CS 455/655 --Mobile Sensor Networks
Lecture Notes 10: Consensus 1: Average Consensus

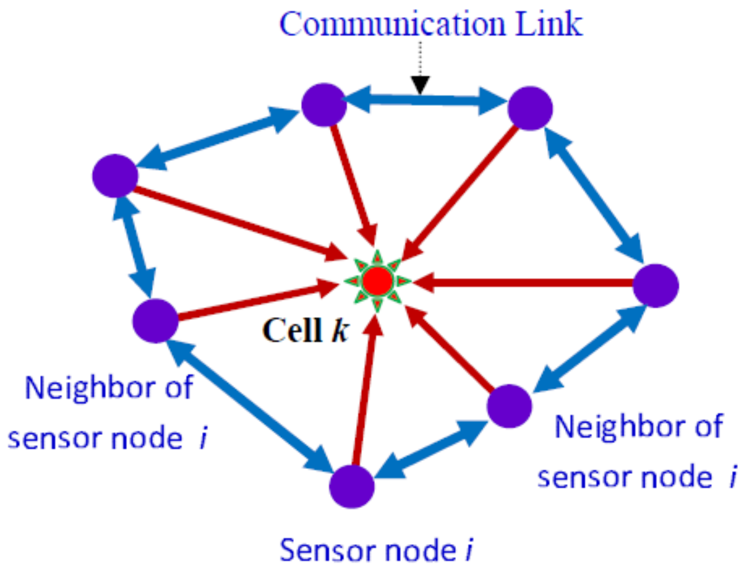
Hung La
University of Nevada Reno

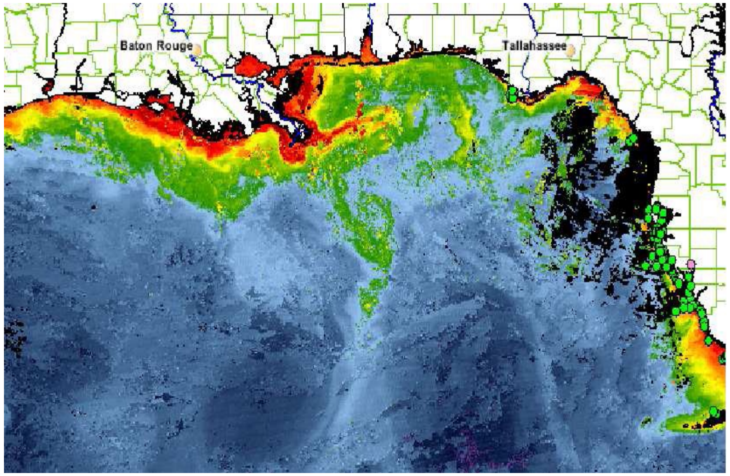
Consensus Meaning

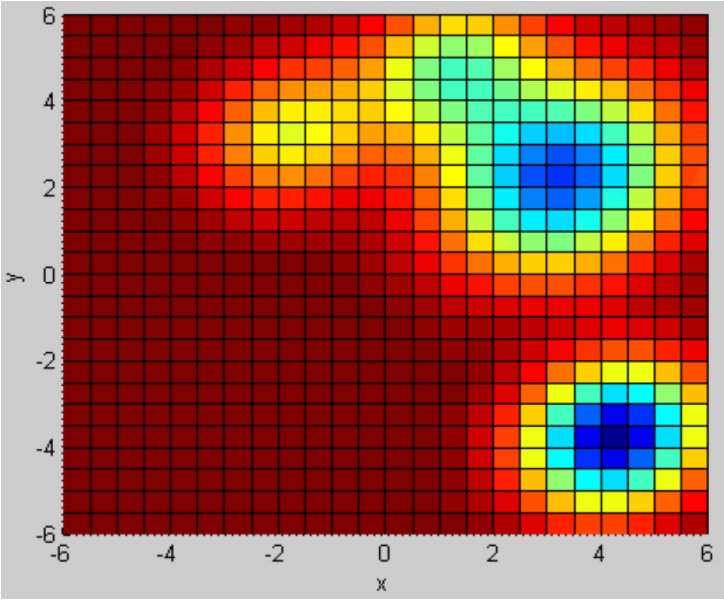
- How fish or birds can flock together
  - Swim/Fly with same/similar velocity
  - Maintain distance among each other (avoid collision)
  - How can they agree on the same velocity? Need consensus

Consensus Algorithms/Filters

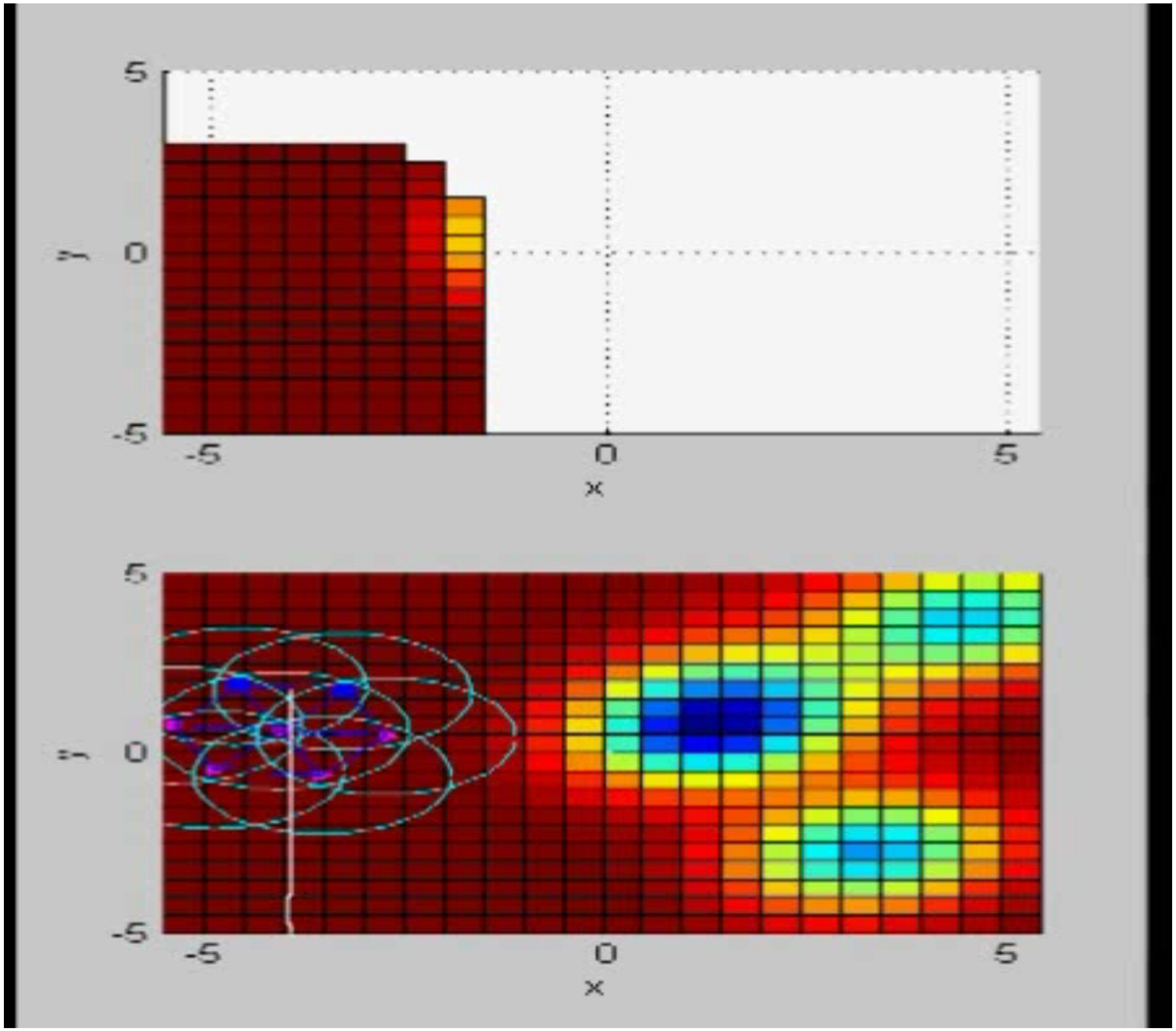
- Why need a consensus in MNSs
  - Different readings on each sensor
  - Agree on estimation





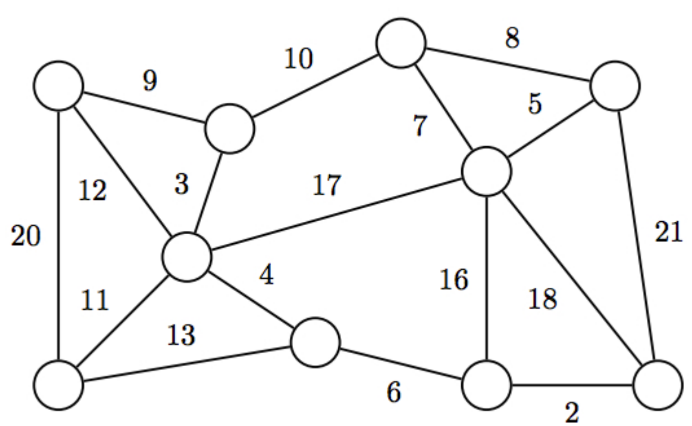


Consensus Application: Scalar field mapping

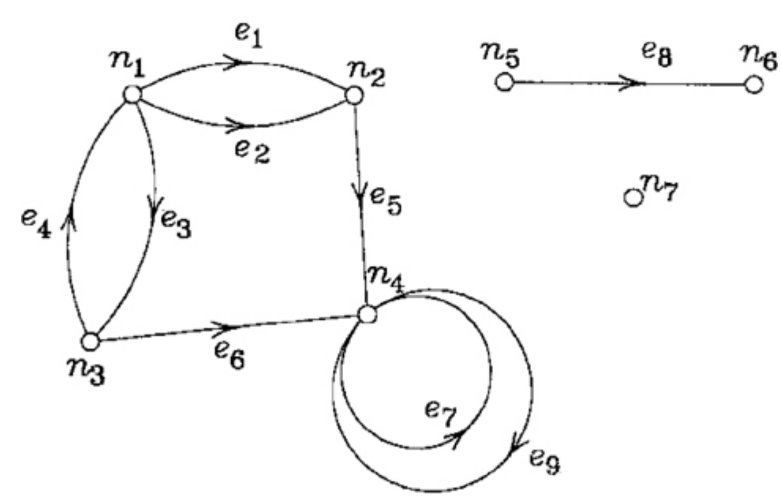


Graph Background (Fixed Graph)

We model the topology of a sensor network by an undirected graph — the communication graph. Let $\mathcal{G} = (\mathcal{E}, \mathcal{V})$ denote an undirected graph with vertex set $\mathcal{V} = \{1, 2, \ldots, n\}$ and edge set $\mathcal{E} \subset \{\{i, j\} \mid i, j \in \mathcal{V}\}$, where each edge $\{i, j\}$ is an unordered pair of distinct nodes. A graph is *connected* if for any two vertices i and j there exists a sequence of edges (a path) $\{i, k_1\}, \{k_1, k_2\}, \ldots, \{k_{s-1}, k_s\}, \{k_s, j\}$ in $\mathcal{E}$.



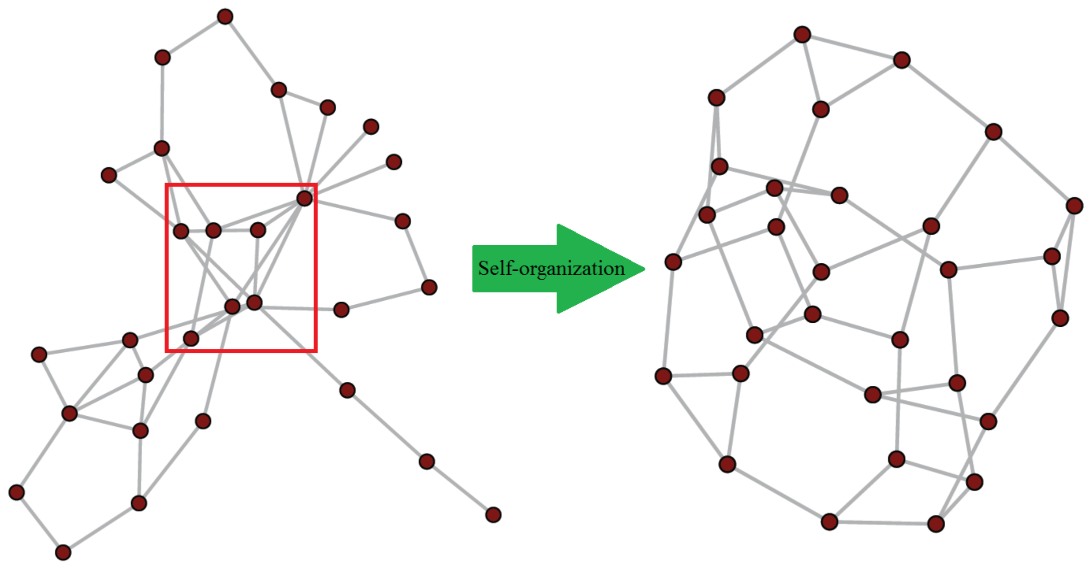
Undirected Graph



Directed Graph

Graph Background (Dynamic Graph)

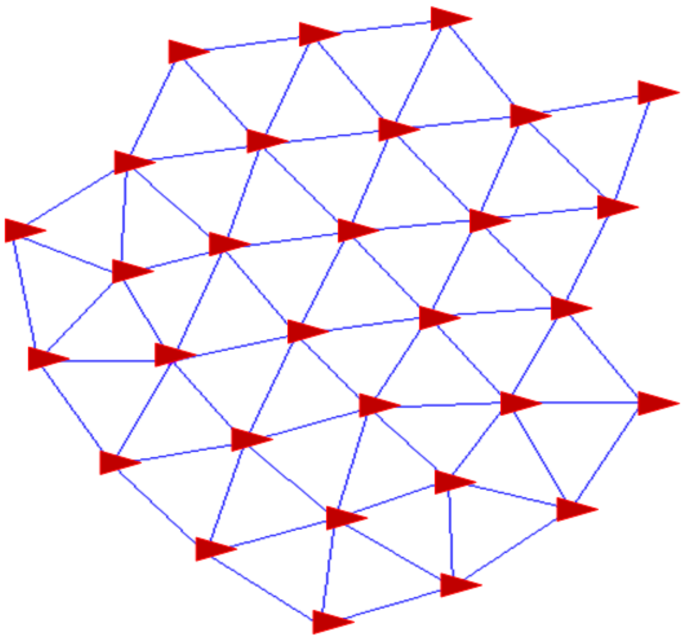
We represent the time-varying communication graph of a sensor network by $\mathcal{G}(t) = (\mathcal{E}(t), \mathcal{V})$, where $\mathcal{E}(t)$ is the set of active edges at time t. Let $\mathcal{N}_i(t) = \{j \in \mathcal{V} \mid \{i, j\} \in \mathcal{E}(t)\}$ denote the set of neighbors of node i at time t, and $d_i(t) = |\mathcal{N}_i(t)|$ denote the degree (number of neighbors) of node i at time t. In this paper, the sequence of communication graphs $\{\mathcal{G}(t)\}_{t=0}^{\infty}$ can be either deterministic or stochastic.



Neighborhood Definition

- q_i: the position of node i
- The set of neighbors of node *i* :

$$N_i(t) = \{j \in \vartheta : \|q_j - q_i\| \leq r, \vartheta = \{1, 2, \ldots, n\}, j \neq i\}$$



Example of α lattice graph

Background

A. Maximum-likelihood parameter estimation

We consider the estimation of a vector of unknown (but constant) parameters $\theta \in \mathbf{R}^m$ using a network of n distributed sensors. Each sensor makes a noisy vector measurement

$$y_i = A_i \theta + v_i, \qquad i = 1, \ldots, n,$$

where $y_i \in \mathbf{R}^{m_i}$, A_i is a known matrix that relates the unknown parameter to the ith sensor measurement, and v_i is noise modeled as a random variable. We assume v_i has zero mean and covariance matrix Σ_i, and that the noises v_i are independent. We will also assume that the sensor noises are (jointly) Gaussian.

The aggregate measurement of all sensors is

$$y = A\theta + v$$

Weighted Least-Squares

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix},$$

and the covariance matrix of v is $\Sigma = \mathbf{diag}(\Sigma_1, \ldots, \Sigma_n)$.

We assume that $m \le \sum_{i=1}^{n} m_i$ and the matrix A is full rank. The maximum-likelihood (ML) estimate of θ, given the measurements $y_1, \ldots, y_n$, is the weighted least-squares (WLS) approximate solution

$$\begin{aligned} \hat{\theta}_{\mathrm{ML}} &= \left(A^T \Sigma^{-1} A\right)^{-1} A^T \Sigma^{-1} y \\ &= \left(\sum_{i=1}^{n} A_i^T \Sigma_i^{-1} A_i\right)^{-1} \sum_{i=1}^{n} A_i^T \Sigma_i^{-1} y_i. \end{aligned} \tag{1}$$

This estimate is unbiased (*i.e.*, $\mathbf{E}\,\hat{\theta}_{\mathrm{ML}} = \theta$) and has error covariance matrix

$$Q = \mathbf{E}\left((\hat{\theta}_{\mathrm{ML}} - \theta)(\hat{\theta}_{\mathrm{ML}} - \theta)^T\right) = \left(A^T \Sigma^{-1} A\right)^{-1}. \tag{2}$$

Recall

Full Rank Matrices

When all of the vectors in a matrix are linearly independent, the matrix is said to be **full rank**. Consider the matrices **A** and **B** below.

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}$$

Notice that row 2 of matrix **A** is a scalar multiple of row 1; that is, row 2 is equal to twice row 1. Therefore, rows 1 and 2 are linearly dependent. Matrix **A** has only one linearly independent row, so its rank is 1. Hence, matrix **A** is not full rank.

Now, look at matrix **B**. All of its rows are linearly independent, so the rank of matrix **B** is 3. Matrix **B** is full rank.

Special Case

We first explain our method for the following special case:

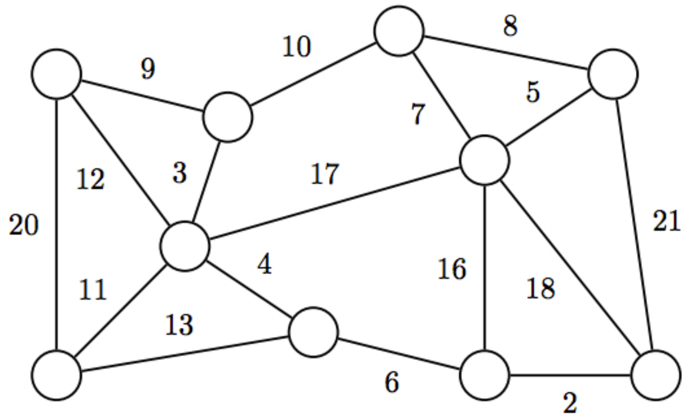
$$y_i = \theta + v_i, \qquad i = 1, \ldots, n,$$

where θ is a scalar to be estimated, and the noises are i.i.d. Gaussian: $v_i \sim \mathcal{N}(0, \sigma^2)$. In this case we have

$$\hat{\theta}_{\mathrm{ML}} = \frac{1}{n}\mathbf{1}^T y,$$

where $\mathbf{1}$ denotes the vector with all components one. In other words, the ML estimate is the average of the measurements y_i at all the sensors. The associated mean-square error is σ^2/n.

Consensus Algorithm



$$x_i(t+1) = W_{ii}(t)x_i(t) + \sum_{j \in \mathcal{N}_i(t)} W_{ij}(t)x_j(t), \quad i = 1, \ldots, n. \tag{3}$$

Here $W_{ij}(t)$ is the linear *weight* on $x_j(t)$ at node i. Setting $W_{ij}(t) = 0$ for $j \notin \mathcal{N}_i(t)$, the above distributed iterative process can be written in vector form as

$$x(t+1) = W(t)x(t), \tag{4}$$

Average Consensus

with the initial condition $x(0) = y$. Note that the weight matrix $W(t) \in \mathbf{R}^{n\times n}$ has the sparsity pattern specified by the communication graph $\mathcal{G}(t)$.

With the definition of a t-step transition matrix

$$\Phi(t) = W(t-1)\cdots W(1)W(0),$$

we have

$$x(t) = \Phi(t)x(0).$$

We would like to choose the weight matrices $W(t)$ such that the states at all the nodes converge to $\hat{\theta}_{\mathrm{ML}} = (1/n)\mathbf{1}^T y$, *i.e.*,

$$\lim_{t\to\infty} x(t) = \hat{\theta}_{\mathrm{ML}}\mathbf{1} = \left(\frac{1}{n}\mathbf{1}^T x(0)\right)\mathbf{1}.$$

Weight Design

Since this should hold for any $x(0) \in \mathbf{R}^n$, it is equivalent to

$$\lim_{t \to \infty} \Phi(t) = \frac{1}{n} \mathbf{1}\mathbf{1}^T. \tag{5}$$

A. Choice of weights

Although we are interested here in time-varying graphs, it is interesting to first consider the case when the communication graph is fixed, and we use a time-invariant weight matrix W. In this case, the following conditions are necessary and sufficient for (5) (in this case $\Phi(t) = W^t$)

$$\mathbf{1}^T W = \mathbf{1}^T, \quad W\mathbf{1} = \mathbf{1}, \quad \rho(W - \mathbf{1}\mathbf{1}^T/n) < 1, \tag{6}$$

where $\rho(\cdot)$ denote the spectral radius of a matrix; see [8], [9]. The asymptotic convergence rate of the average consensus, defined as

$$r_{\text{asym}}(W) = \sup_{x(0) \neq \theta_{\text{ML}} \mathbf{1}} \lim_{t \to \infty} \left(\frac{\|x(t) - \theta_{\text{ML}} \mathbf{1}\|}{\|x(0) - \theta_{\text{ML}} \mathbf{1}\|} \right)^{1/t}, \tag{7}$$

Weight Design

- **Maximum-degree weights.** Here we use the constant weight $1/n$ on all the edges, and choose the self-weights so that the sum of weights at each node is 1:

$$W_{ij}(t) = \begin{cases} \dfrac{1}{n} & \text{if } \{i,j\} \in \mathcal{E}(t), \\ 1 - \dfrac{d_i(t)}{n} & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

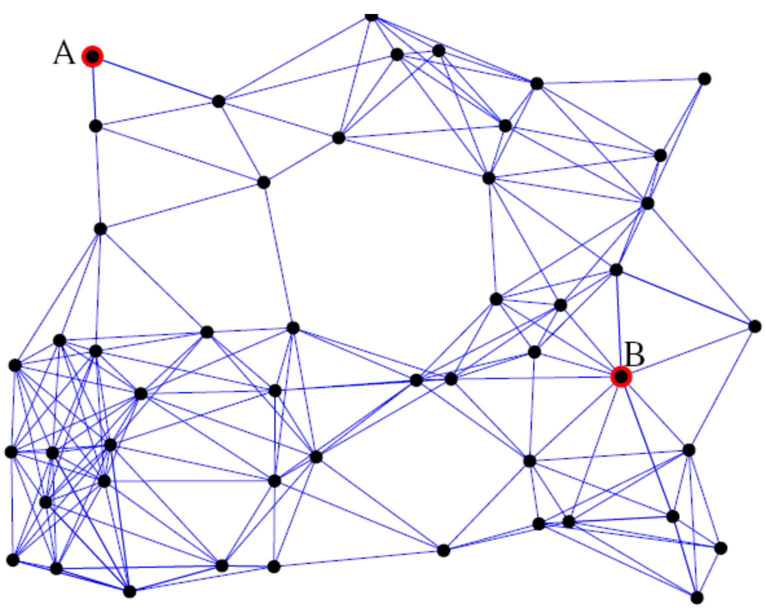
- **Metropolis weights.** The Metropolis weight matrix is defined as

$$W_{ij}(t) = \begin{cases} \dfrac{1}{1+\max\{d_i(t), d_j(t)\}} & \text{if } \{i,j\} \in \mathcal{E}(t), \\ 1 - \displaystyle\sum_{\{i,k\} \in \mathcal{E}(t)} W_{ik}(t) & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

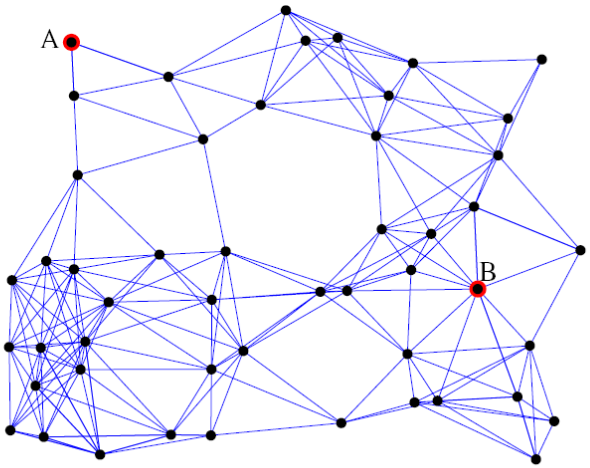
Results

50 Sensor nodes
Distributed on the unit square [0,1]x[0,1]
Neighbor nodes: if distance < ¼ unit

The vector of unknown parameters, θ, has dimension $m = 5$. Each sensor takes a scalar measurement $y_i = a_i^T \theta + v_i$, where the vectors a_i were chosen from a uniform distribution on the unit sphere in $\mathbf{R}^5$, and the noises are i.i.d. Gaussian with unit variance: $v_i \sim \mathcal{N}(0, 1)$.



Results

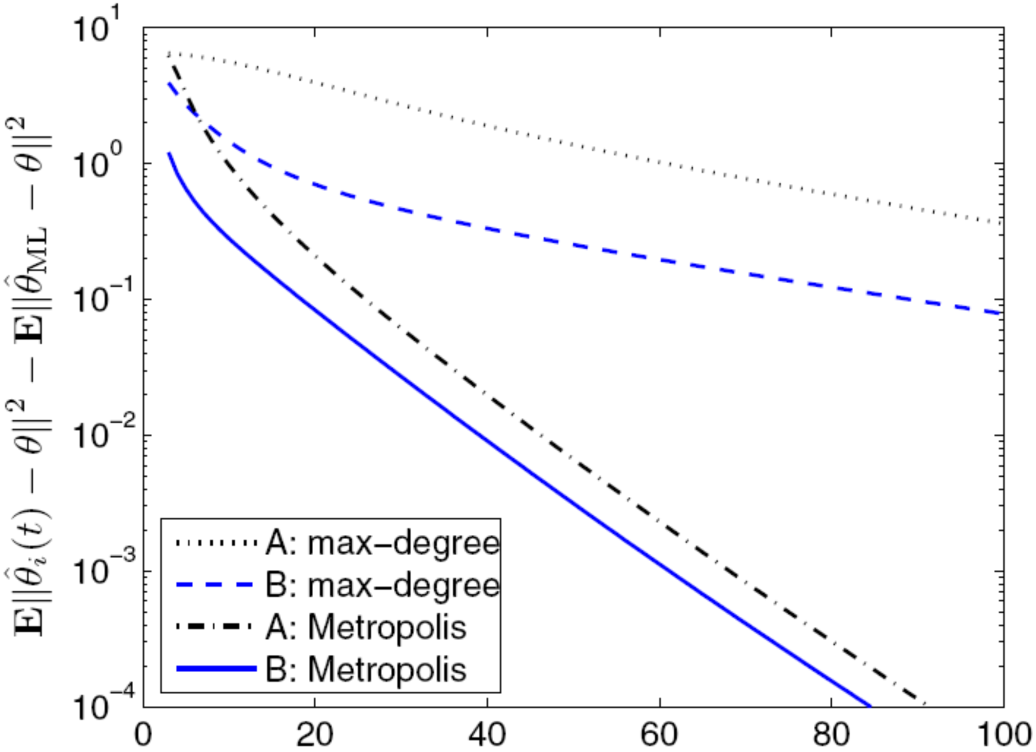


Max-Degree

$$W_{ij}(t) = \begin{cases} \frac{1}{n} & \text{if } \{i,j\} \in \mathcal{E}(t), \\ 1 - \frac{d_i(t)}{n} & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Metropolis

$$W_{ij}(t) = \begin{cases} \frac{1}{1+\max\{d_i(t), d_j(t)\}} & \text{if } \{i,j\} \in \mathcal{E}(t) \\ 1 - \sum_{\{i,k\} \in \mathcal{E}(t)} W_{ik}(t) & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$



Iteration Number

Mean Square Error (MSE) at nodes A and B: fixed communication graph.

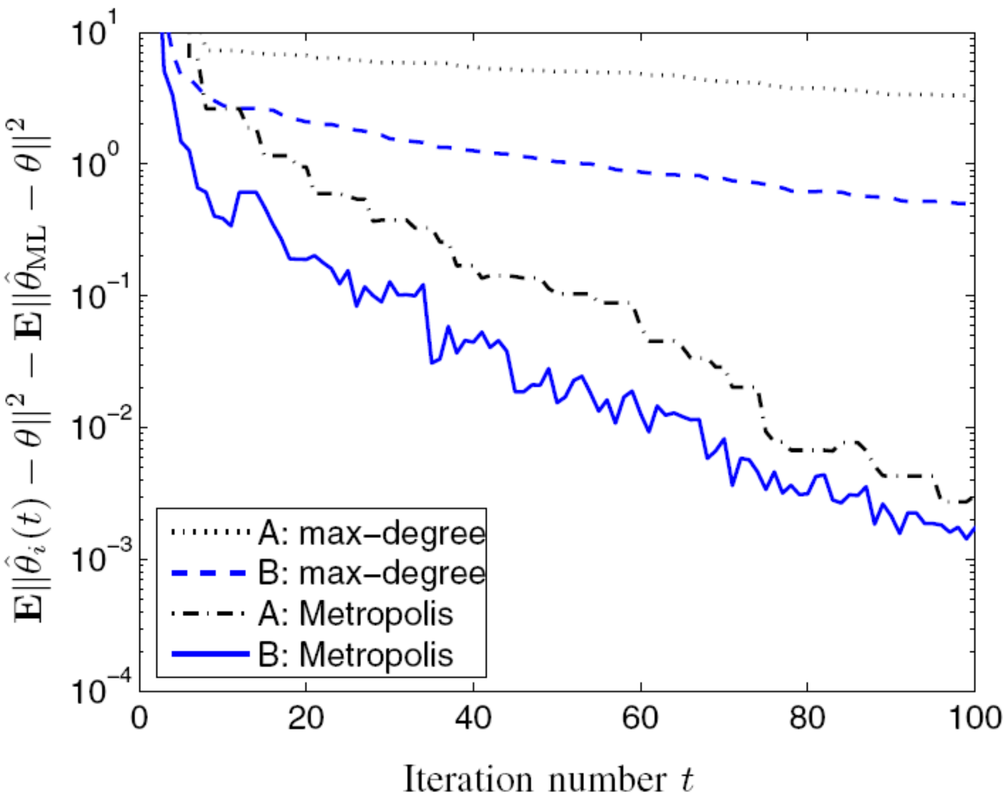
Results

Max-Degree

$$W_{ij}(t) = \begin{cases} \frac{1}{n} & \text{if } \{i,j\} \in \mathcal{E}(t), \\ 1 - \frac{d_i(t)}{n} & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Metropolis

$$W_{ij}(t) = \begin{cases} \frac{1}{1+\max\{d_i(t), d_j(t)\}} & \text{if } \{i,j\} \in \mathcal{E}(t) \\ 1 - \sum_{\{i,k\} \in \mathcal{E}(t)} W_{ik}(t) & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$



Iteration Number

MSE at nodes A and B: dynamically changing communication graph.

Reference

A Scheme for Robust Distributed Sensor Fusion Based on Average Consensus

Lin Xiao
Center for the Mathematics of Information
California Institute of Technology
Pasadena, CA 91125-9300
Email: lxiao@caltech.edu
Stephen Boyd
Department of Electrical Engineering
Stanford University
Stanford, CA 94305-9510
Email: boyd@stanford.edu
Sanjay Lall
Department of Aeronautics and
Astronautics
Stanford University
Stanford, CA 94305-4035
Email: lall@stanford.edu