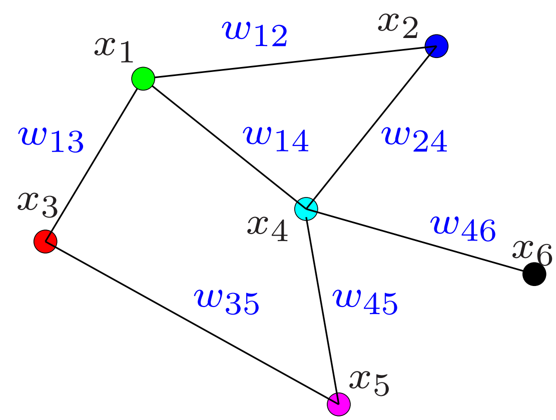
Distributed Consensus: Convergence, Robustness, and Optimization

Lin Xiao
Microsoft Research

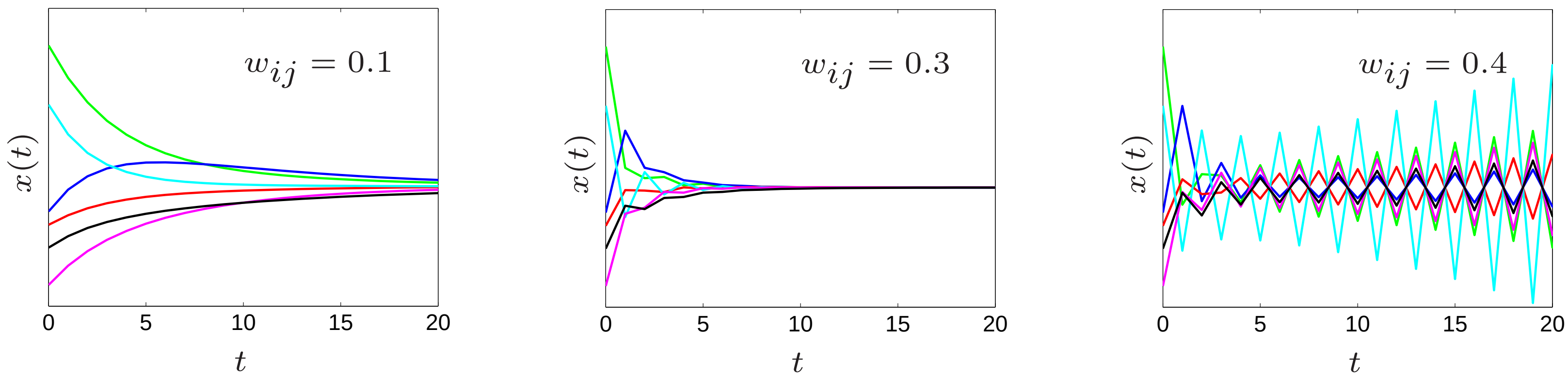
Mathematical Challenges and Opportunities in Sensor Networking
IPAM, UCLA
January 10, 2007

Distributed average consensus

- **basic algorithm**
  - initial values at each node $x_i(0)$
  - every node to compute $\frac{1}{n}\sum x_i(0)$



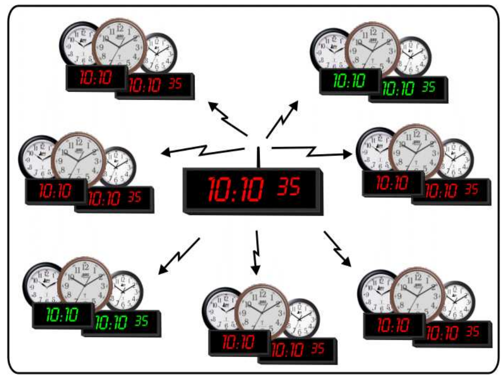
$$x_i(t+1) = x_i(t) + \sum_{j \in \mathcal{N}_i} w_{ij}\Big(x_j(t) - x_i(t)\Big)$$

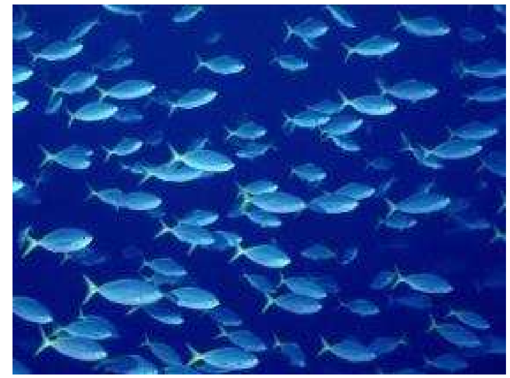


- design **local parameters** w_{ij} to achieve **global performance**
  - convergence, robustness to unreliable links, noises, topology changes
  - optimization for fastest convergence, minimum communication costs

Applications of distributed consensus

- distributed coordination, synchronization, flocking, and load balancing



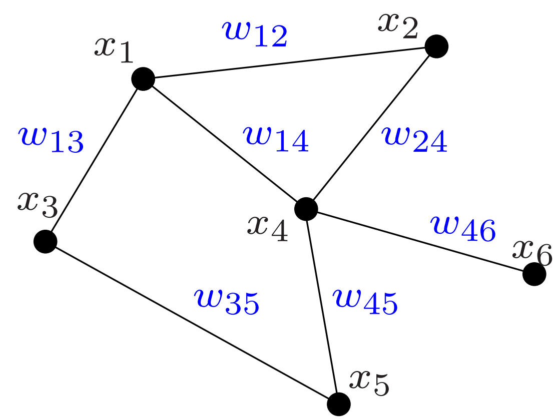


Tsitsiklis (1984), Jadbabaie, Lin & Morse (2003), Olfati-Saber & Murray (2004), X. & Boyd (2004), Blondel, Hendricks, Olshevsky & Tsitsiklis (2005), Moreau (2005), etc.

- information processing in sensor networks
  - distributed parameter estimation, distributed Kalman filtering
  - distributed detection/inference

Grime & Durrant-Whyte (1994), Scherber & Papadopolos (2004), Alanyali, Venkatesh, Savas & Aeron (2004), Spanos, Olfati-Saber & Murray (2005), Aldosari & Moura (2005), X., Boyd & Lall (2005), etc.

Mathematical challenges and opportunities



- basic algorithm

$$x_i(t+1) = x_i(t) + \sum_{j \in \mathcal{N}_i} w_{ij} \Big(x_j(t) - x_i(t) \Big)$$

- challenges

  convergence, robustness, optimization

- mathematical tools
  - linear algebra
  - spectral graph theory
  - convex optimization
  - duality

- connections: Markov chains, sensor localization, dimensionality reduction

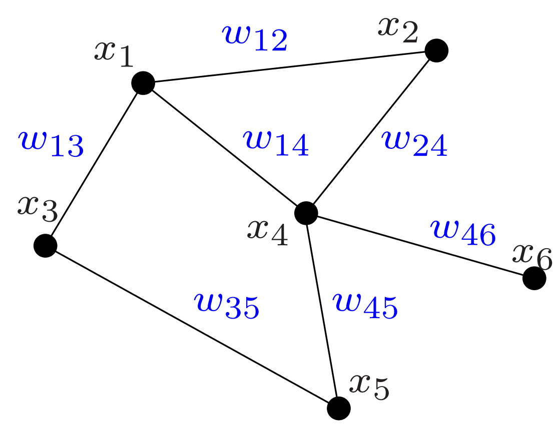
Outline

- **convergence on fixed graph**
- **optimization on fixed graph**
- **robustness to random link failures**
- **optimization of randomized gossip**

Convergence on fixed graph

- basic algorithm

$$x_i(t+1) = x_i(t) + \sum_{j \in \mathcal{N}_i} w_{ij}\Big(x_j(t) - x_i(t)\Big)$$



- vector form: $x(t+1) = Wx(t)$

$$W_{ij} = \begin{cases} w_{ij} & \{i,j\} \in \mathcal{E} \\ 0 & \{i,j\} \notin \mathcal{E} \\ 1 - \sum_k w_{ik} & i = j \end{cases} \qquad (W = I - L)$$

- **theorem** (XB'04): $\lim_{t\to\infty} x_i(t) = \frac{1}{n}\sum_{j=1}^n x_j(0)$ for all $x(0) \in \mathbf{R}^n$ iff

$$\mathbf{1}^T W = \mathbf{1}^T, \qquad W\mathbf{1} = \mathbf{1}, \qquad \rho\left(W - (1/n)\mathbf{1}\mathbf{1}^T\right) < 1$$

Optimal design for fastest convergence

- for symmetric weights, define convergence factor

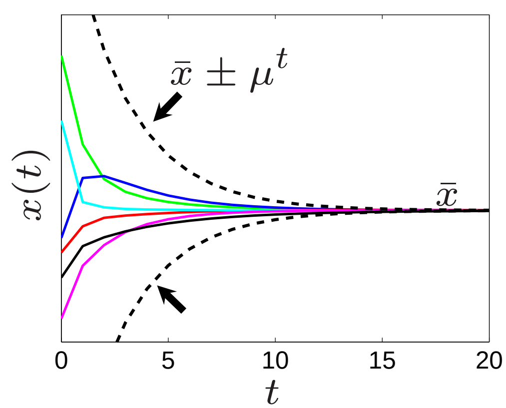
$$\mu(W) = \rho\left(W - (1/n)\mathbf{1}\mathbf{1}^T\right) = \|W - (1/n)\mathbf{1}\mathbf{1}^T\|$$

- convergence bound

$$\|x(t) - \bar{x}\mathbf{1}\|_2 \leq \mu(W)^t \, \|x(0) - \bar{x}\mathbf{1}\|_2$$

- optimization with symmetric weights

$$\begin{array}{ll} \text{minimize} & \mu(W) \\ \text{subject to} & W = W^T, \quad W\mathbf{1} = \mathbf{1} \\ & W_{ij} = 0, \quad \{i,j\} \notin \mathcal{E} \end{array}$$

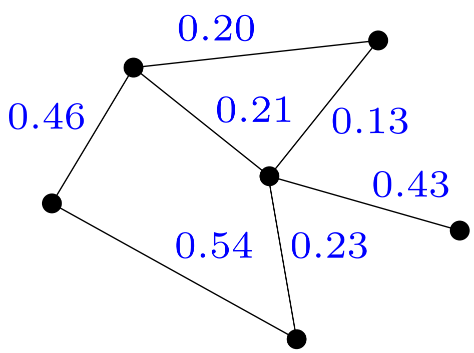


a **convex** optimization problem, globally and efficiently solved

A small example

FMMC

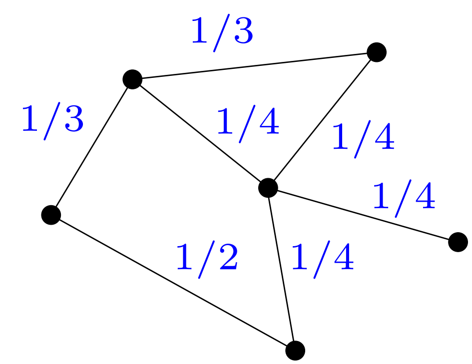
$$W^\star = \arg\min_{W \in \mathcal{C}} \mu(W)$$



$$\mu^\star \quad = \quad 0.72$$

Metropolis

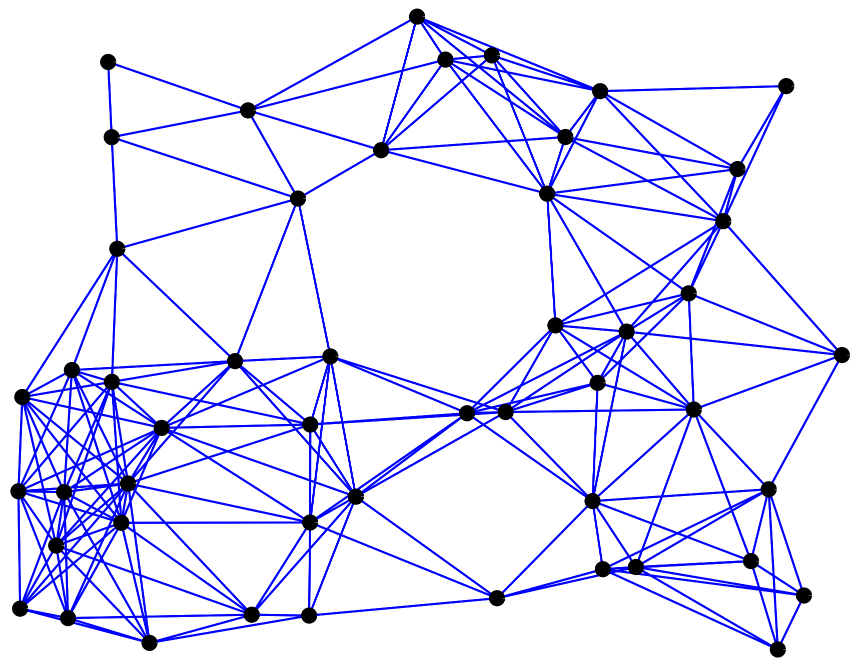
$$W_{ij} = \frac{1}{\max\{d_i,\ d_j\}}$$



$$\mu \quad = \quad 0.77$$

A larger example

- randomly generated network with 50 nodes, 200 edges



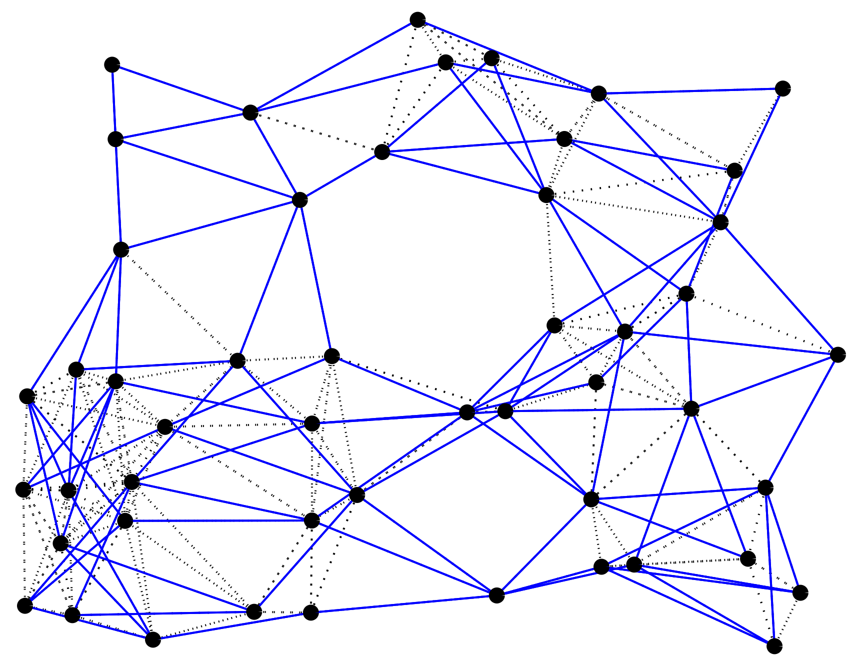
	Metropolis	best constant	optimal
$\rho(W - \mathbf{1}\mathbf{1}^T/n)$	0.949	0.947	0.902
$\tau = 1/\log(1/\rho)$	19.104	18.363	9.696

Sparse network design

find most sparse subgraph with guaranteed convergence factor μ_0

- a hard combinatorial optimization problem; use ℓ_1 heuristic:

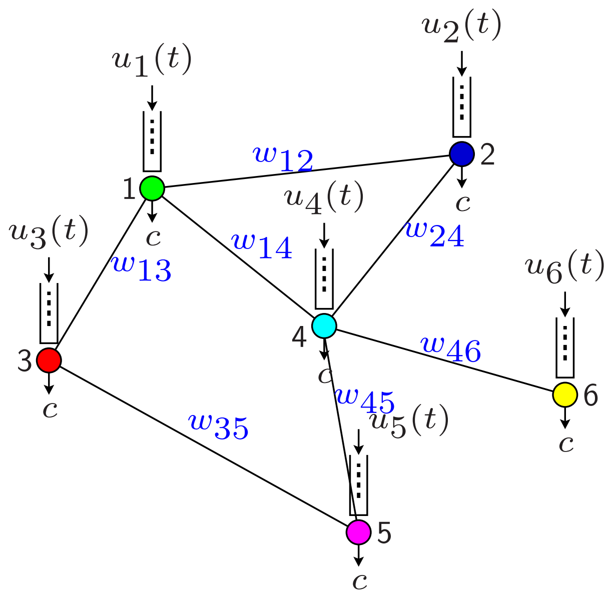
$$\begin{array}{ll} \text{minimize} & \sum_{\{i,j\}\in\mathcal{E}} |w_{ij}| \\ \text{subject to} & \mu(W) \leq \mu_0 \\ & W = W^T,\ W\mathbf{1} = \mathbf{1} \\ & w_{ij} = 0,\ \{i,j\} \notin \mathcal{E} \end{array}$$

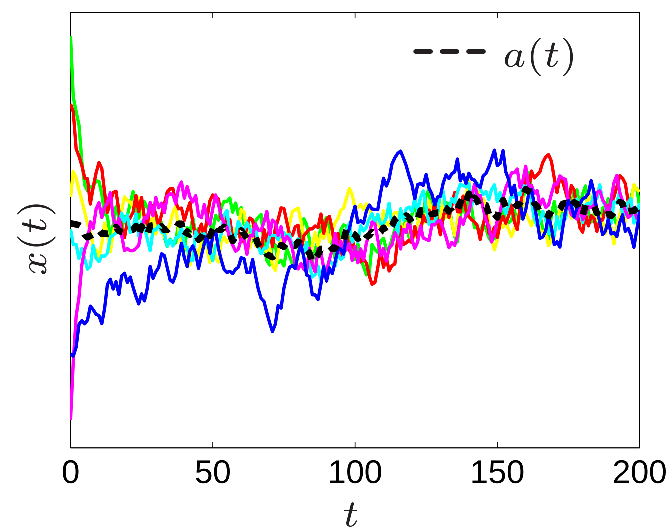


- best possible rate: $\mu^\star = 0.902$
  set design target: $\mu_0 = 0.910$

- total number of edges: 200
  used by sparse design: 96

Distributed consensus with additive noise





- example: distributed load balancing with new jobs arriving randomly

$$x_i(t+1) = x_i(t) + \sum_{j \in \mathcal{N}_i} w_{ij}\big(x_j(t) - x_i(t)\big) + v_i(t)$$

$v_i(t)$ i.i.d., zero mean & unit variance (consider $v_i(t) = u_i(t) - c$)

- instantaneous average: $a(t) = \dfrac{1}{n}\sum_{i=1}^{n} x_i(t)$, does random walk

Mean-squares deviation

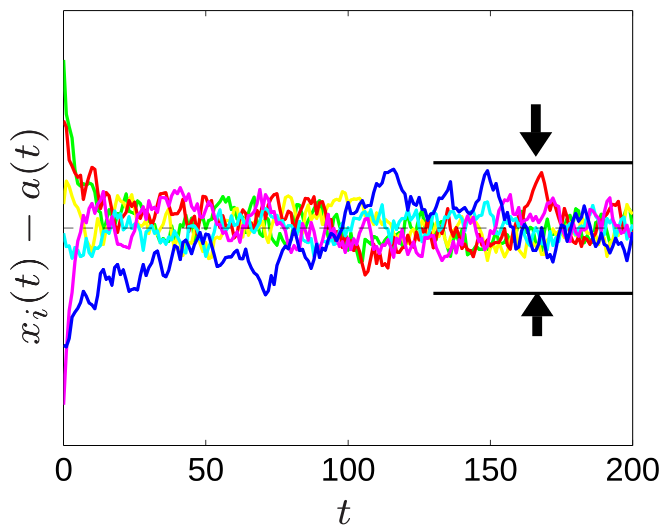
- $x(t+1) = Wx(t) + v(t)$

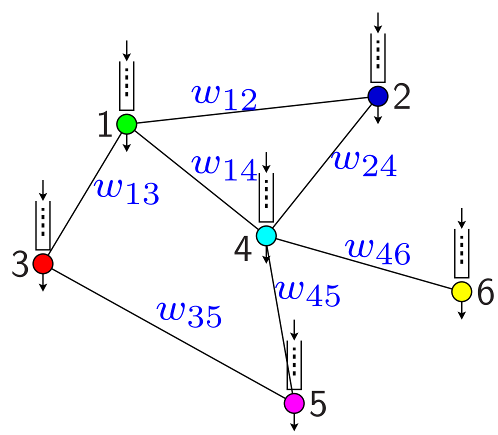
- consider **mean-square deviation**

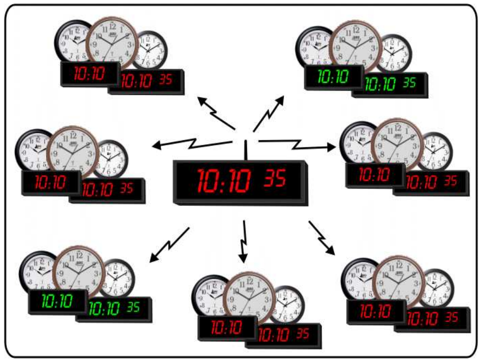
$$\mathsf{MSD}(W) = \lim_{t\to\infty} \mathbf{E} \sum_{i=1}^{n} (x_i(t) - a(t))^2$$

(Cybenko, 1989)

- optimize W to minimize $\mathsf{MSD}(W)$







Least-mean-square (LMS) distributed consensus

- **theorem** (XBK'06): explicit expression of $\mathrm{MSD}(W)$

$$\lim_{t\to\infty} \mathbf{E} \sum_{i=1}^{n} \left(x_i(t) - a(t)\right)^2 \;=\; \sum_{i=2}^{n} \frac{1}{1-\lambda_i(W)^2}$$

eigenvalues: $1 = \lambda_1(W) > \lambda_2(W) \geq \cdots \geq \lambda_n(W) \geq -1$

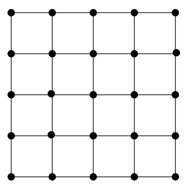
- optimal design for least-mean-square deviation

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{i=2}^{n} \frac{1}{1-\lambda_i(W)^2} \\ \text{subject to} & W = W^T, \quad W\mathbf{1} = \mathbf{1} \\ & W_{ij} = 0, \quad \{i,j\} \notin \mathcal{E} \end{array}$$

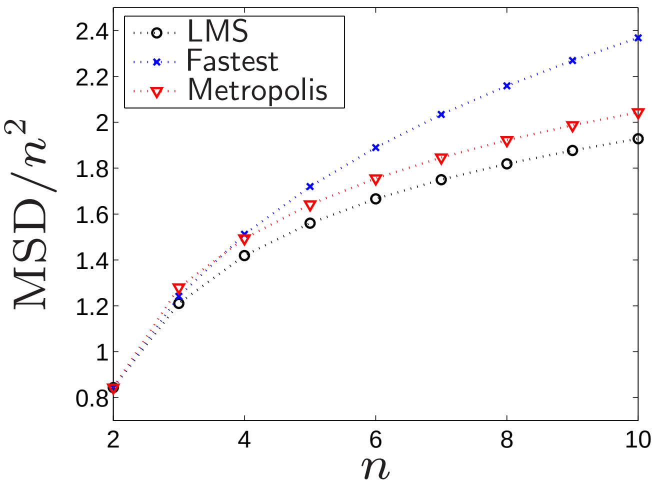
again, a **convex optimization** problem (convex, and differentiable)

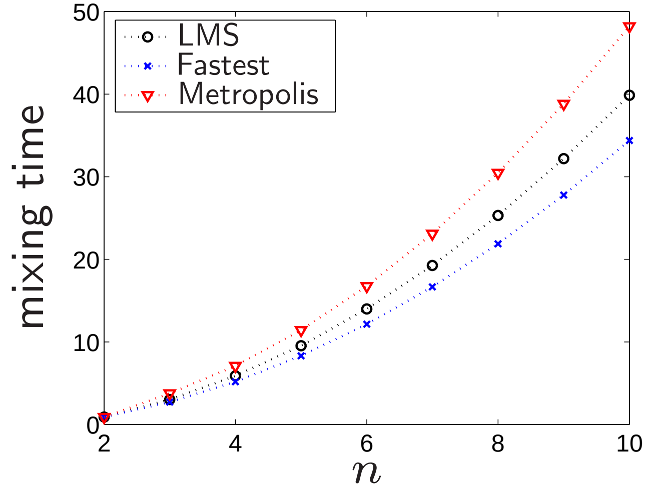
Examples of LMS load balancing

- **2-D grids**

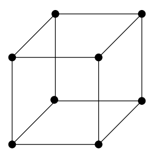


$(n = 5)$

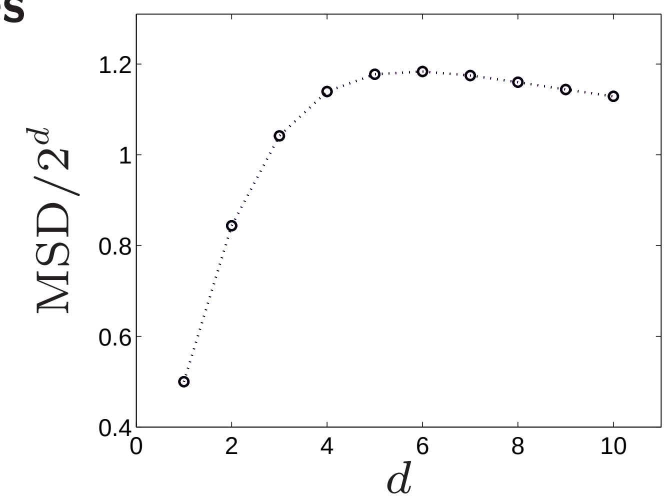


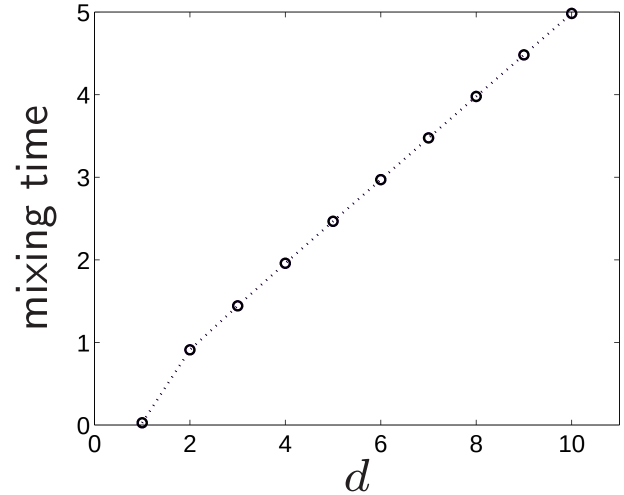


- **hypercubes**



$(d = 3)$





Outline

- convergence on fixed graph
- optimization on fixed graph
- **robustness to random link failures**
- optimization of randomized gossip

Distributed consensus with unreliable links

- communication links may work or fail at random (due to mobility, fading, power constraints); network topology changes with time

- some notations
  - $\mathcal{G}(t) = (\mathcal{E}(t), \mathcal{V})$ time-varying communication graph
  - $\mathcal{N}_i(t) = \{j \in \mathcal{V} \mid \{i,j\} \in \mathcal{E}(t)\}$ instantaneous neighborhood
  - $\{\mathcal{G}(t)\}_{t=0}^{\infty}$ can be either deterministic or stochastic

- distributed average consensus (same form, time-varying weights)

$$x_i(t+1) = x_i(t) + \sum_{j \in \mathcal{N}_i(t)} w_{ij}(t) \left(x_j(t) - x_i(t)\right)$$

what conditions on $\{\mathcal{G}(t), W(t)\}_{t=0}^{\infty}$ guarantee convergence?

Metropolis weights

- distributed average consensus

$$x_i(t+1) = x_i(t) + \sum_{j \in \mathcal{N}_i(t)} w_{ij}(t)\,(x_j(t) - x_i(t))$$

- time-varying Metropolis weights

$$w_{ij}(t) = \frac{1}{1+\max\{d_i(t), d_j(t)\}}, \qquad \{i,j\} \in \mathcal{E}(t)$$

$d_i(t) = |\mathcal{N}_i(t)|$, number of neighors at time t

- only use local information, suitable for distributed implementation

Robustness of convergence

- **theorem** (XBL:05): if the communication graphs that occur infinitely often in $\{\mathcal{G}(t)\}_{t=0}^{\infty}$ are *jointly connected*, then the iteration

$$x(t+1) = W(t)x(t)$$

converges to the average for any $x(0) \in \mathbf{R}^n$

- a finite set of graphs with common vertex set $\mathcal{G}_i = (\mathcal{E}_i, \mathcal{V})$, $i = 1, \ldots, r$, are **jointly connected** if their *union graph* $\mathcal{G} = (\cup_{i=1}^{r} \mathcal{E}_i, \mathcal{V})$ is connected

- **intuition:** convergence happens if graphs *"connected in a long run"*

Outline

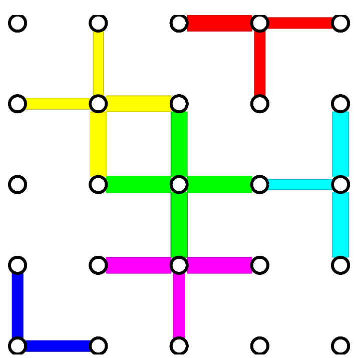
- convergence on fixed graph
- optimization on fixed graph
- robustness to random link failures
- **optimization of randomized gossip**

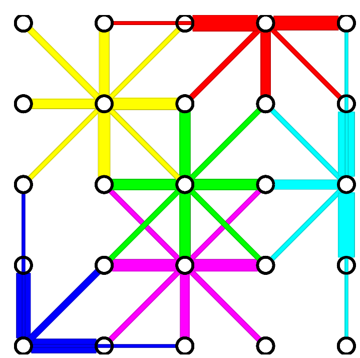
Example

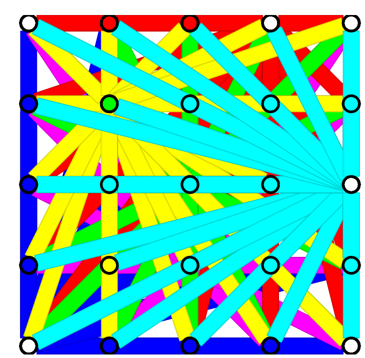
- randomized gossip on complete graph arranged as 5×5 grid

- communication costs

$$C_{ij} \propto d_{ij}^{\alpha}, \qquad d_{ij}\text{: Euclidean distance between node } i \text{ and } j$$







$\alpha = 2.1$ $\qquad$ $\alpha = 2.0$ $\qquad$ $\alpha = 1.9$

Summary

- distributed consensus in sensor networks
  - convergence
  - robustness
  - optimization
- right level of abstraction leads to interesting yet tractable problems
- mathematical tools used
  - linear algebra
  - spectral graph theory
  - convex optimization
  - duality