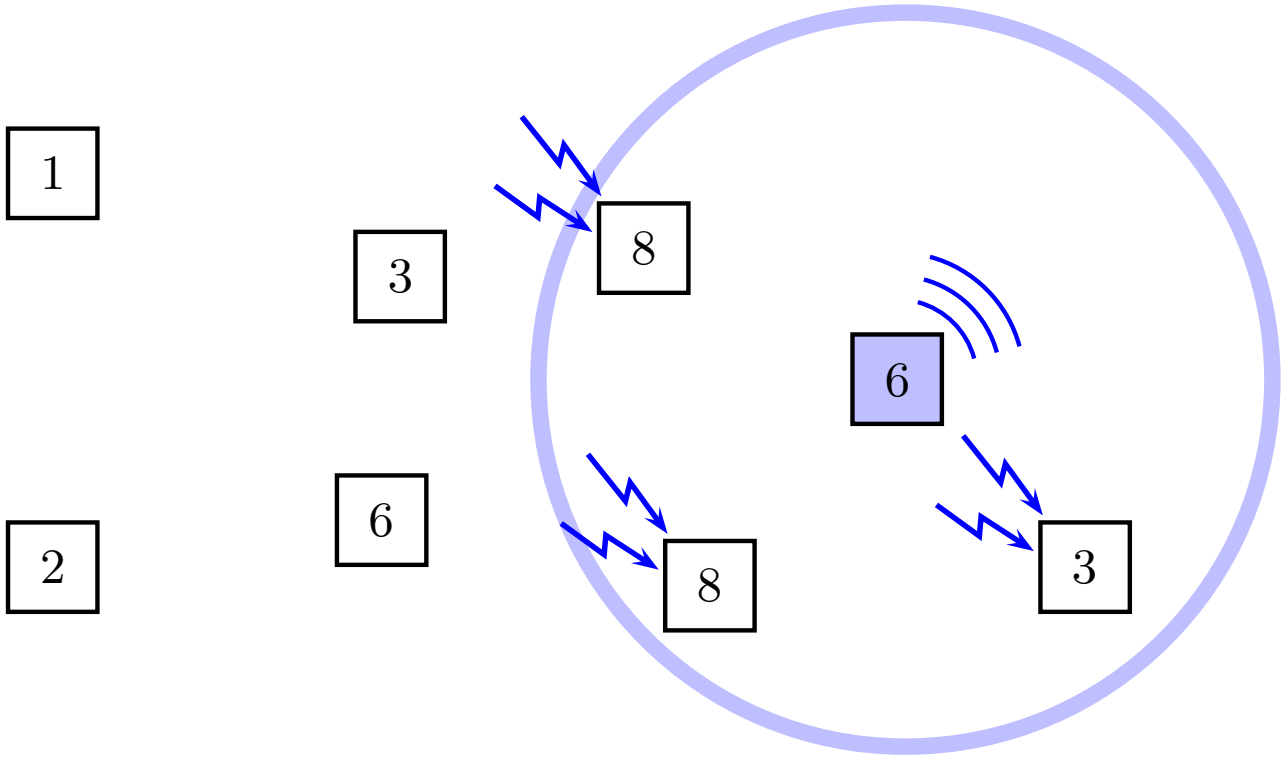
- What if instead of averaging with one neighbor at a time we could average over a neighborhood?
- How can we exploit the additive interference to do this?
- Does this save any time or energy?

Neighborhood Gossip



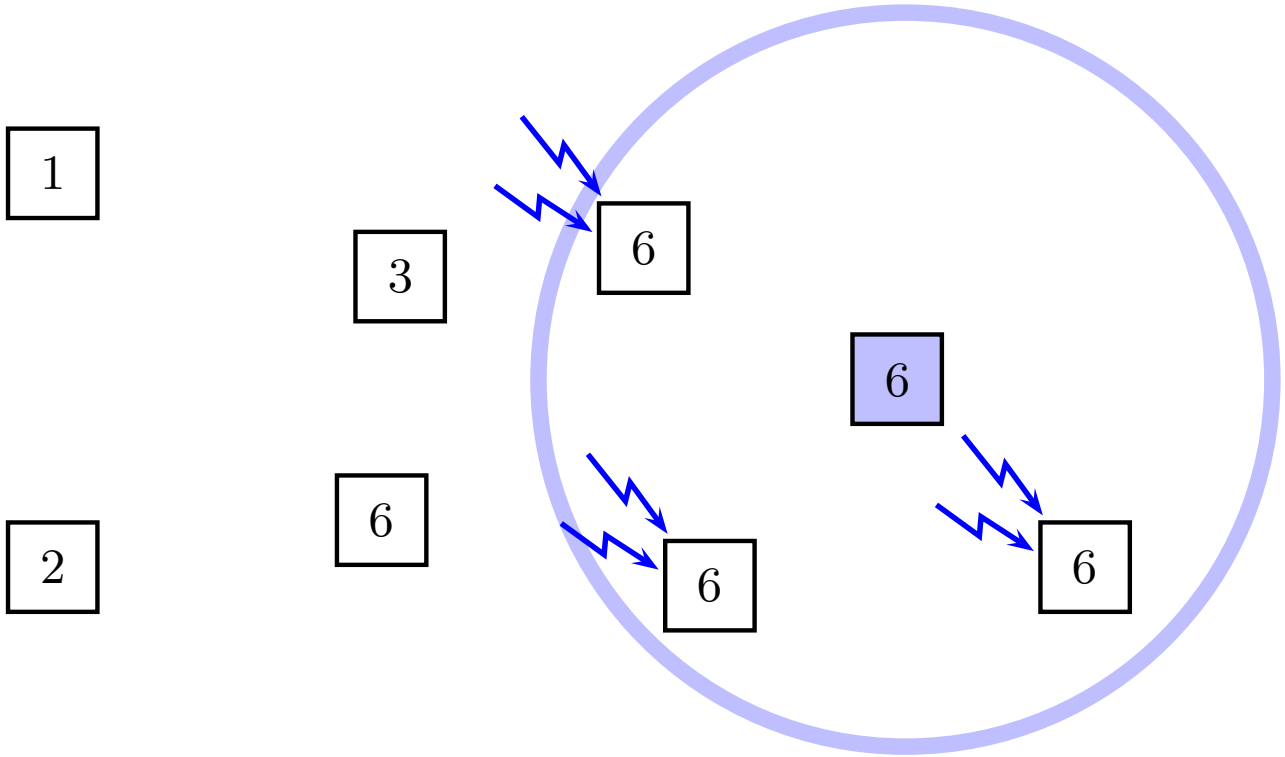
- What if instead of averaging with one neighbor at a time we could average over a neighborhood?
- How can we exploit the additive interference to do this?
- Does this save any time or energy?

Neighborhood Gossip



- What if instead of averaging with one neighbor at a time we could average over a neighborhood?
- How can we exploit the additive interference to do this?
- Does this save any time or energy?

Neighborhood Gossip

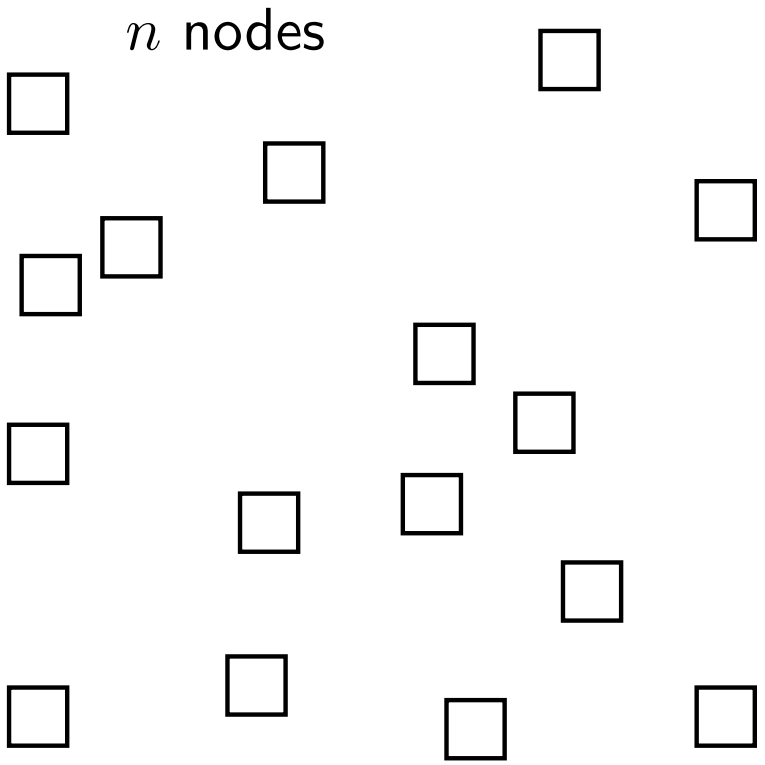


- What if instead of averaging with one neighbor at a time we could average over a neighborhood?
- How can we exploit the additive interference to do this?
- Does this save any time or energy?

Problem Statement

- n nodes randomly placed in a square
- Each node has a real-valued measurement
- AWGN multiple-access channel model:

$$Y_\ell = \sum r_{m\ell}^{-\alpha/2} \phi_{m\ell} X_m + Z_\ell$$

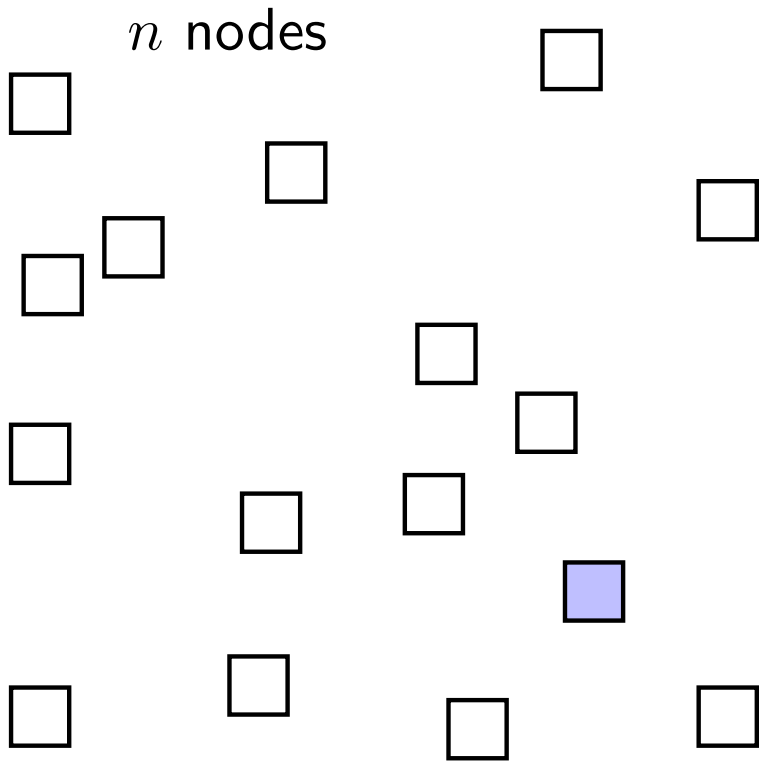


  - Want all nodes to learn the global average.

Problem Statement

- n nodes randomly placed in a square
- Each node has a real-valued measurement
- AWGN multiple-access channel model:

$$Y_\ell = \sum r_{m\ell}^{-\alpha/2} \phi_{m\ell} X_m + Z_\ell$$

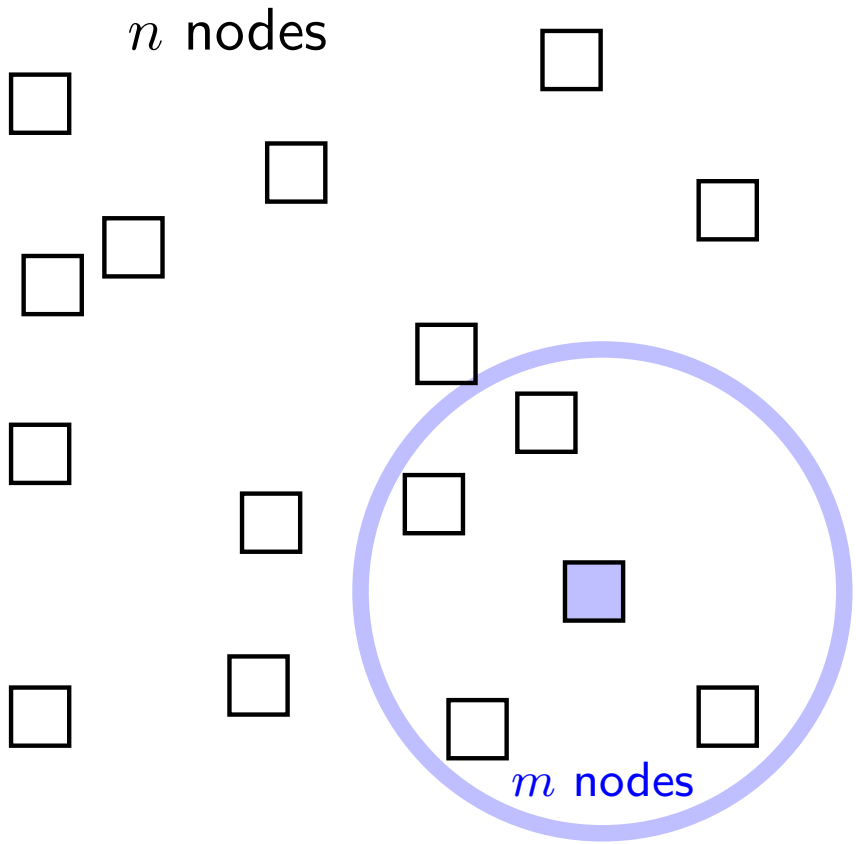


  - Want all nodes to learn the global average.

Problem Statement

- n nodes randomly placed in a square
- Each node has a real-valued measurement
- AWGN multiple-access channel model:

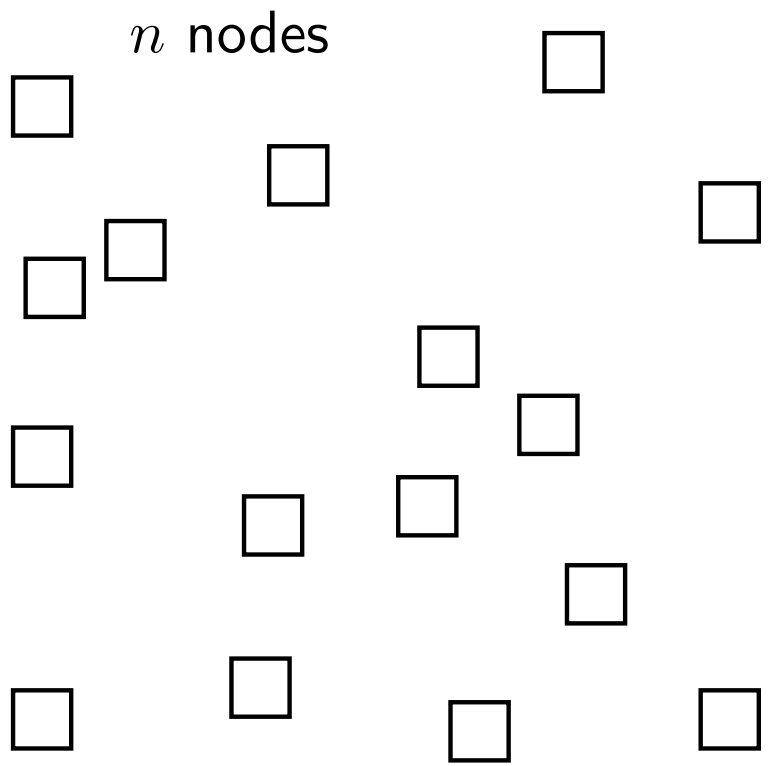
$$Y_\ell = \sum r_{m\ell}^{-\alpha/2} \phi_{m\ell} X_m + Z_\ell$$



- Want all nodes to learn the global average.
- Channel knowledge, $r_{m\ell}, \phi_{m\ell}$, available about size m local neighborhood.

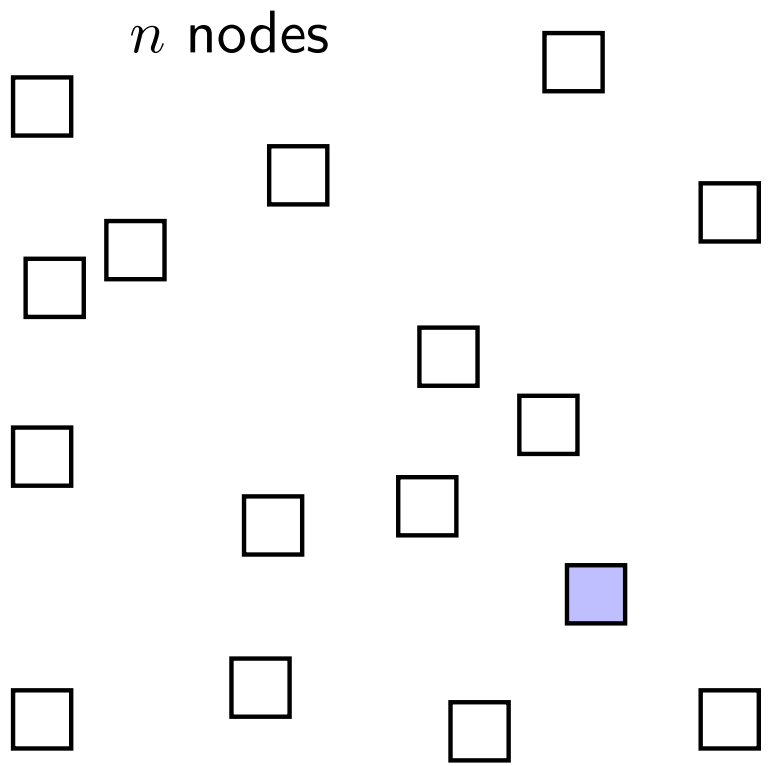
Pairwise Gossip: Convergence

- Gossip Round: a node randomly wakes up and averages with a random neighbor.
- Random geometric graph
- **Boyd-Ghosh-Prabhakar-Shah '06**: Converges in $\Theta(n^2)$ rounds.
- Comes from spectral gap of $\bar{W}$, the averaging matrix.



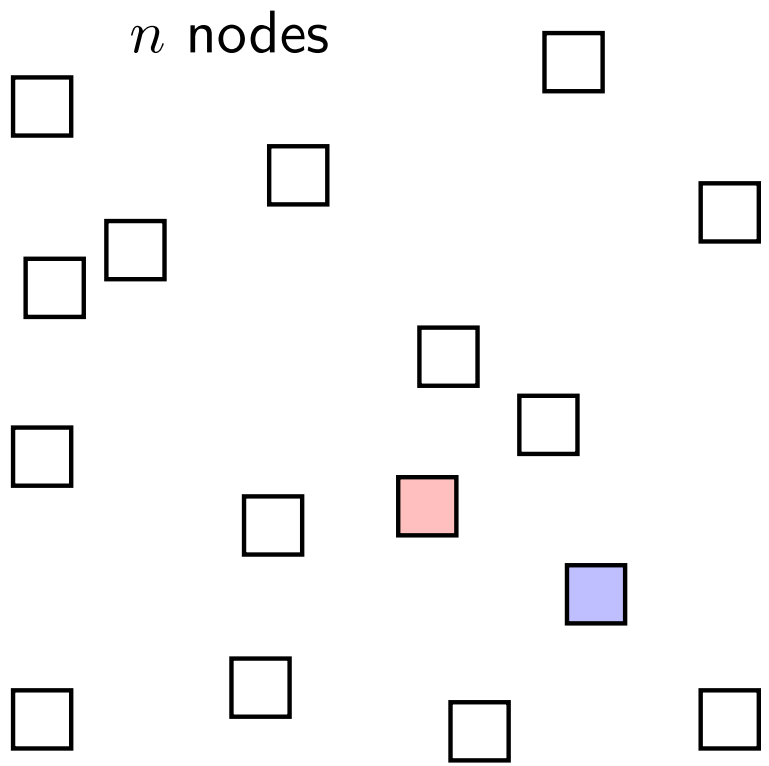
Pairwise Gossip: Convergence

- Gossip Round: a node randomly wakes up and averages with a random neighbor.
- Random geometric graph
- **Boyd-Ghosh-Prabhakar-Shah '06**: Converges in $\Theta(n^2)$ rounds.
- Comes from spectral gap of $\bar{W}$, the averaging matrix.



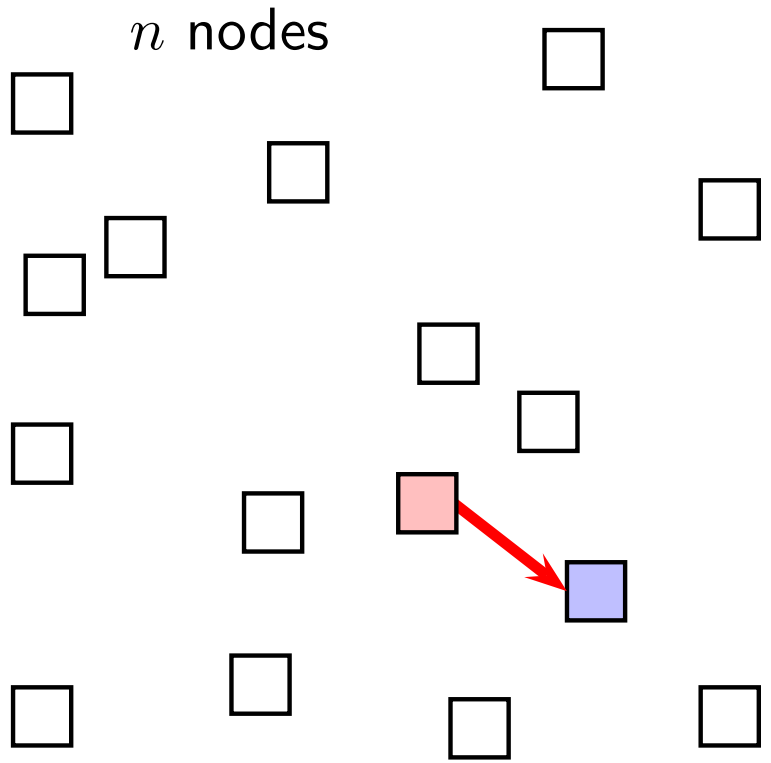
Pairwise Gossip: Convergence

- Gossip Round: a node randomly wakes up and averages with a random neighbor.
- Random geometric graph
- **Boyd-Ghosh-Prabhakar-Shah '06**: Converges in $\Theta(n^2)$ rounds.
- Comes from spectral gap of $\bar{W}$, the averaging matrix.



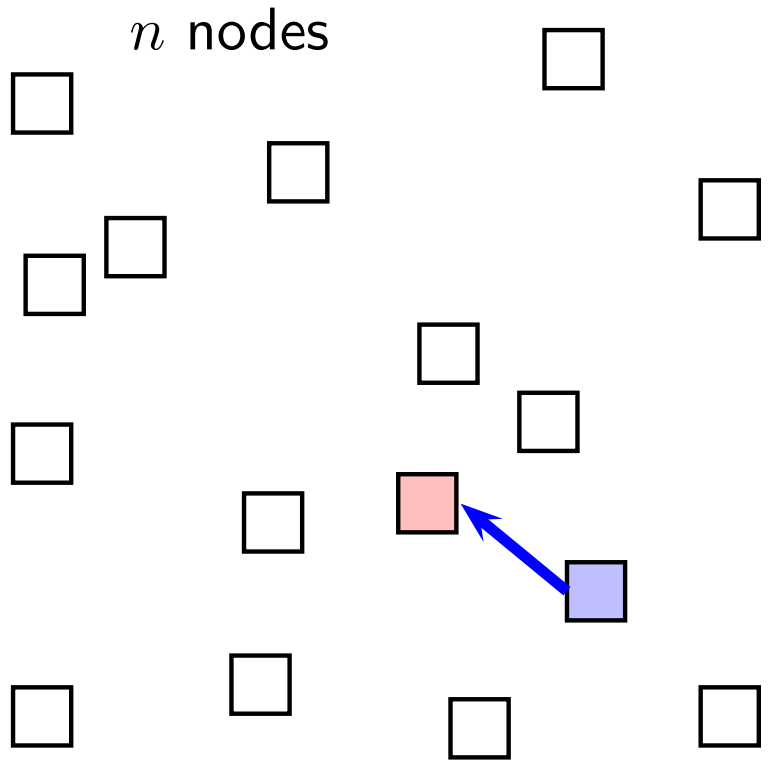
Pairwise Gossip: Convergence

- Gossip Round: a node randomly wakes up and averages with a random neighbor.
- Random geometric graph
- **Boyd-Ghosh-Prabhakar-Shah '06**: Converges in $\Theta(n^2)$ rounds.
- Comes from spectral gap of $\bar{W}$, the averaging matrix.



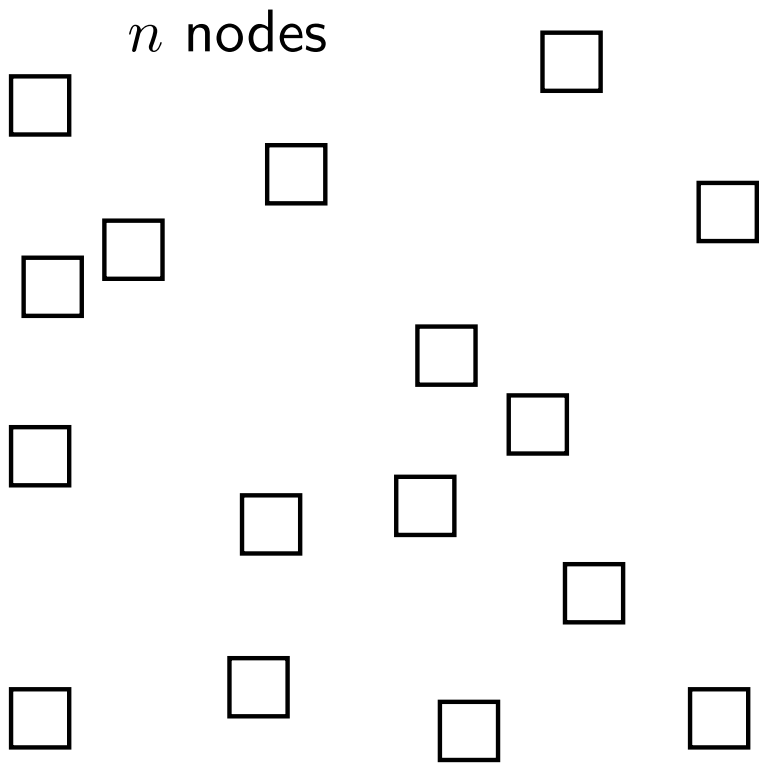
Pairwise Gossip: Convergence

- Gossip Round: a node randomly wakes up and averages with a random neighbor.
- Random geometric graph
- **Boyd-Ghosh-Prabhakar-Shah '06**: Converges in $\Theta(n^2)$ rounds.
- Comes from spectral gap of $\bar{W}$, the averaging matrix.



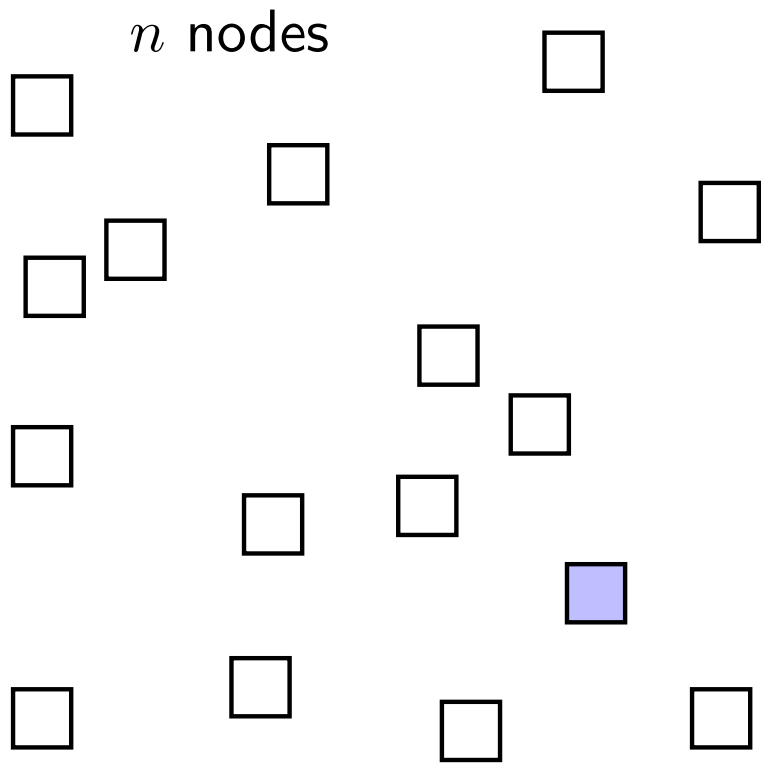
Neighborhood Gossip: Convergence

- Gossip Round: a node randomly wakes up and averages with its entire neighborhood.
- Averaging matrix $\bar{W}$ difficult to compute.
- Idea: Lower bound conductance of a related Markov chain.



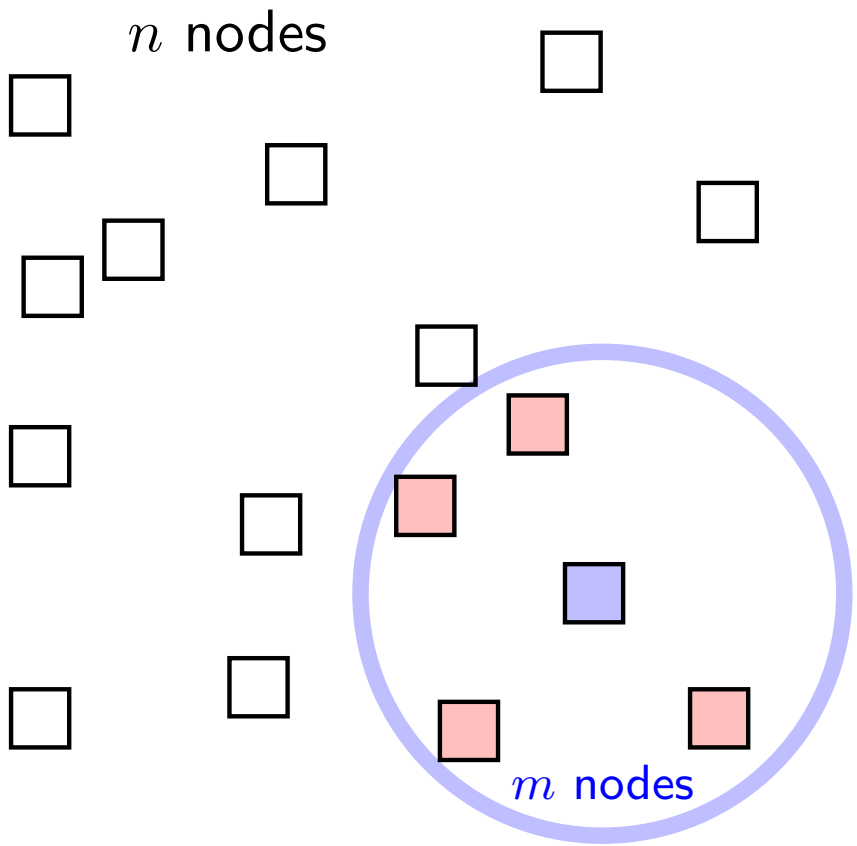
Neighborhood Gossip: Convergence

- Gossip Round: a node randomly wakes up and averages with its entire neighborhood.
- Averaging matrix $\bar{W}$ difficult to compute.
- Idea: Lower bound conductance of a related Markov chain.



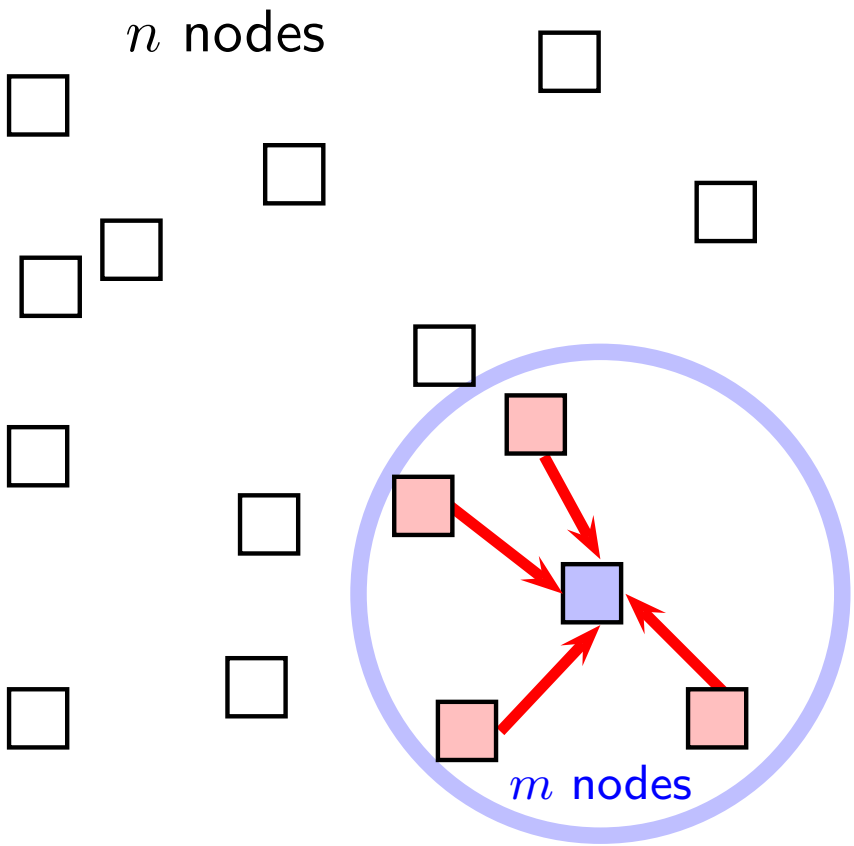
Neighborhood Gossip: Convergence

- Gossip Round: a node randomly wakes up and averages with its entire neighborhood.
- Averaging matrix $\bar{W}$ difficult to compute.
- Idea: Lower bound conductance of a related Markov chain.



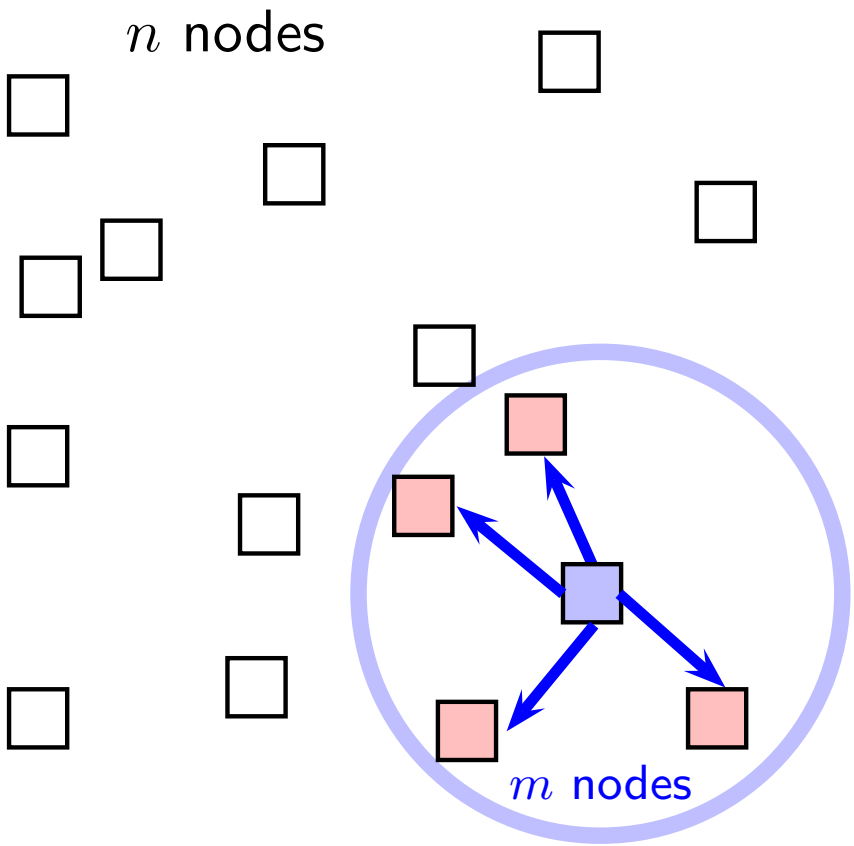
Neighborhood Gossip: Convergence

- Gossip Round: a node randomly wakes up and averages with its entire neighborhood.
- Averaging matrix $\bar{W}$ difficult to compute.
- Idea: Lower bound conductance of a related Markov chain.



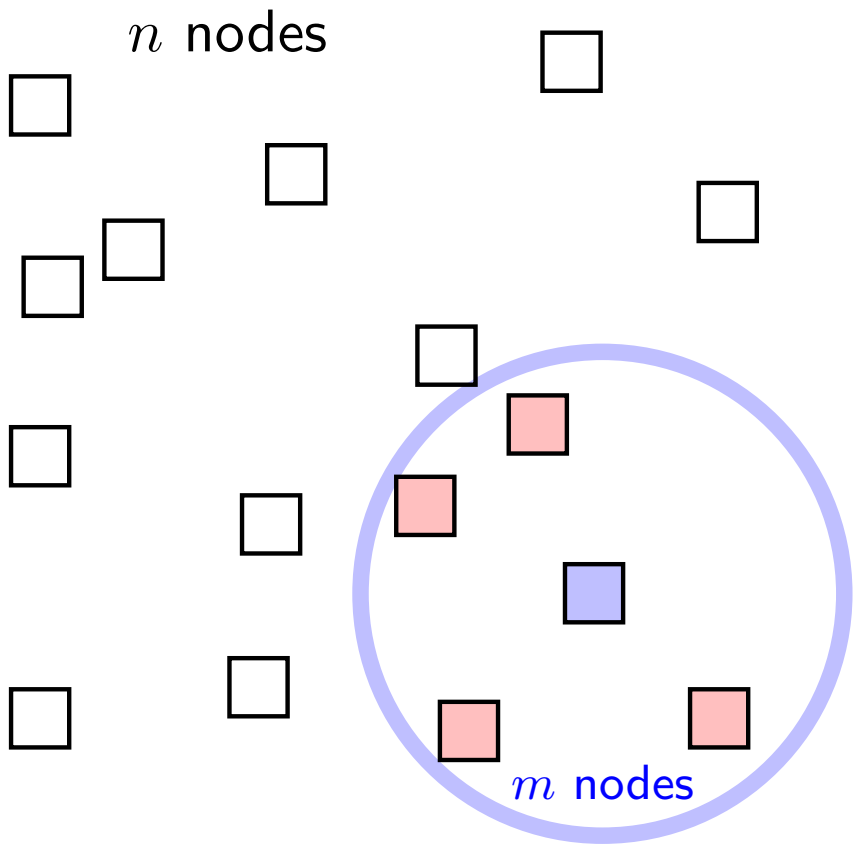
Neighborhood Gossip: Convergence

- Gossip Round: a node randomly wakes up and averages with its entire neighborhood.
- Averaging matrix $\bar{W}$ difficult to compute.
- Idea: Lower bound conductance of a related Markov chain.



Neighborhood Gossip: Convergence

- Gossip Round: a node randomly wakes up and averages with its entire neighborhood.
- Averaging matrix $\bar{W}$ difficult to compute.
- Idea: Lower bound conductance of a related Markov chain.



Theorem: Converges in $O\left(\frac{n^2}{m^2}\right)$ rounds.

Energy Analysis

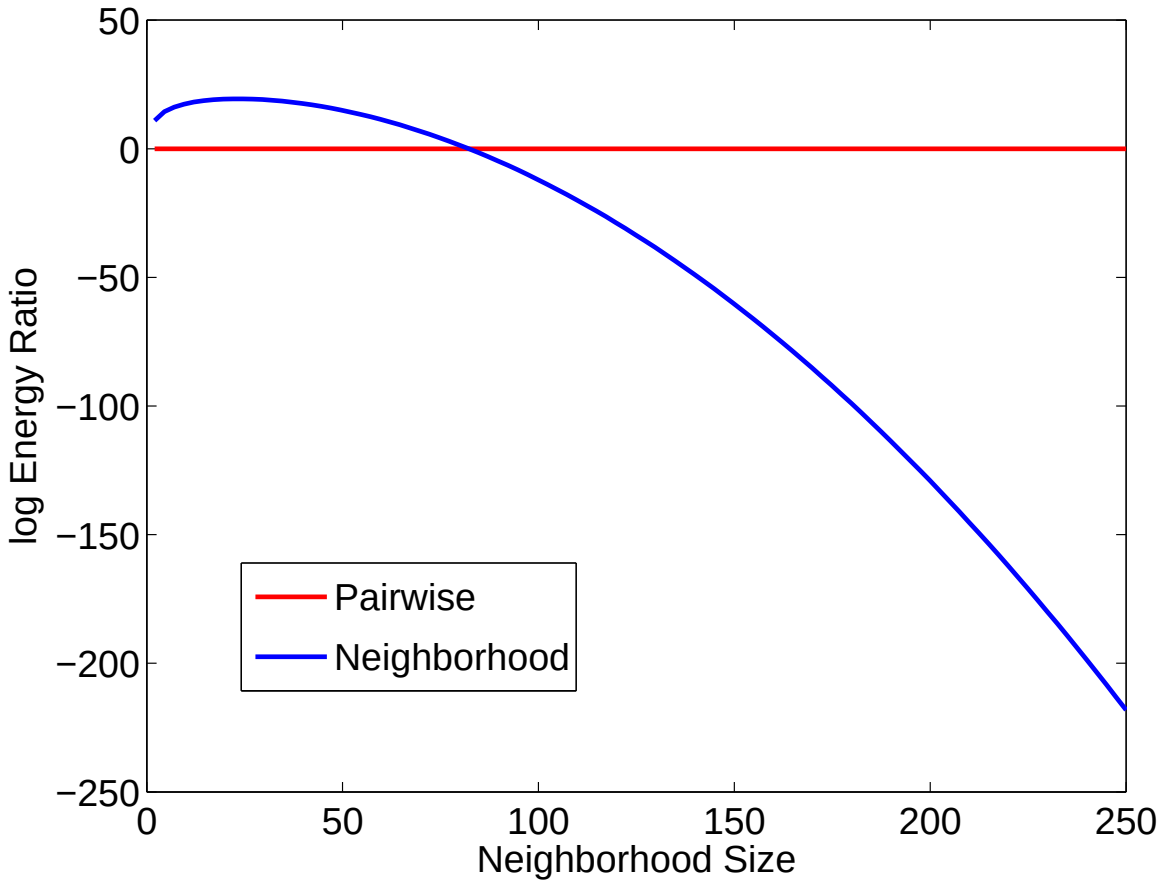
- Want to compare how much energy each scheme uses to converge in time T.
- Measured in total transmit energy:

$$\text{Total Energy} = \sum_{i=1}^{n} \sum_{t=1}^{T} |x_i(t)|^2$$

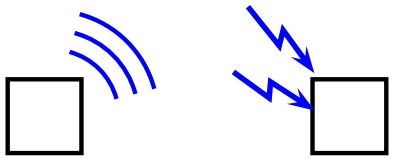
- Need to analyze how much energy used for communication per gossip round in pairwise and neighborhood gossip.

Critical Neighborhood Size

Main Result: If the neighborhood size is larger than a critical value then we get exponential energy savings.



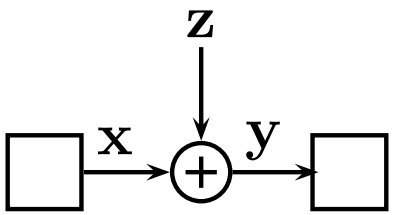
Channel Coding



Channel Coding

- Pairwise Gossip:

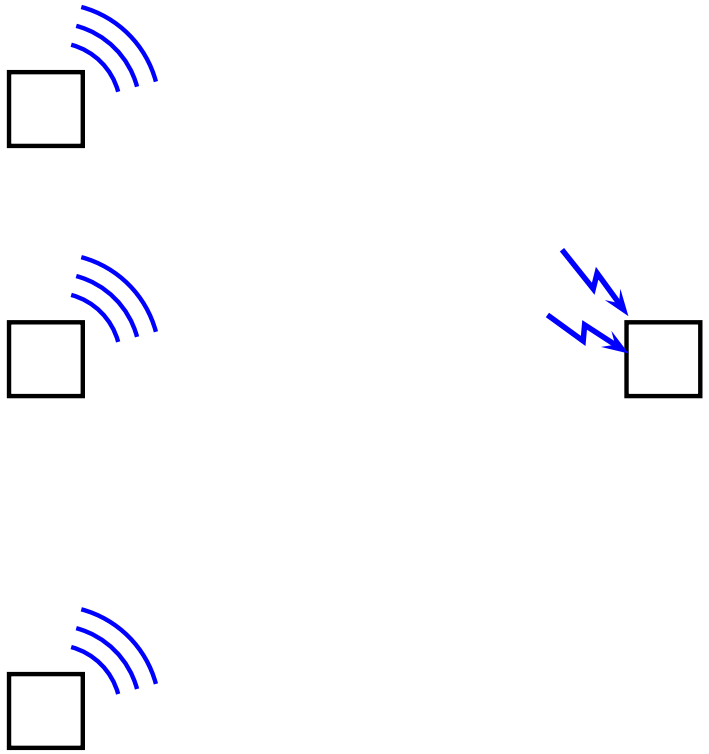
$$R_{\text{PAIR}} = \log(1 + P)$$



Channel Coding

- Pairwise Gossip:

$$R_{\text{PAIR}} = \log(1 + P)$$



Channel Coding

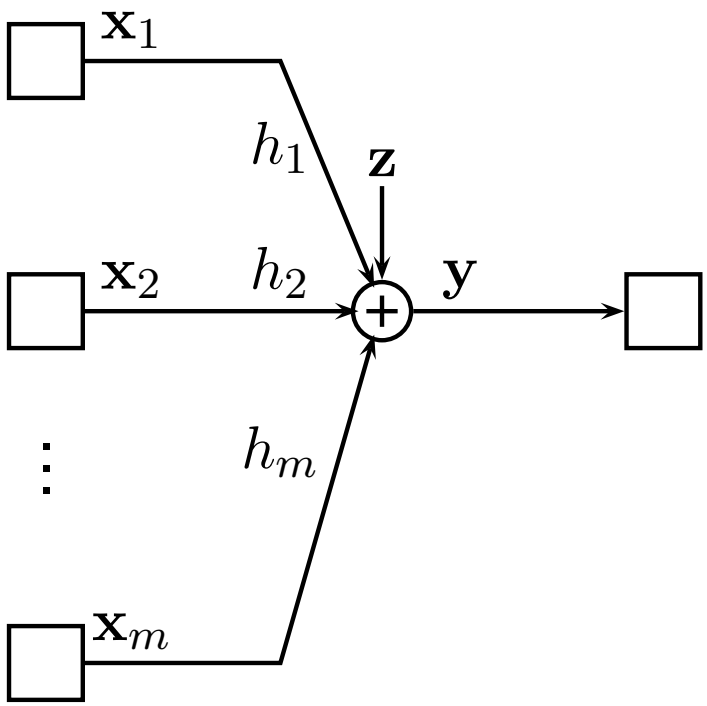
- Pairwise Gossip:

$$R_{\text{PAIR}} = \log(1+P)$$

- Standard Multiple-Access:

$$R_{\text{MAC}} = \frac{1}{m}\log\left(1+m^{-\alpha/2}P\right)$$

$\alpha =$ power path loss coefficient



Channel Coding

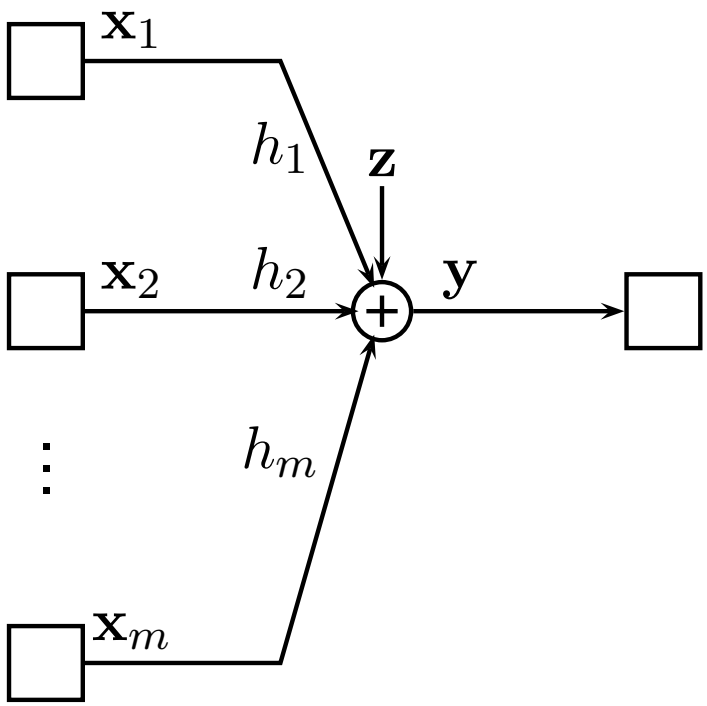
- Pairwise Gossip:

$$R_{\text{PAIR}} = \log(1+P)$$

- Standard Multiple-Access:

$$R_{\text{MAC}} = \frac{1}{m}\log\left(1+m^{-\alpha/2}P\right)$$

$\alpha =$ power path loss coefficient



- Neighborhood gossip with usual channel codes **always** uses more energy than pairwise gossip.

Channel Coding

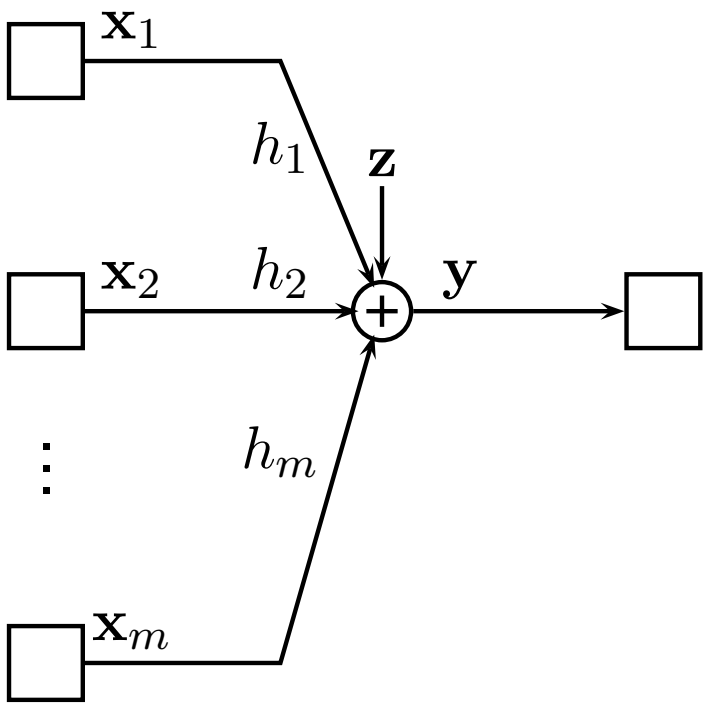
- Pairwise Gossip:

$$R_{\text{PAIR}} = \log(1+P)$$

- Standard Multiple-Access:

$$R_{\text{MAC}} = \frac{1}{m}\log\left(1+m^{-\alpha/2}P\right)$$

$\alpha =$ power path loss coefficient



- Neighborhood gossip with usual channel codes **always** uses more energy than pairwise gossip.
- We want the channel to do the averaging for us!

Channel Coding

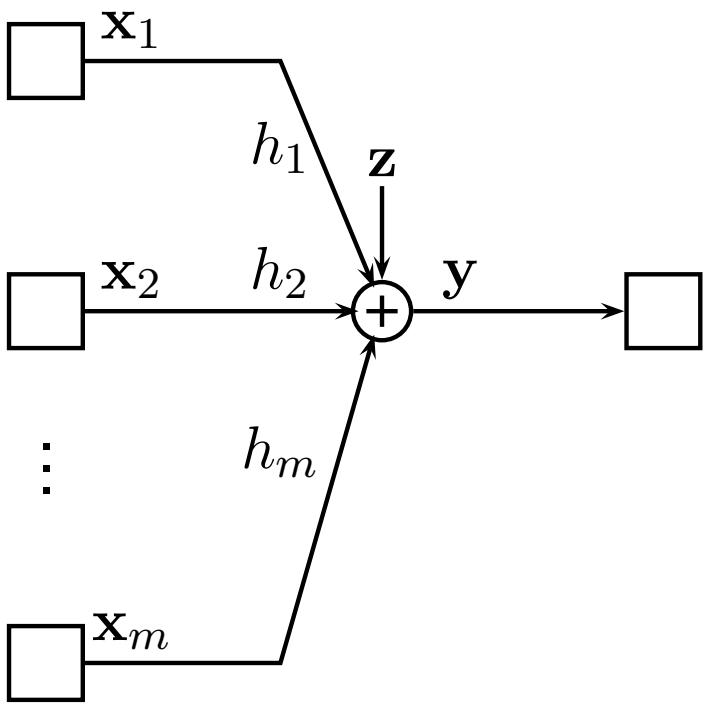
- Pairwise Gossip:

$$R_{\text{PAIR}} = \log(1+P)$$

- Standard Multiple-Access:

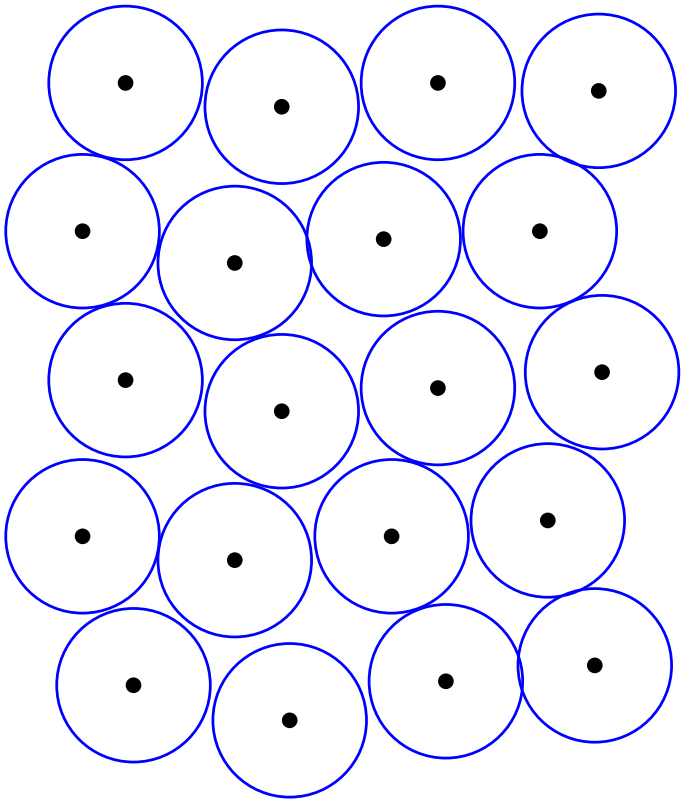
$$R_{\text{MAC}} = \frac{1}{m}\log\left(1+m^{-\alpha/2}P\right)$$

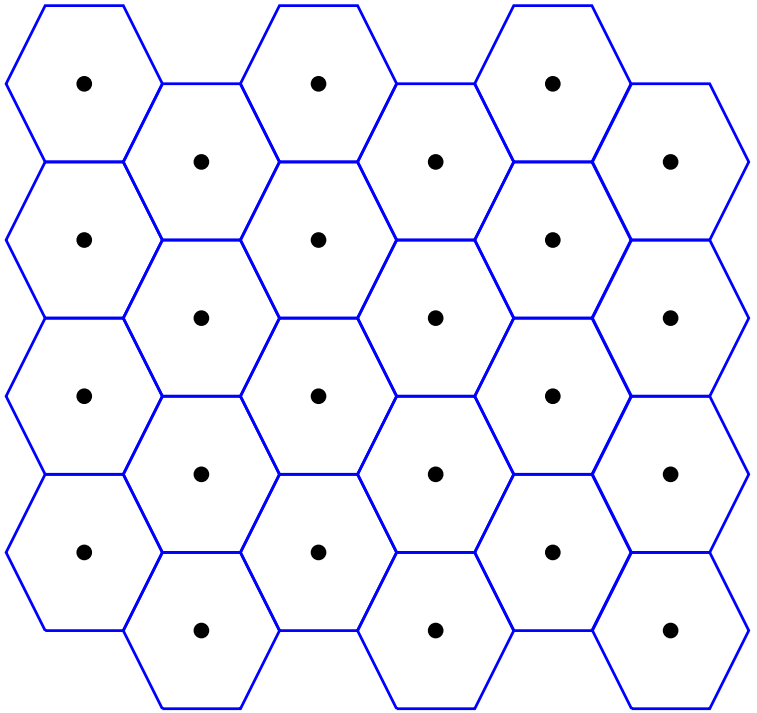
$\alpha =$ power path loss coefficient



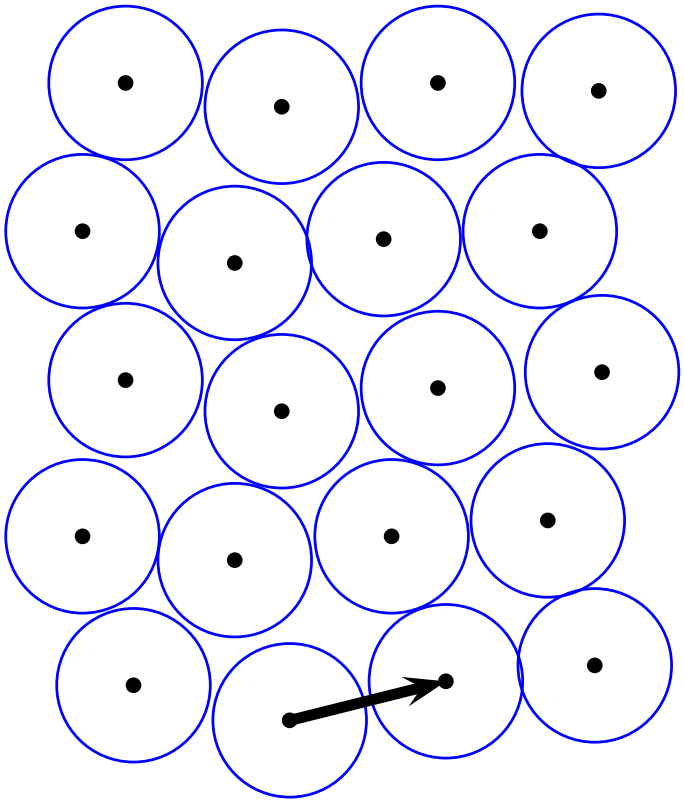
- Neighborhood gossip with usual channel codes **always** uses more energy than pairwise gossip.
- We want the channel to do the averaging for us!
- This is possible if each transmitter uses the same linear code.

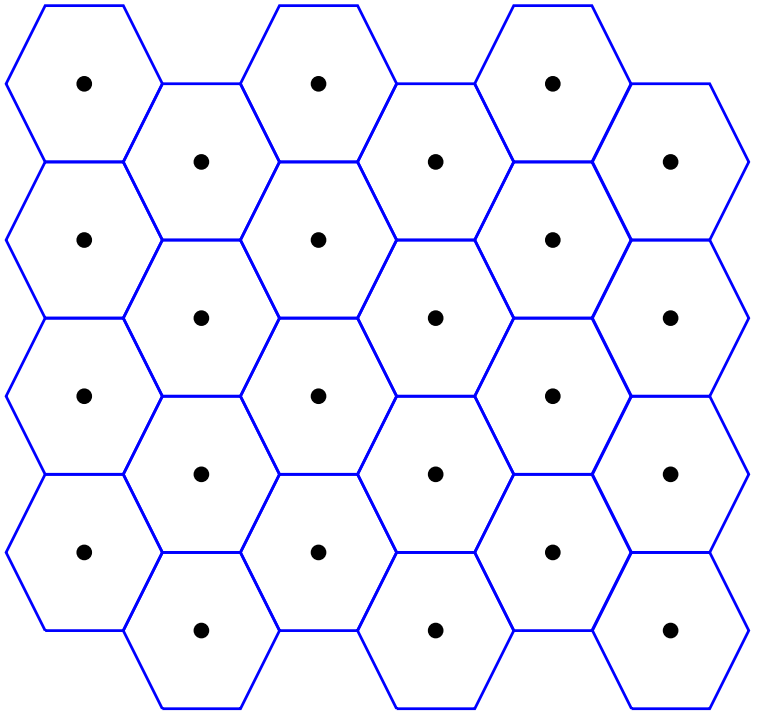
Benefits of Linear Coding



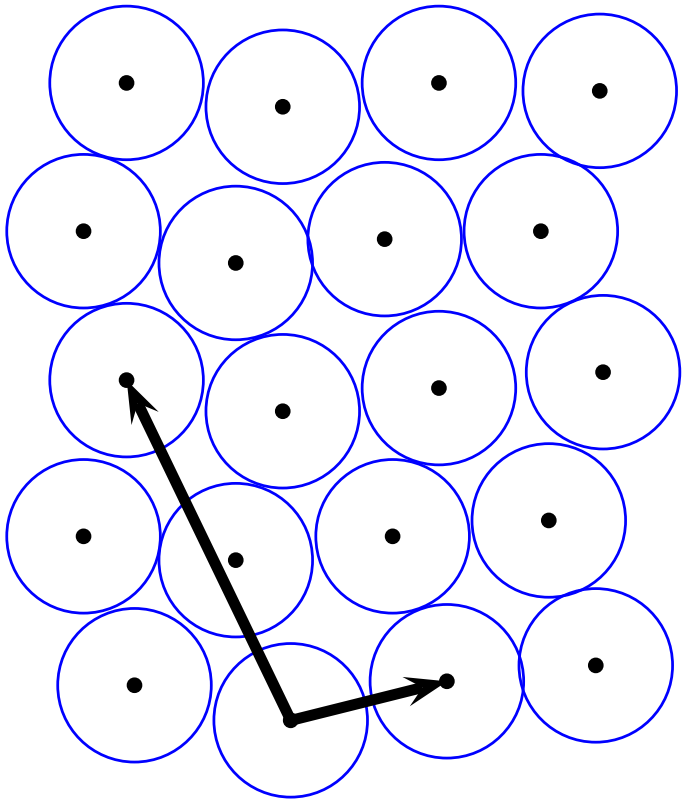


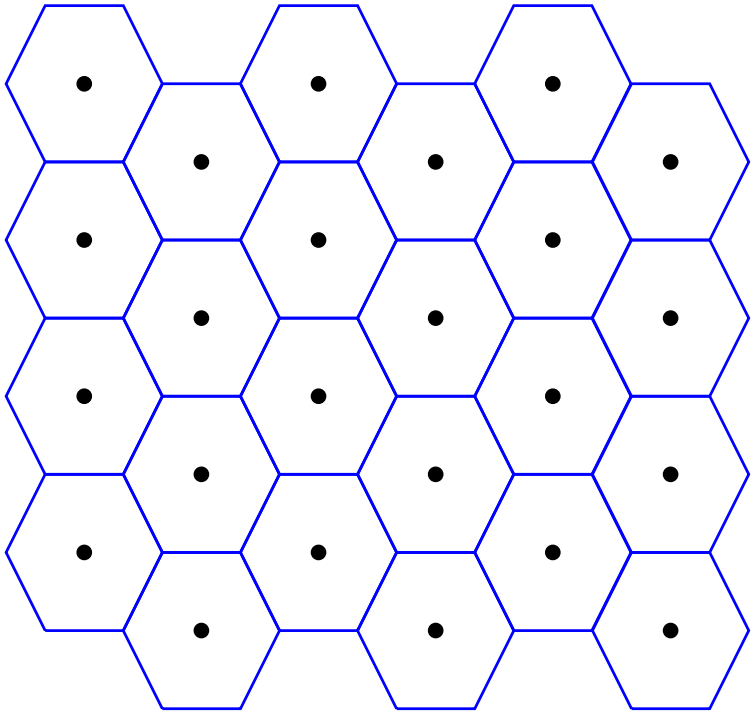
Benefits of Linear Coding



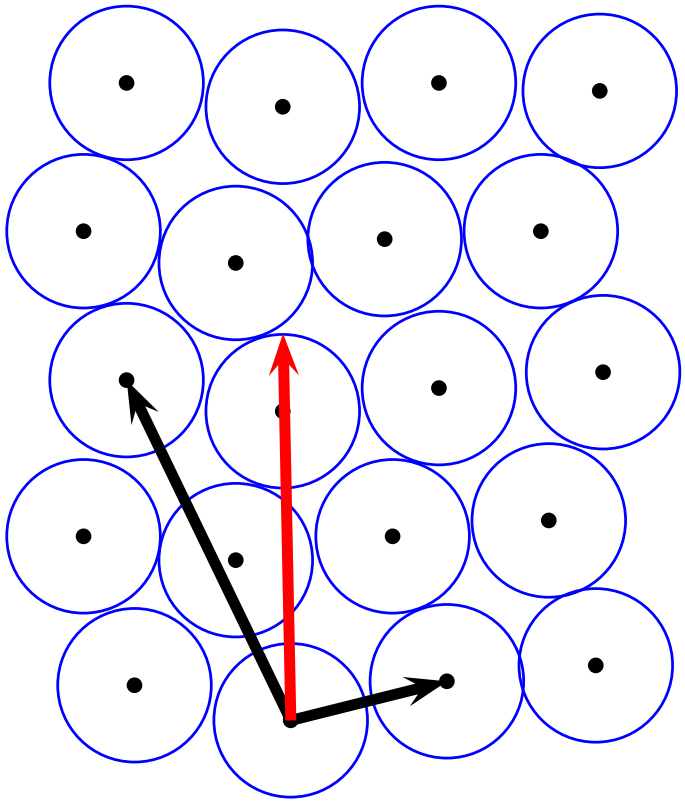


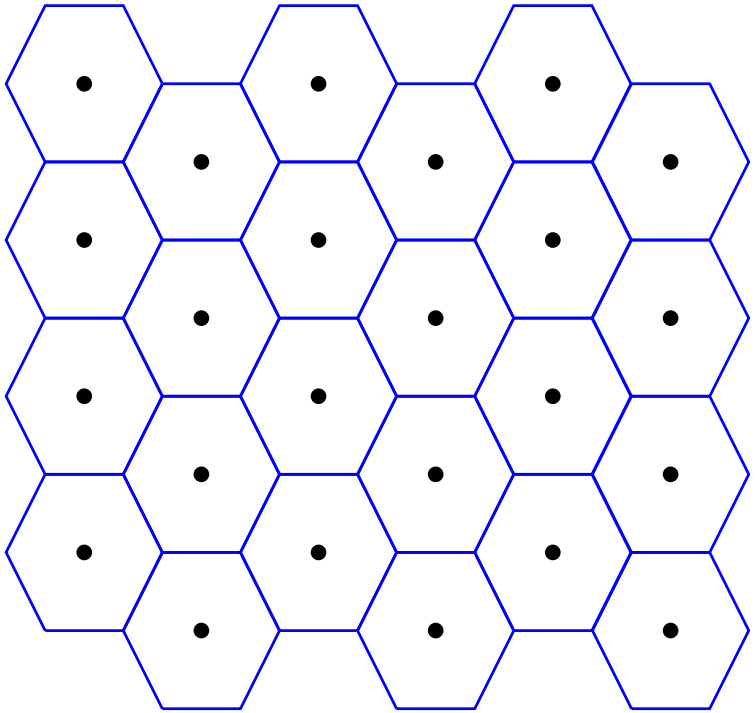
Benefits of Linear Coding





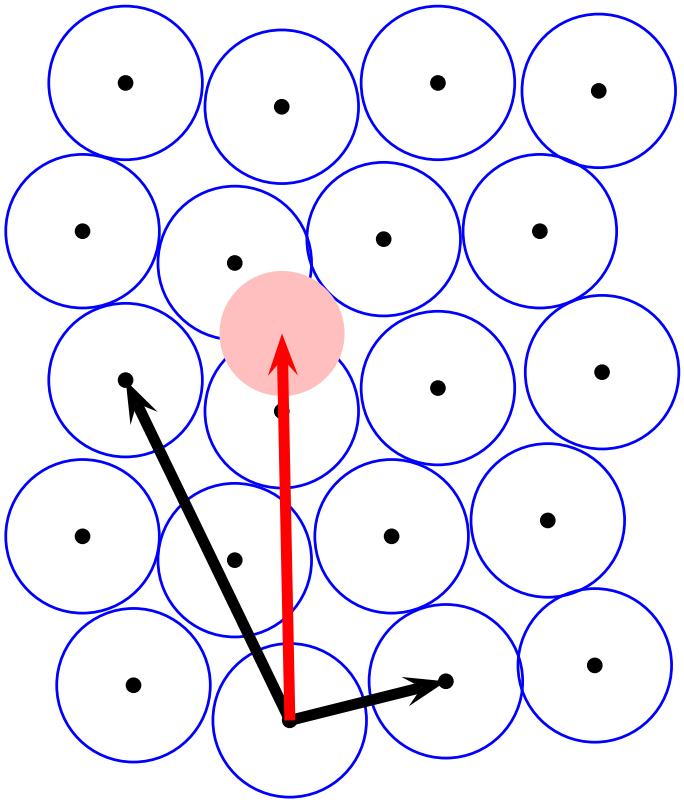
Benefits of Linear Coding

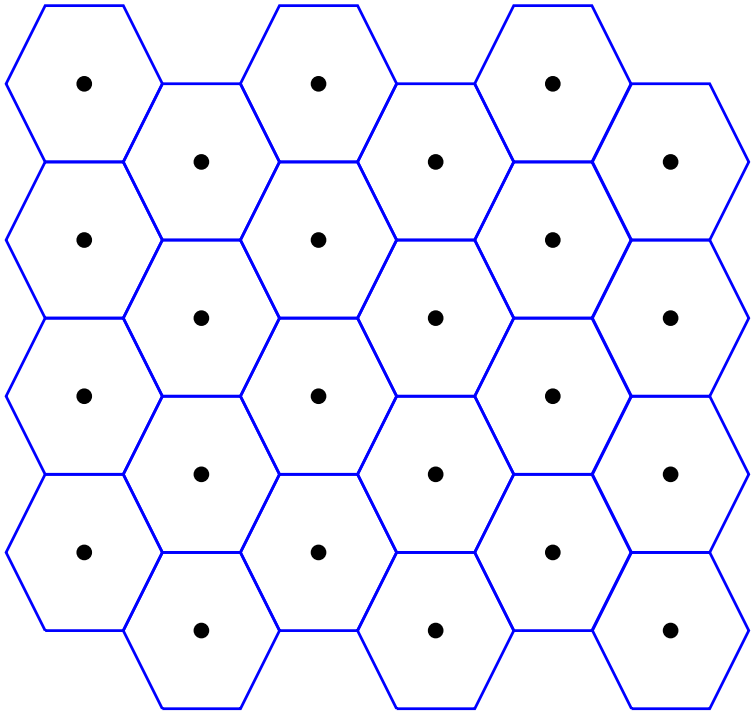




- Sum of codewords is not a codeword.
- Must decode individual messages.

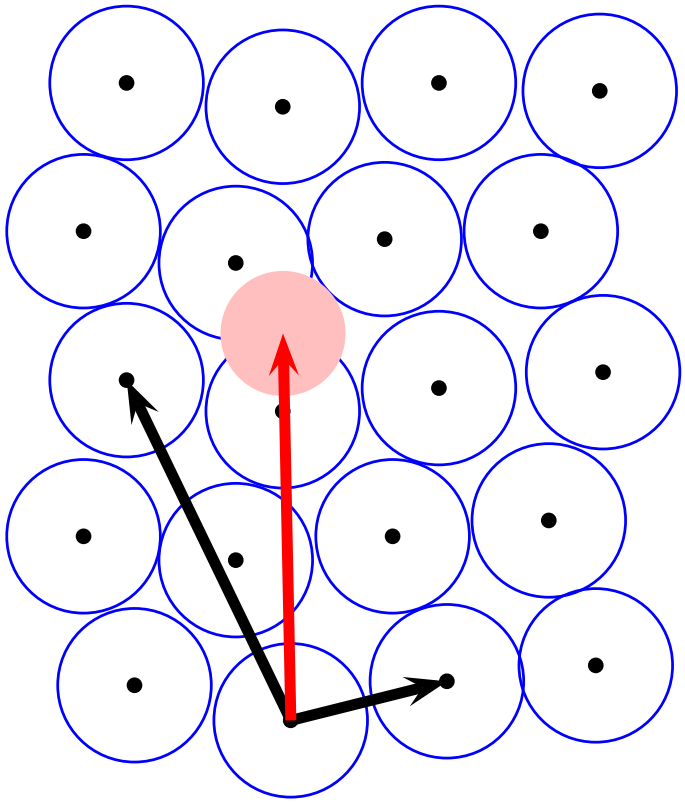
Benefits of Linear Coding

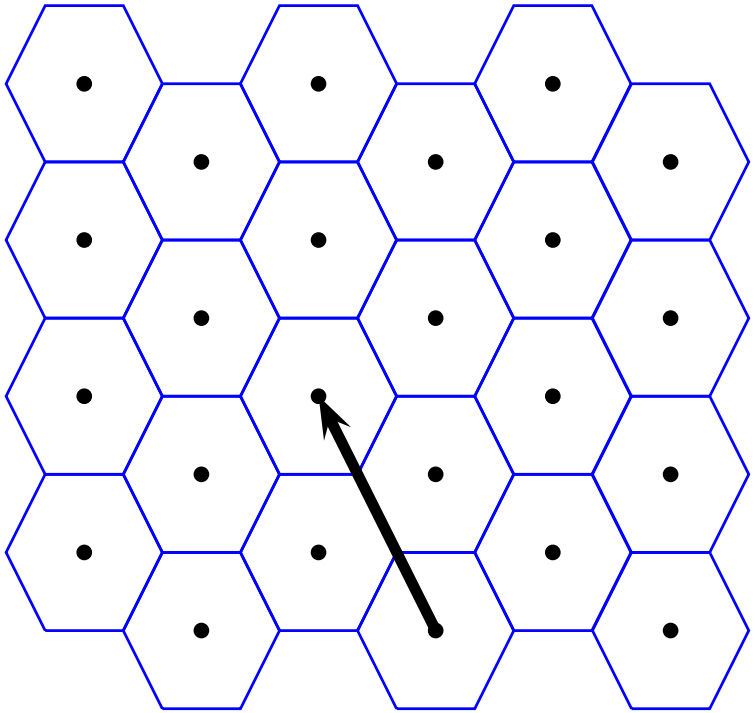




- Sum of codewords is not a codeword.
- Must decode individual messages.

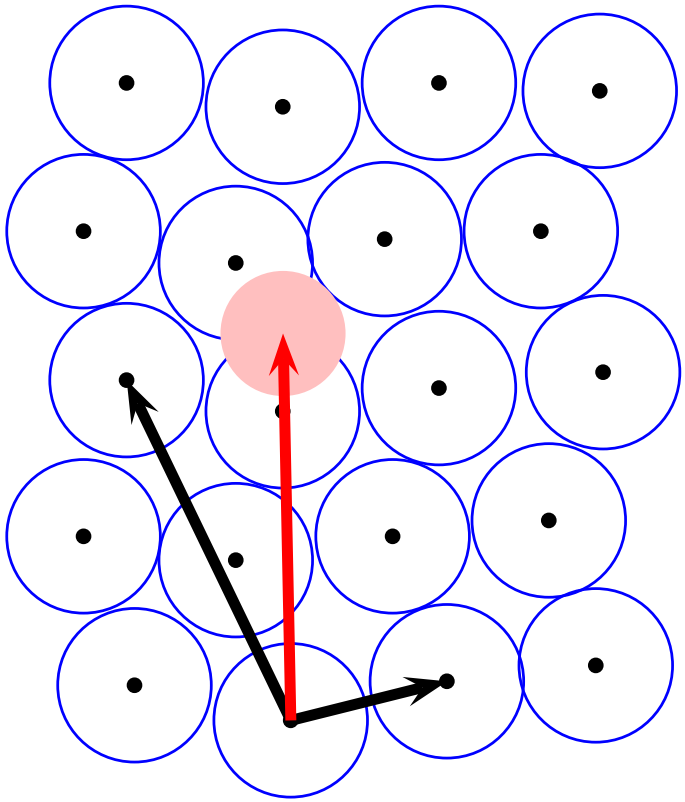
Benefits of Linear Coding



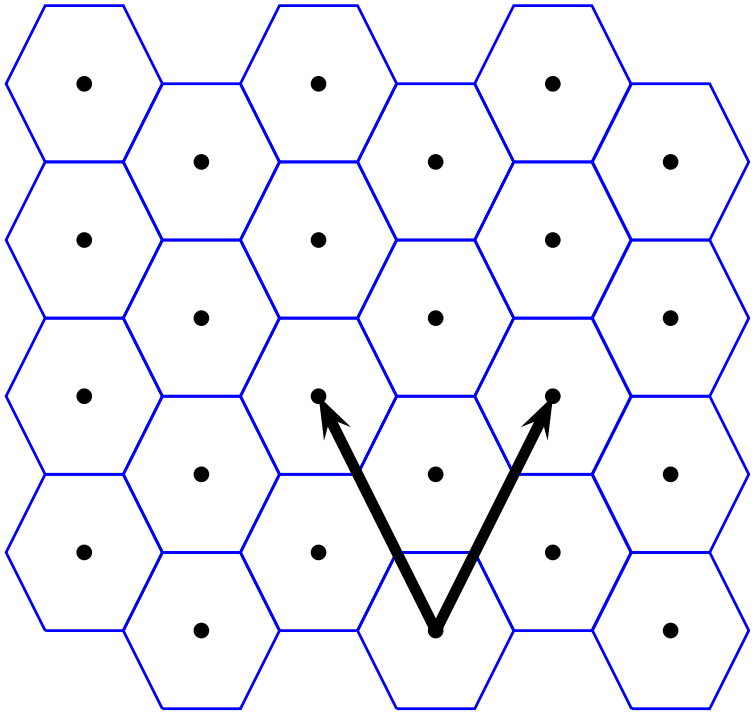


- Sum of codewords is not a codeword.
- Must decode individual messages.

Benefits of Linear Coding

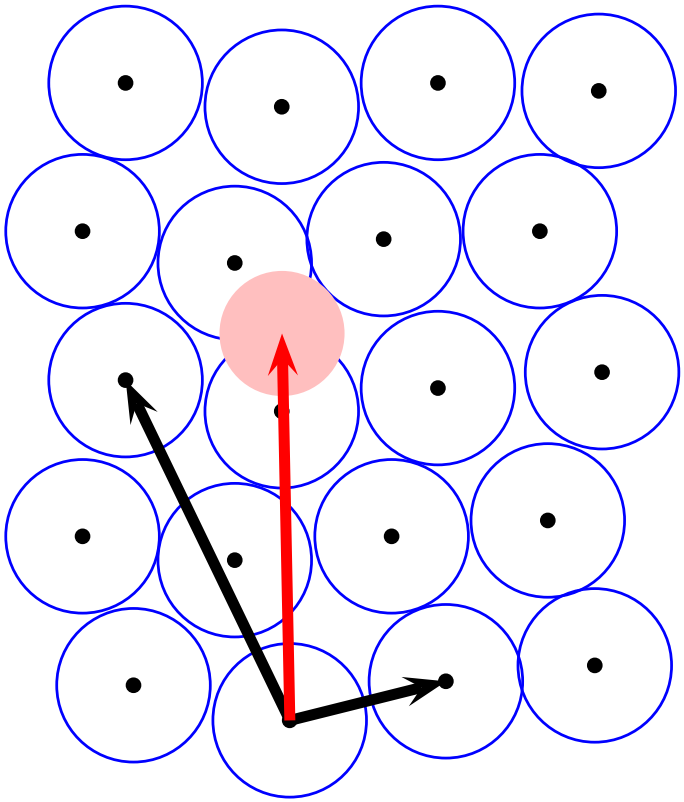


- Sum of codewords is not a codeword.
- Must decode individual messages.

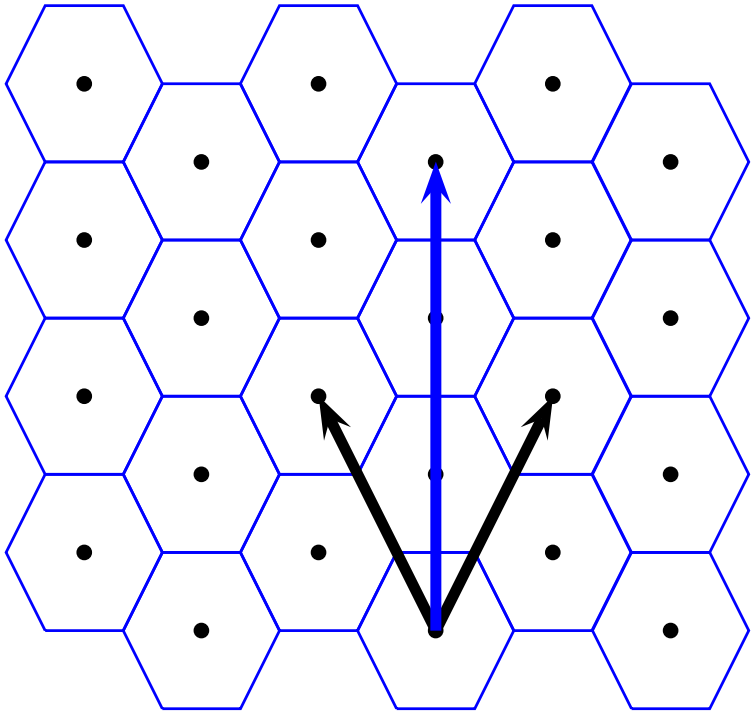


- Sum of codewords is a codeword.
- Can decode linear functions of messages.

Benefits of Linear Coding

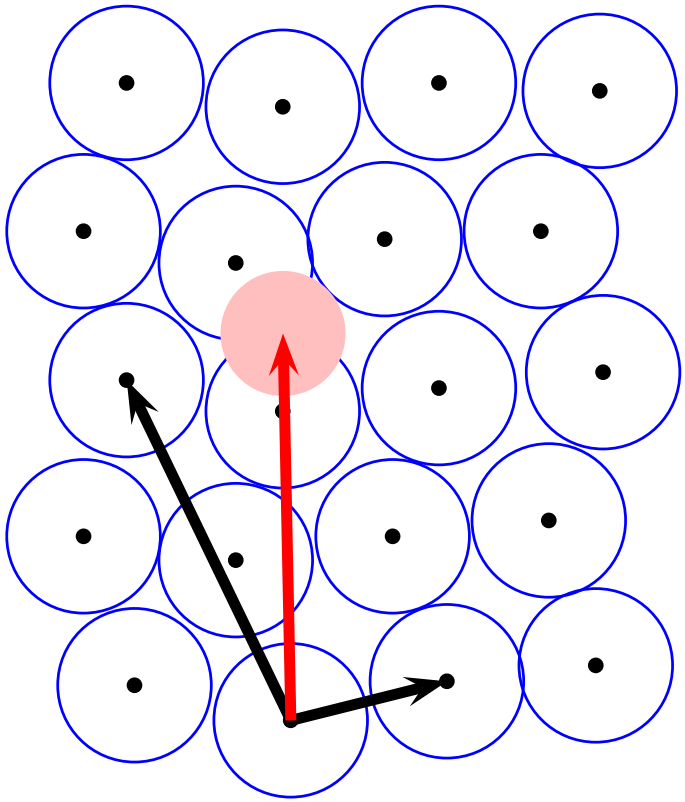


- Sum of codewords is not a codeword.
- Must decode individual messages.

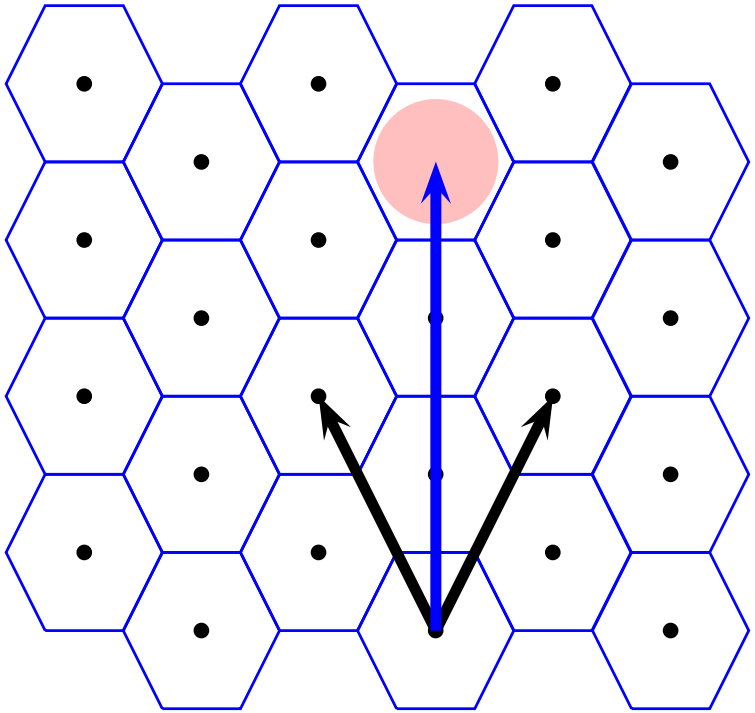


- Sum of codewords is a codeword.
- Can decode linear functions of messages.

Benefits of Linear Coding



- Sum of codewords is not a codeword.
- Must decode individual messages.



- Sum of codewords is a codeword.
- Can decode linear functions of messages.

Computation Coding

- Pairwise Gossip:

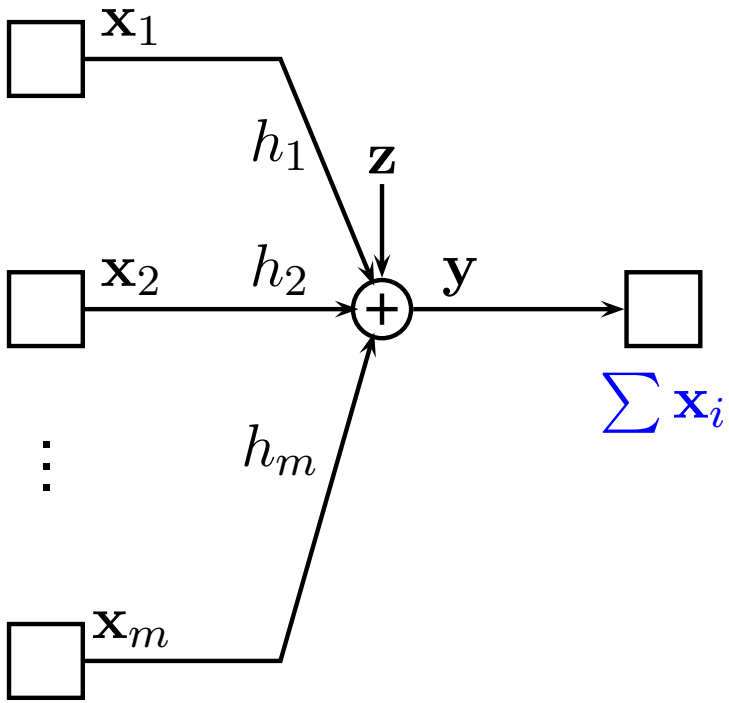
$$R_{\mathsf{PAIR}} = \log\left(1+P\right)$$

- Standard Multiple-Access:

$$R_{\mathsf{MAC}} = \frac{1}{m}\log\left(1+m^{-\alpha/2}P\right)$$

- Computation Coding:

$$R_{\mathsf{COMP}} = \log\left(\frac{1}{m}+m^{-\alpha/2}P\right)$$



- See **Nazer-Gastpar IT Trans '07** for more details.
- Does this save us any energy?

Unlimited Time

- If we do not care about convergence time, then pairwise gossip always uses less energy.
- This is because neighborhood gossip has to transmit over much longer distances ($\sqrt{m}$ distance for neighborhood of size m).
- We are really interested in the time-energy tradeoff.

Time-Energy Tradeoff

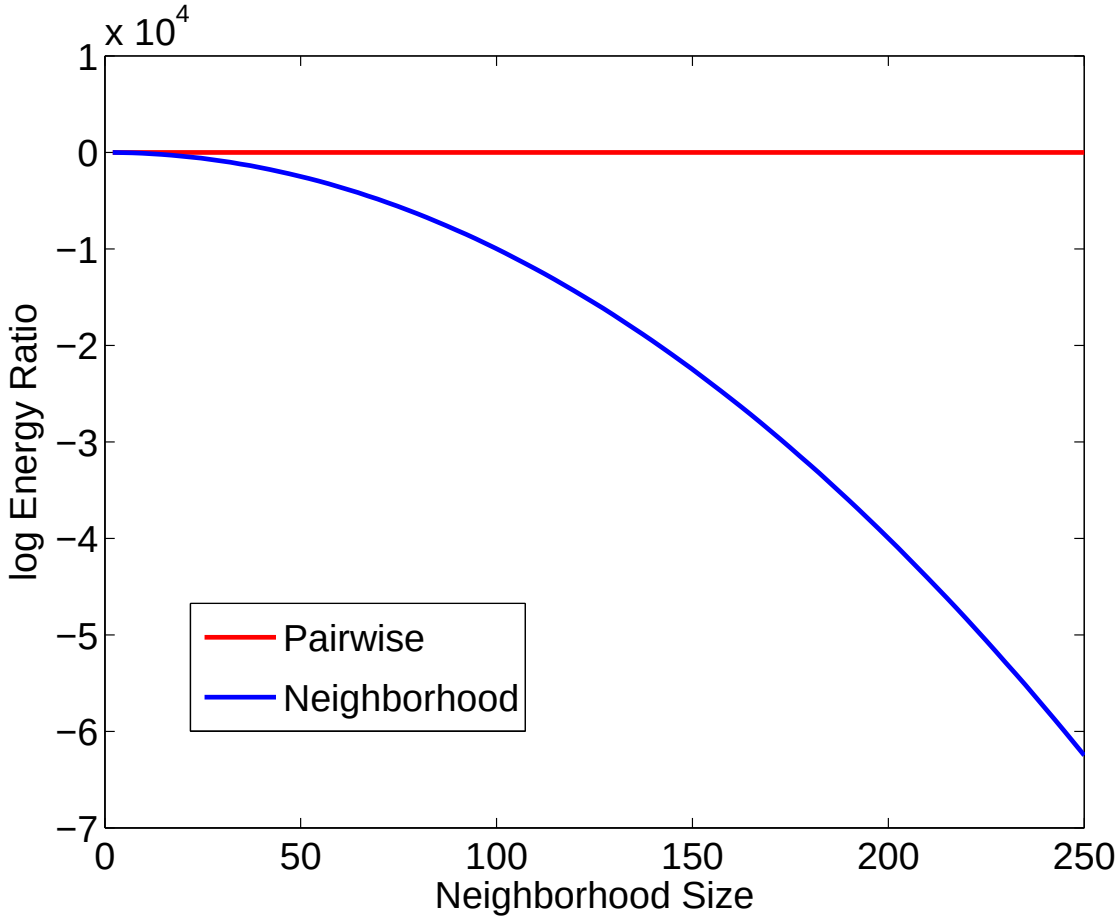
- $n =$ number of nodes
- $m =$ neighborhood size
- $\alpha =$ power path-loss coefficient
- $\tau =$ speed-up factor $= \dfrac{\text{Pairwise Convergence Time}}{\text{Neighborhood Convergence Time}}$

$$\log \frac{\text{Pairwise Energy}}{\text{Neighborhood Energy}} = \left(\frac{\alpha}{2} + 2\right) \log m - \log \tau - \frac{m^2}{\tau}$$

Exponential energy savings possible if the neighborhood size scales with the network size!

Exponential Energy Savings

- Energy savings increase as the neighborhood size increases.
- Is this a fair comparison?



Making a Fair Comparison

- **Issue 1:** Multiple gossip rounds could happen at once.

Making a Fair Comparison

- **Issue 1:** Multiple gossip rounds could happen at once.
- **Solution 1:** Assume all nodes can pairwise gossip at once.
  Only one neighborhood gossip round at at time.

Making a Fair Comparison

- **Issue 1:** Multiple gossip rounds could happen at once.
- **Solution 1:** Assume all nodes can pairwise gossip at once.
  Only one neighborhood gossip round at at time.
- Penalty 1: n factor slow-down

Making a Fair Comparison

- **Issue 1:** Multiple gossip rounds could happen at once.
- **Solution 1:** Assume all nodes can pairwise gossip at once.
  Only one neighborhood gossip round at at time.
- Penalty 1: n factor slow-down

- **Issue 2:** Quantization error could build up.

Quantization references: **Nedic et al. '07, Frasca et al. '08, Aysal et al. '08, Kar et al. '09**

Making a Fair Comparison

- **Issue 1:** Multiple gossip rounds could happen at once.
- **Solution 1:** Assume all nodes can pairwise gossip at once.
  Only one neighborhood gossip round at at time.
- Penalty 1: n factor slow-down

- **Issue 2:** Quantization error could build up.
- **Solution 2:** Assume no build up for pairwise gossip.
  Assume worst-case build up for neighborhood gossip.

Quantization references: **Nedic et al. '07, Frasca et al. '08, Aysal et al. '08, Kar et al. '09**

Making a Fair Comparison

- **Issue 1:** Multiple gossip rounds could happen at once.
- **Solution 1:** Assume all nodes can pairwise gossip at once.
  Only one neighborhood gossip round at at time.
- Penalty 1: n factor slow-down

- **Issue 2:** Quantization error could build up.
- **Solution 2:** Assume no build up for pairwise gossip.
  Assume worst-case build up for neighborhood gossip.
- Penalty 2: $\log n$ extra quantization bits

Quantization references: **Nedic et al. '07, Frasca et al. '08, Aysal et al. '08, Kar et al. '09**

Time-Energy Tradeoff Revised

- $n =$ number of nodes
- $m =$ neighborhood size
- $\alpha =$ power path-loss coefficient
- $\tau =$ speed-up factor $= \dfrac{\text{Pairwise Convergence Time}}{\text{Neighborhood Convergence Time}}$

$$\log \frac{\text{Pairwise Energy}}{\text{Neighborhood Energy}} = \left(\frac{\alpha}{2} + 2\right) \log m - \log \tau - \frac{m^2}{\tau}$$

Exponential energy savings still possible if the neighborhood is large enough!

Time-Energy Tradeoff Revised

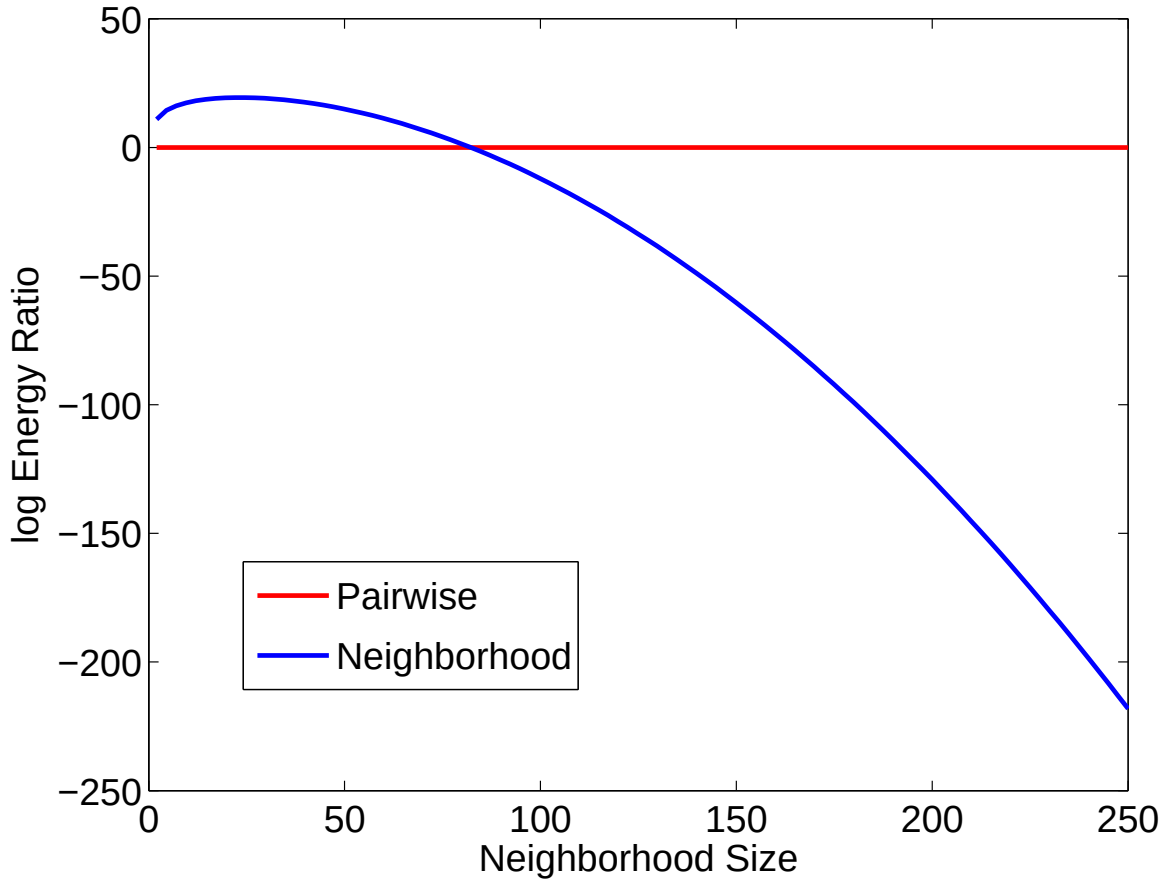
- $n =$ number of nodes
- $m =$ neighborhood size
- $\alpha =$ power path-loss coefficient
- $\tau =$ speed-up factor $= \dfrac{\text{Pairwise Convergence Time}}{\text{Neighborhood Convergence Time}}$

$$\log \frac{\text{Pairwise Energy}}{\text{Neighborhood Energy}} = \left(\frac{\alpha}{2} + 2\right) \log m + \log n - \log \tau - \frac{m^2}{\tau n}$$

Exponential energy savings still possible if the neighborhood is large enough!

Critical Neighborhood Size

Exponential energy savings if the neighborhood is larger than a critical value that depends on the power path-loss coefficient and the speed-up factor.



Equations Instead of Bits

- By employing a **new abstraction** for the physical layer, we can save both time and energy.
- Computation coding exploits the interference of the wireless channel using structured codes to reliably and efficiently compute equations.
- Can be used in a modular fashion for other gossip algorithms and sensor network topologies.
- Computation codes useful in many other multi-user communication scenarios. See **Nazer-Gastpar IT '07, ETT '08**.

Conclusions

Neighborhood gossip can exploit wireless interference for faster and cheaper distributed averaging.

Related work exploits other properties for better gossip:

- Broadcast Gossip (**Aysal-Yildiz-Sarwate-Scaglione '08**)
- Geographic Gossip (**Dimakis-Sarwate-Wainwright '06,'08, Rabbat-Haupt-Singh-Nowak '06, Li-Dai '08**)
- Path Averaging (**Benezit-Dimakis-Thiran-Vetterli '07**)
- Mobility (**Sarwate-Dimakis**)
- Much more ...