Binary Consensus in Wireless Sensor Networks Using Distributed Cophasing

Harish V. Venugopalakrishna Y. R. Chandra R. Murthy

Indian Institute of Science, Bangalore

January 26, 2013

Outline

- Introduction to Consensus and Distributed Cophasing
- System Model & Problem Statement
- Processing at Nodes
- Performance Analysis
- Simulation Results
- Conclusions & Future Work

Introduction to Consensus

- Consensus: A number of nodes coming to an agreement with each other
- **Motivation:**
  - The cognitive radio system
    Nodes: cognitive users
    Desired Value: presence of primary
- Very important in cooperative control problems
- **Classifications of Consensus:**
  - Distributed vs. centralized
  - Average, majority,. . .
  - Detection vs. estimation
  - Physical vs. higher layers
  - Static vs. dynamic

Introduction to Consensus

- Consensus: A number of nodes coming to an agreement with each other
- **Motivation:**
  - The cognitive radio system
    Nodes: cognitive users
    Desired Value: presence of primary
- Very important in cooperative control problems
- **Classifications of Consensus:**
  - Distributed vs. centralized
  - Average, majority,. . .
  - Detection vs. estimation
  - Physical vs. higher layers
  - Static vs. dynamic

Introduction to Consensus

- Consensus: A number of nodes coming to an agreement with each other
- **Motivation:**
  - The cognitive radio system
    Nodes: cognitive users
    Desired Value: presence of primary
- Very important in cooperative control problems
- **Classifications of Consensus:**
  - Distributed vs. centralized
  - Average, majority, . . .
  - Detection vs. estimation
  - Physical vs. higher layers
  - Static vs. dynamic

Introduction to Consensus

- Consensus: A number of nodes coming to an agreement with each other
- **Motivation:**
  - The cognitive radio system
    Nodes: cognitive users
    Desired Value: presence of primary
- Very important in cooperative control problems
- **Classifications of Consensus:**
  - Distributed vs. centralized
  - Average, majority,...
  - Detection vs. estimation
  - Physical vs. higher layers
  - Static vs. dynamic

Introduction to Consensus

- Consensus: A number of nodes coming to an agreement with each other
- **Motivation:**
  - The cognitive radio system
    Nodes: cognitive users
    Desired Value: presence of primary
- Very important in cooperative control problems
- **Classifications of Consensus:**
  - Distributed vs. centralized
  - Average, majority,. . .
  - Detection vs. estimation
  - Physical vs. higher layers
  - Static vs. dynamic

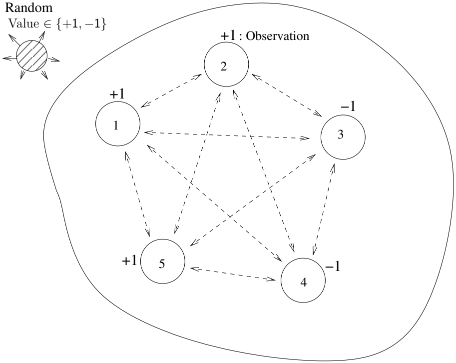
Introduction to Consensus

- Consensus: A number of nodes coming to an agreement with each other
- **Motivation:**
  - The cognitive radio system
    Nodes: cognitive users
    Desired Value: presence of primary
- Very important in cooperative control problems
- **Classifications of Consensus:**
  - Distributed vs. centralized
  - Average, majority,...
  - Detection vs. estimation
  - Physical vs. higher layers
  - Static vs. dynamic

How to Achieve Consensus?

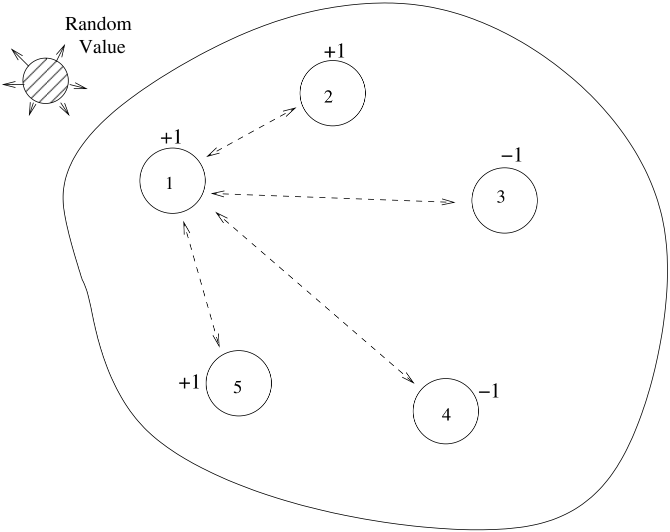
- Need to exchange information: Transmission scheme
- Transmission scheme affects the performance
- Examples: Point-to-point, broadcast, multiple access, distributed cophasing etc.
- Impact of transmission scheme is not well studied
- Typically consensus problems assume error-free links
- Consensus under noisy communication is not well studied

Problem Statement

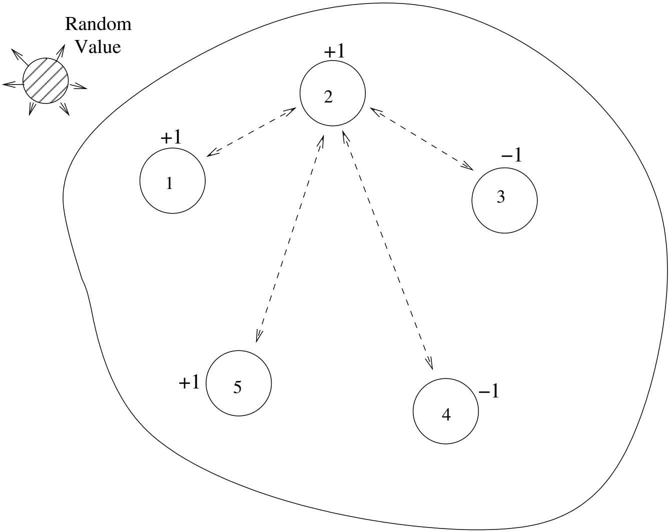


- Each node has an estimate of the binary random variable
- Nodes are allowed to exchange information & update in a fully connected network topology, till consensus is reached

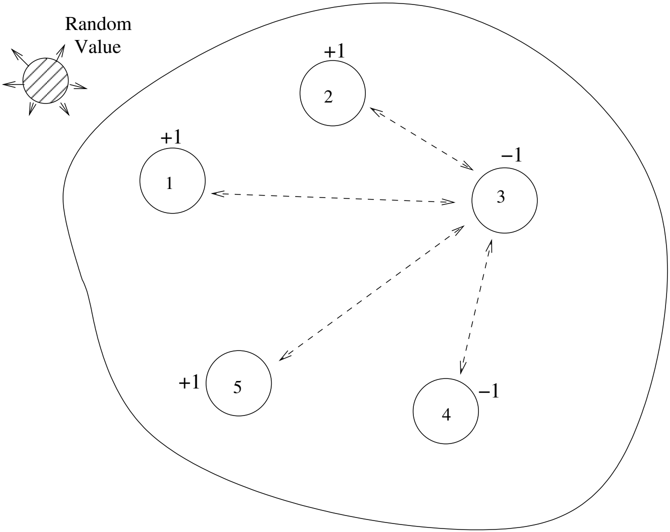
Problem Statement



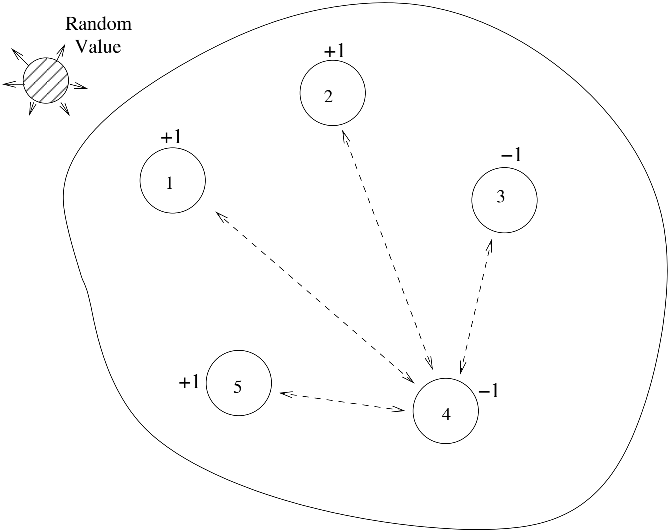
Problem Statement



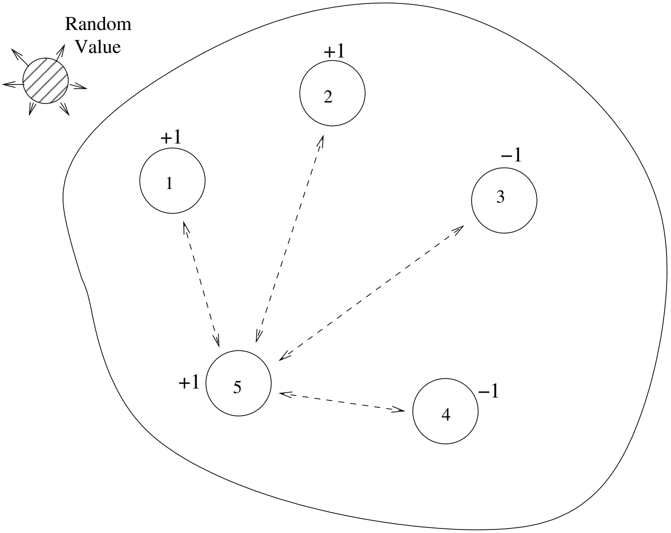
Problem Statement



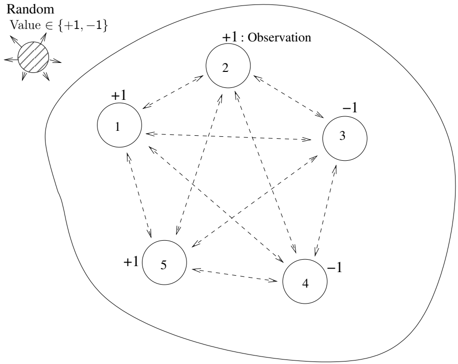
Problem Statement



Problem Statement



Problem Statement



Q: Will it reach consensus? If so, how long does it take?
Is it better than earlier schemes?

The Tx Scheme of Distributed Cophasing (DCP)

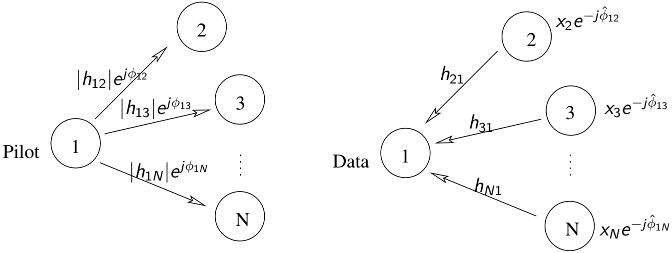


Figure: A DCP Session

- Pilot assisted transmission, no power control
- Nodes intend to transmit such that their signals coherently add at the fusion center
- Channels are assumed to be reciprocal, i.i.d. and Rayleigh faded

The Tx Scheme of Distributed Cophasing (DCP)

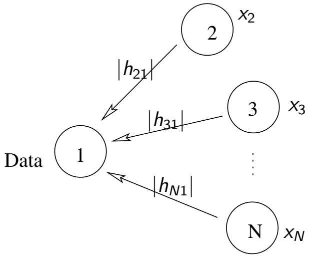
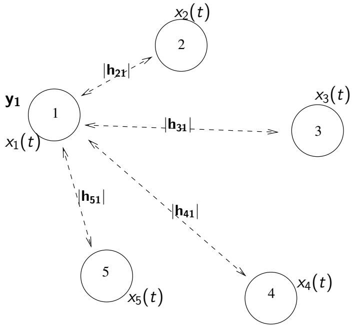


Figure: A DCP Session

- Pilot assisted transmission, no power control
- Nodes intend to transmit such that their signals coherently add at the fusion center
- Channels are assumed to be reciprocal, i.i.d. and Rayleigh faded

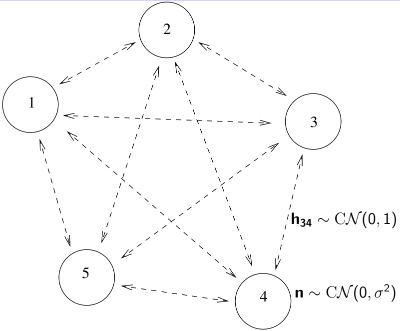
System Model



- Assume perfect channel estimation at nodes
- After a DCP step, a node has a pair of values x_i, y_i where y_i is the received DCP symbol given by:

$$y_i(t) = \sum_{j \neq i} |h_{ji}| x_j(t) + n\,, \quad n \sim \mathcal{CN}(0, \sigma^2)$$

System Model



- Information at node i: $x_i(t) \in \{+1, -1\}$
- The set of all nodes: $\mathcal{N} \triangleq \{1, 2 \ldots, N\}$
- The channel matrix: $H = [h_{ij}]$
- State-vector of binary values: $\mathcal{D}(t) = [x_1(t), x_2(t), \ldots, x_N(t)]$
- The set of all possible 2^N states: Φ
- The subset of Φ where majority is $+1$: Φ_1

Node Update Rules for DCP

- At each node:
  1. Available Data: Own observation, received DCP symbol and channel gains $(x_i(t), y_i(t), \{h_{ji}, j \neq i\})$
  2. To estimate: Majority bit across the nodes
  3. Question: *What is the best estimate of majority?*

- We propose two techniques for estimation:
  1. Maximum Likelihood (ML) based estimation
  2. Low complexity Linear Minimum Mean Squared Error (LMMSE) based estimation

ML Based Update Rule

- Then the ML estimate can be written as:

$$x_i^{ML}(t+1) = \begin{cases} +1, & \text{if } \Theta^{(i)} \geq 0.5 \\ -1, & \text{else} \end{cases}$$

where $\Theta^{(i)}$ is defined as the probability of $+1$ majority

Contd.

ML Rule (Contd.)

- By Bayes' rule:

$$
\begin{aligned}
\Theta^{(i)} &\triangleq \Pr\Big(\text{n/w state has majority} + 1 \mid \text{available data}\Big) \\
&= \Pr\Big\{\mathcal{D}(t) \in \Phi_1 \mid (x_i, y_i), H(t)\Big\} \\
&= \frac{\sum_{\phi \in \Phi_1} \Pr\Big\{(x_i, y_i) \mid \mathcal{D}(t) = \phi, H(t)\Big\}}{\sum_{\phi \in \Phi} \Pr\Big\{(x_i, y_i) \mid \mathcal{D}(t) = \phi, H(t)\Big\}}
\end{aligned}
$$

Notation:

$\mathcal{D}(t) = [x_1 \;\; x_2 \ldots x_N]$ — the network state at time t

Φ — the set of all possible states

Φ_1 — the set of states where majority is $+1$

LMMSE Based Update Rule

- Uses LMMSE estimate of the sum $\sum_j x_j$ based on the data (y_i, x_i)
- Less complex and much easier to implement
- The estimate is given by:

$$x_i^{LMMSE}(t+1) = \text{sign}(\hat{\mathbf{s}}_\mathbf{i})$$

where $\hat{\mathbf{s}}_\mathbf{i}$ is the LMMSE estimate of the sum $\mathbf{s} \triangleq \sum_j x_j$ at node i, at time t

Contd.

LMMSE Rule (Contd.)

- Let $\hat{\mathbf{s}}_{\mathbf{i}}'$ denote the estimate of $\mathbf{s_i}' \triangleq \sum_{j \neq i} x_j$
- $\mathbf{s_i}'$ is a function of the DCP symbol y_i only. Therefore, a linear estimate $\hat{\mathbf{s}}_{\mathbf{i}}'$ of $\mathbf{s_i}'$ suffices for the desired $\hat{\mathbf{s}}_{\mathbf{i}}$

$$\hat{\mathbf{s}}_{\mathbf{i}} \triangleq (\hat{\mathbf{s}}_{\mathbf{i}}' + x_i),$$

$$\hat{\mathbf{s}}_{\mathbf{i}}' = \alpha_i^* y_i + \beta_i^*,$$

$$\text{where } \alpha_i^* = \frac{\sum_{j \neq i} |h_{ji}|}{\sum_{j \neq i} |h_{ji}|^2 + \sigma^2}, \quad \beta_i^* = 0, \quad \forall i \in \mathcal{N}$$

Performance Analysis

- The state transition depends solely on earlier state and the channel state (channel gain matrix)
- Channel gains vary at each cycle

  $\implies$ Transition probability matrix (TPM) varies with time

  $\implies$ Its a time-varying Markov chain!
- Can study the average statistics of this dynamic system

Contd.

Performance Analysis (Contd.)

- Performance metrics for consensus:

  1. Probability of *accurate* consensus
  2. Speed of convergence

Probability of Accurate Consensus

- It can be seen that the average TPM P has all positive elements

$$P = [p_{ij}], p_{ij} > 0 \; \forall i, j$$

(Every state is attainable from an arbitrary state)

Probability of Accurate Consensus

- It can be seen that the average TPM P has all positive elements
$$P = [p_{ij}], p_{ij} > 0 \ \forall i, j$$
(Every state is attainable from an arbitrary state)

- From Perron's theorem, the stationary prob. distribution exists & the same is attained for any initial distribution

Probability of Accurate Consensus

- It can be seen that the average TPM P has all positive elements

  $$P = [p_{ij}], p_{ij} > 0 \; \forall i, j$$

  (Every state is attainable from an arbitrary state)

- From Perron's theorem, the stationary prob. distribution exists & the same is attained for any initial distribution

  $\implies$ the final state is independent of the initial state!

Probability of Accurate Consensus

- It can be seen that the average TPM P has all positive elements
$$P = [p_{ij}], p_{ij} > 0 \ \forall i, j$$
(Every state is attainable from an arbitrary state)

- From Perron's theorem, the stationary prob. distribution exists & the same is attained for any initial distribution

$\implies$ the final state is independent of the initial state!

- However, in the transient stage, we have observed through simulations that the probability of accurate consensus increases monotonically and approaches 1

Probability of Accurate Consensus

- It can be seen that the average TPM P has all positive elements
$$P = [p_{ij}], p_{ij} > 0 \ \forall i, j$$
(Every state is attainable from an arbitrary state)

- From Perron's theorem, the stationary prob. distribution exists & the same is attained for any initial distribution

$\implies$ the final state is independent of the initial state!

- However, in the transient stage, we have observed through simulations that the probability of accurate consensus increases monotonically and approaches 1

$\implies$ In finite number of cycles, accurate consensus can in fact be achieved with very high probability

Speed of Convergence Indicator

- Convergence (on an average) under consideration is:

$$\mathbb{E}[\pi_\infty] = \lim_{n\to\infty} \pi_0 P^n$$

- The convergence of a matrix like P^n can be seen in its *diagonalized* form:

$$P^n = S\Lambda^n S^{-1} = \sum_{i=1}^{n} \lambda_i^n u_i {v_i}^T,$$

where we denote the matrix S formed by eigenvectors $\{u_i, i = 1, 2 \dots N\}$, matrix Λ formed by eigenvalues as:

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_N \end{bmatrix}, \quad S = [\underline{u_1}\ \underline{u_2} \dots \ \underline{u_N}] \text{ and } S^{-1} = \begin{bmatrix} \underline{v_1}^T \\ \underline{v_2}^T \\ \vdots \\ \underline{v_N}^T \end{bmatrix}$$

$$P^n = S\Lambda^n S^{-1} = \sum_{i=1}^{n} \lambda_i^n u_i v_i^T,$$

- If $(|\lambda_1| = 1) \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$, as $n \to \infty$, P^n is dominated more and more by the term with λ_2^n
- We can take $|\lambda_2|$ as a measure of convergence rate
- The closer $|\lambda_2|$ is to one, slower the speed of convergence to the memoryless state and longer the system depends on the initial state
- The proof extends in a similar way to non-diagonalizable P matrices, using Jordan Canonical form

The Second Eigenvalue Computation (Approx.)

- Closed-form expression for λ_2 is difficult in general
- We need an approximation for the second eigenvalue λ_2
- An approximation to λ_2 is:

$$\lambda_2 \approx 1 - 2\overline{\gamma}^{(i)}_{\text{all-zero}},$$

where $\overline{\gamma}^{(i)}_{\text{all-zero}}$ is the average error probability at node i in all-zero state, i.e.,

$$\overline{\gamma}^{(i)}_{\text{all-zero}} \triangleq \mathbb{E}_H \left[\Pr \left\{ x_i(t+1) = +1 |_{\text{all-zero state}, H(t)} \right\} \right]$$

Simulation Setup

- Number of nodes is $N = 8$. The TPM of the Markov chain is a 256×256 matrix
- Averaged over $10,000$ channel instantiations to generate TPM
- Channel to noise ratio:

$$CNR \triangleq \frac{\mathbb{E}\left[|h|^2\right]}{\sigma^2}$$

- Results:
  1. Performance of LMMSE update rule
  2. Comparison of the performance of DCP algorithm with an existing scheme called Basic Affine Estimation (BAE)
  3. Verifying the second eigenvalue approximation

Simulations

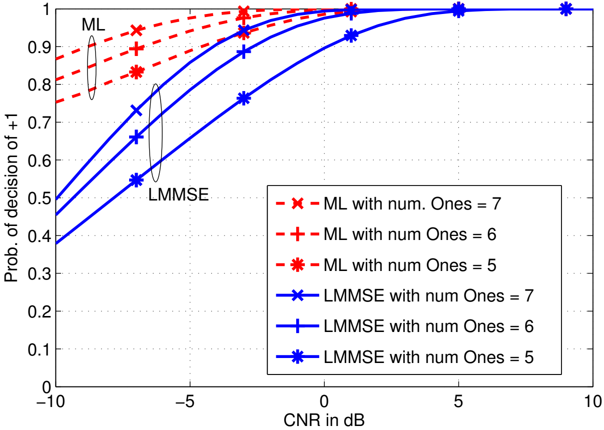


Figure: LMMSE vs ML for different *initial* majorities in a network of 8 nodes, in one cycle

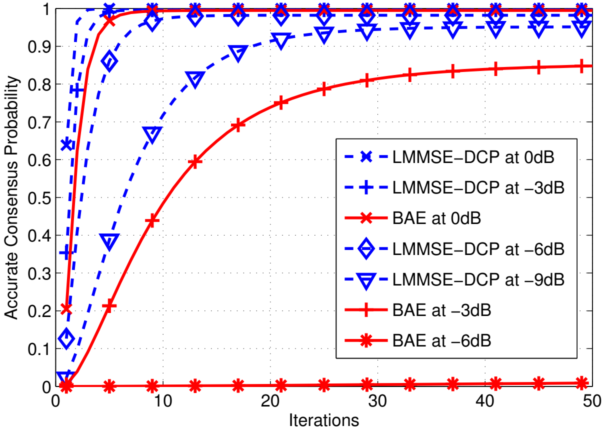


Figure: DCP LMMSE vs BAE[1] algorithm, when 6 out of 8 nodes initially vote $+1$

[1] Mostofi Y. and Malmirchegini M., "Binary Consensus Over Fading Channels", *IEEE Trans. Signal Proc.*, vol.58, no.12, Dec. 2010.

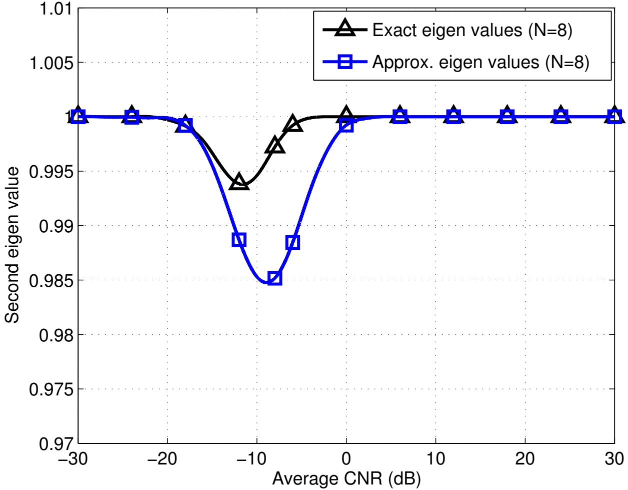


Figure: Second eigen value: approx. vs actual at various CNRs when N=8

Conclusions

- We proposed a feasible model for achieving improved performance of physical layer binary consensus in fading environment
- We have proposed a low complexity linear update rule at nodes which performs comparable to the ML rule
- Significantly better performance over existing consensus algorithms

Future Work

- The explicit node scheduling difficulty in distributed setup — The "Randomized Wake Policy" or "Pull" model
- A node randomly wakes up and updates itself after DCP protocol
- Simple & attractive in practical implementation
- Simulations suggest that its performance is on par with the case where nodes are precisely scheduled!
- Current Challenges:
  1. Theoretical analysis of convergence
  2. Second eigenvalue computation to characterize the convergence behavior

Thank You!