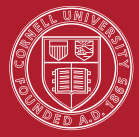


Cornell University

Lecture 11 - Network dynamics: the Random Walk and Consensus

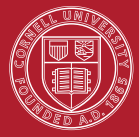
Examples of network dynamics
ECE 5260 – ORIE 5735

Anna Scaglione

Cornell Tech

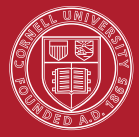
March 2, 2026

Content



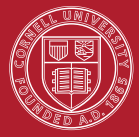
Cornell University

Outline



Cornell University

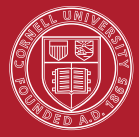
Review and overview



Cornell University

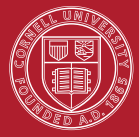
- In the first module we introduced:
  - Basic notions from graph theory, including the algebraic definition of three key matrices:
    1. The adjacency matrix
    2. The incidence matrix
    3. The laplacian matrix
  - We discussed how these matrices can help understand the structure of the network

Review and overview



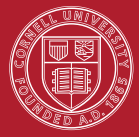
- In the first module we introduced:
  - Basic notions from graph theory, including the algebraic definition of three key matrices:
    1. The adjacency matrix
    2. The incidence matrix
    3. The laplacian matrix
  - We discussed how these matrices can help understand the structure of the network

Review and overview



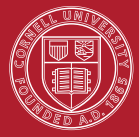
- In the first module we introduced:
  - Basic notions from graph theory, including the algebraic definition of three key matrices:
    1. The adjacency matrix
    2. The incidence matrix
    3. The laplacian matrix
  - We discussed how these matrices can help understand the structure of the network

Review and overview



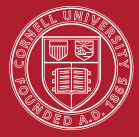
- In the first module we introduced:
  - Basic notions from graph theory, including the algebraic definition of three key matrices:
    1. The adjacency matrix
    2. The incidence matrix
    3. The laplacian matrix
  - We discussed how these matrices can help understand the structure of the network

Review and overview



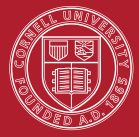
- Then we focused on:
  - metrics that help understand the network structure and provide an embedding of the network for comparisons
  - models for random graphs to gain insight on the similarities and differences of real networks from these abstract models
- Today we introduce some basic models of network dynamics that are used to model a variety of networked systems.

Review and overview



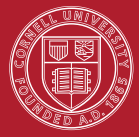
- Then we focused on:
  - metrics that help understand the network structure and provide an embedding of the network for comparisons
  - models for random graphs to gain insight on the similarities and differences of real networks from these abstract models
- Today we introduce some basic models of network dynamics that are used to model a variety of networked systems.

Review and overview



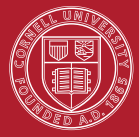
- Then we focused on:
  - metrics that help understand the network structure and provide an embedding of the network for comparisons
  - models for random graphs to gain insight on the similarities and differences of real networks from these abstract models
- Today we introduce some basic models of network dynamics that are used to model a variety of networked systems.

Review and overview



- Then we focused on:
  - metrics that help understand the network structure and provide an embedding of the network for comparisons
  - models for random graphs to gain insight on the similarities and differences of real networks from these abstract models
- Today we introduce some basic models of network dynamics that are used to model a variety of networked systems.

Outline



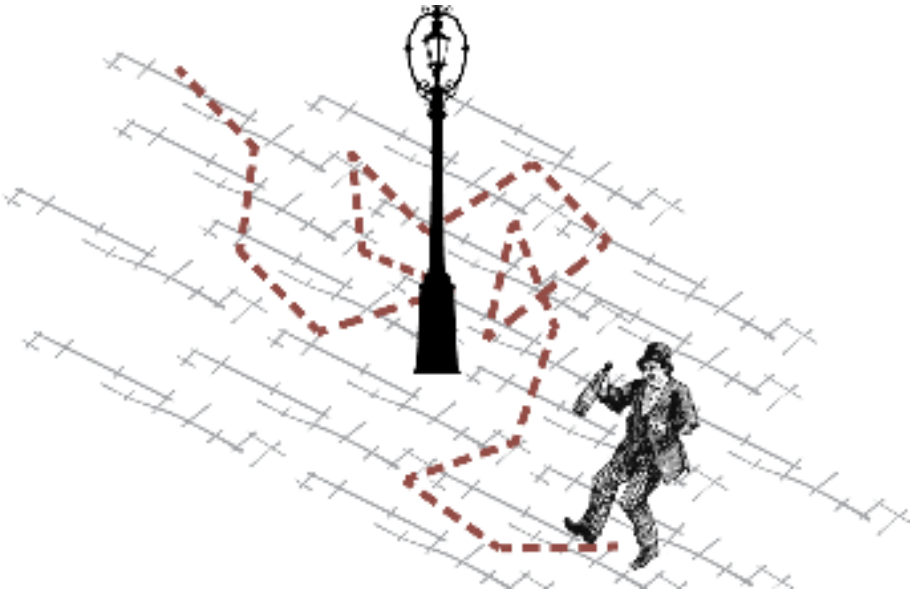
Cornell University

Random walk

A first example of dynamics

Random Walk

A random walk starting at node u is a walk where the edges are chosen at random, by selecting the next node to land on randomly in the (out) neighborhood.



- In the *uniform random walk*, the probability of selecting a certain edge at step e_k is uniform among all edges, i.e. $P(e_k = (v_{k-1}, v_k)) = \frac{1}{|\mathcal{N}^+_{v_{k-1}}|}$, or is proportional to the edge weights (equal if they add up to one).

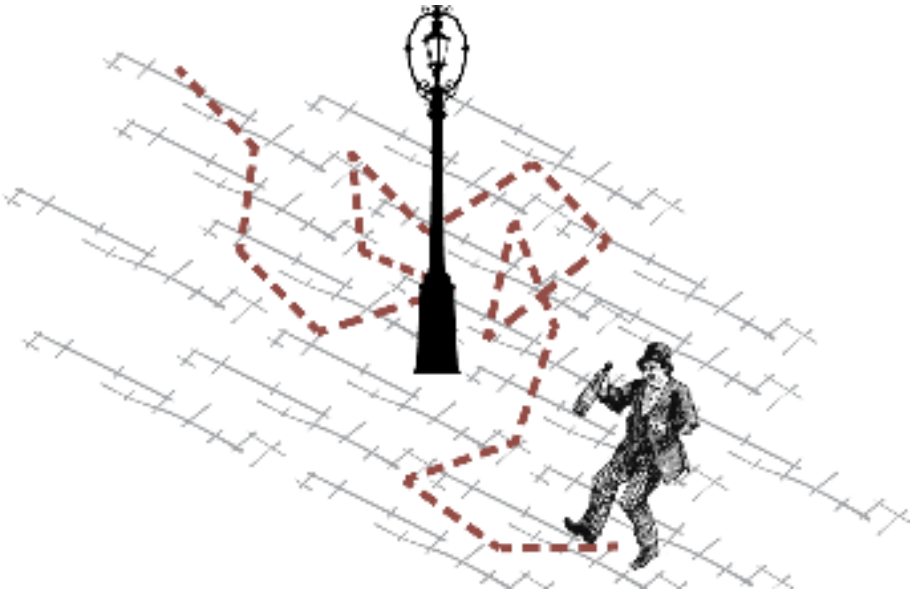
Random walk

A first example of dynamics

Cornell University

Random Walk

A random walk starting at node u is a walk where the edges are chosen at random, by selecting the next node to land on randomly in the (out) neighborhood.



- In the *uniform random walk*, the probability of selecting a certain edge at step e_k is uniform among all edges, i.e. $P(e_k = (v_{k-1}, v_k)) = \frac{1}{|\mathcal{N}^+_{v_{k-1}}|}$, or is proportional to the edge weights (equal if they add up to one).

Probability of visiting a node

- The random walk does not need to be uniform
- In general there can be a different probability of selecting a neighbor
- There is also the possibility of remaining in the same position
- Random walk can model anything that can spread from a site to another
  - Epidemics
  - Navigating at random a network of links
  - Spread of opinions in social networks
  - Flocking behavior
  - Synchronization

Probability of visiting a node

- The random walk does not need to be uniform
- In general there can be a different probability of selecting a neighbor
- There is also the possibility of remaining in the same position
- Random walk can model anything that can spread from a site to another
  - Epidemics
  - Navigating at random a network of links
  - Spread of opinions in social networks
  - Flocking behavior
  - Synchronization

Probability of visiting a node

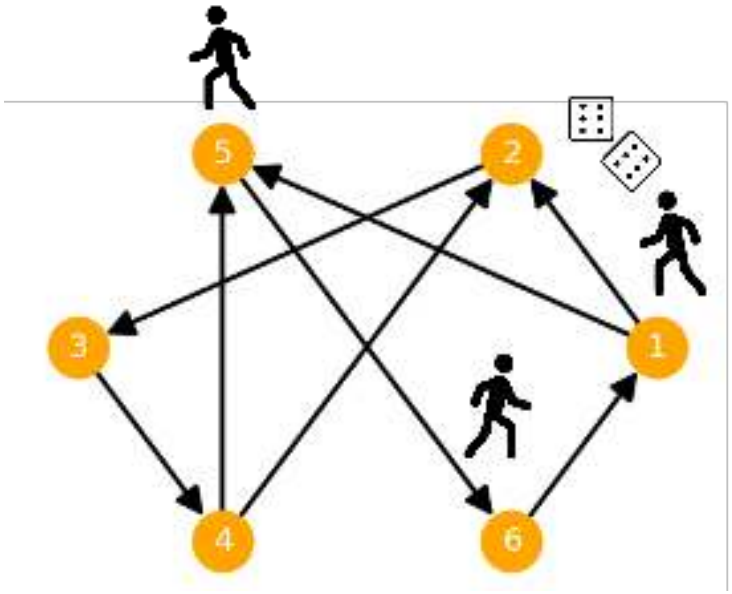
- The random walk does not need to be uniform
- In general there can be a different probability of selecting a neighbor
- There is also the possibility of remaining in the same position
- Random walk can model anything that can spread from a site to another
  - Epidemics
  - Navigating at random a network of links
  - Spread of opinions in social networks
  - Flocking behavior
  - Synchronization

Probability of visiting a node

- The random walk does not need to be uniform
- In general there can be a different probability of selecting a neighbor
- There is also the possibility of remaining in the same position
- Random walk can model anything that can spread from a site to another
  - Epidemics
  - Navigating at random a network of links
  - Spread of opinions in social networks
  - Flocking behavior
  - Synchronization

Reviewing Markov chain

- In a random walk on a network, the current position limits where I can go in the next step.
- The stochastic model fits in that of a Markov chain (MC).

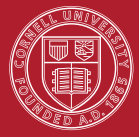


Markov chain (MC)

A MC is a random repeated experiment where the outcome at the tth iteration depends on what happened in the *previous* one **only**. The future depends on the present but not on the past.

- The outcomes of an MC are called states.

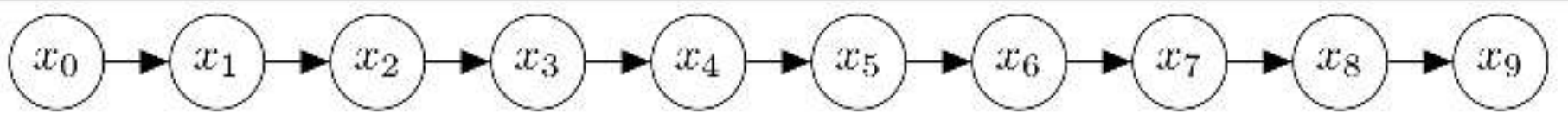
Markov chain as a Bayesian Network



- MC are very simple graphical models, but the graph that describes the MC shall not be confused with the Bayesian graph of an MC.
- The Bayesian Network of a MC is the line graph: each node has only one parent, the previous random experiment, i.e.:

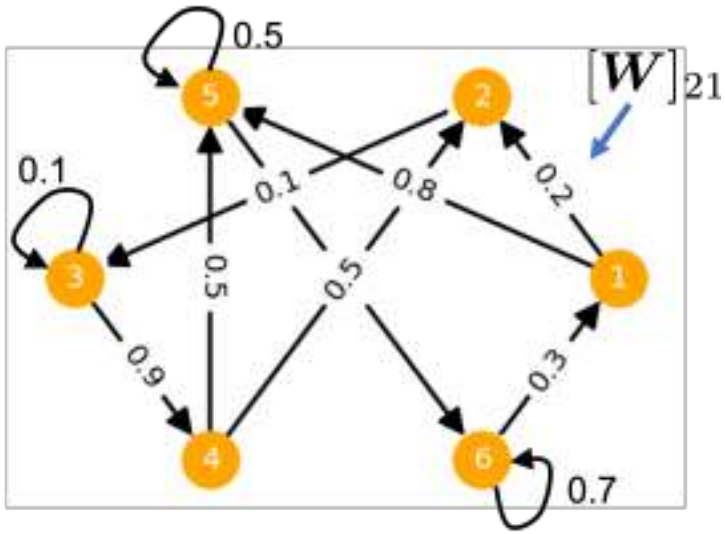
$$P(x(t)|\boldsymbol{Pa}(x(t))) \equiv P(x(t)|x(t-1))$$ transition probabilities

Each $x(t)$ is the random value seen at the t^{rh} experiment



Markov chain as stochastic state machine/graph

- In a Bayesian Network we need the probability of the *child given the parent* outcomes.
- A matrix $\boldsymbol{W}$ can specify the all transition probabilities
- The MC graph nodes are outcomes in the sample space and the edge weights are the transition probabilities.



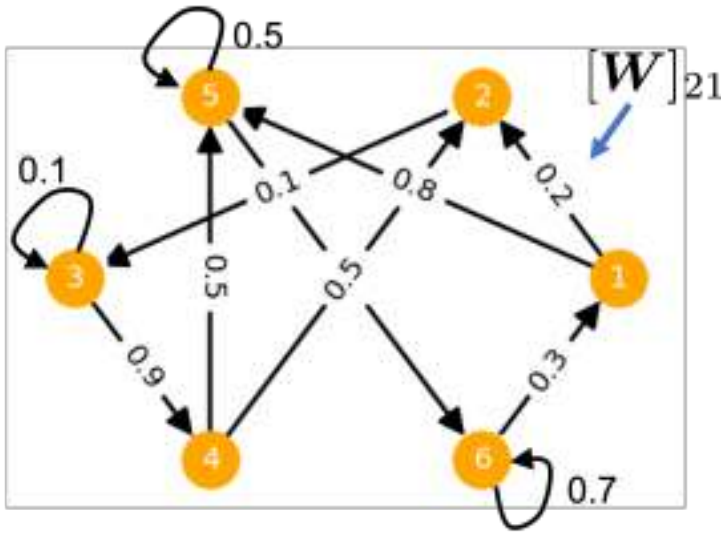
$$\boldsymbol{W}: \ [\boldsymbol{W}]_{uv} = w_{uv} = P(x(t) = u | x(t-1) = v),$$

$$w_{uu} = 1 - \sum_{v \neq u} w_{uv} \quad \rightarrow \text{ they add up to one}$$

$$\Rightarrow \ \mathbf{1}^T \boldsymbol{W} = \mathbf{1}^\top \quad \rightarrow \ \boldsymbol{W} \text{ stochastic matrix}$$

Markov chain as stochastic state machine/graph

- In a Bayesian Network we need the probability of the *child given the parent* outcomes.
- A matrix $\boldsymbol{W}$ can specify the all transition probabilities
- The MC graph nodes are outcomes in the sample space and the edge weights are the transition probabilities.



$$\boldsymbol{W}: \ [\boldsymbol{W}]_{uv} = w_{uv} = P(x(t) = u | x(t-1) = v),$$

$$w_{uu} = 1 - \sum_{v \neq u} w_{uv} \quad \rightarrow \text{ they add up to one}$$

$$\Rightarrow \ \mathbf{1}^T \boldsymbol{W} = \mathbf{1}^\top \quad \rightarrow \ \boldsymbol{W} \text{ stochastic matrix}$$

Markov chain Statistics

- Let $\boldsymbol{\pi}(t)$ be the probability mass function of an outcome, which in MC is also called the state probability vector:

$$[\boldsymbol{\pi}(t)]_v = \pi_v(t) = P(x(t) = v)$$

Chapman Kolmogorov Theorem

The probability of being in a certain state v (i.e. having a certain outcome) in an MC can be obtained from the probability of being in a state at the previous time instant as follows:

$$\pi_v(t) = \sum_{u=1}^{S} w_{vu}\pi_u(t-1) \quad \Longrightarrow \quad \boldsymbol{\pi}(t) = \boldsymbol{W}\boldsymbol{\pi}(t-1)$$

- Note that one may know the first outcome $x_0 = u$ and in that case $\boldsymbol{\pi}(0)$ has all entry equal to zero except for the $u$tn entry.

Markov chain Statistics

- Let $\boldsymbol{\pi}(t)$ be the probability mass function of an outcome, which in MC is also called the state probability vector:

$$[\boldsymbol{\pi}(t)]_v = \pi_v(t) = P(x(t) = v)$$

Chapman Kolmogorov Theorem

The probability of being in a certain state v (i.e. having a certain outcome) in an MC can be obtained from the probability of being in a state at the previous time instant as follows:

$$\pi_v(t) = \sum_{u=1}^{S} w_{vu}\pi_u(t-1) \quad \Longrightarrow \quad \boldsymbol{\pi}(t) = \boldsymbol{W}\boldsymbol{\pi}(t-1)$$

- Note that one may know the first outcome $x_0 = u$ and in that case $\boldsymbol{\pi}(0)$ has all entry equal to zero except for the $u$tn entry.

Markov chain Statistics

- Let $\boldsymbol{\pi}(t)$ be the probability mass function of an outcome, which in MC is also called the state probability vector:

$$[\boldsymbol{\pi}(t)]_v = \pi_v(t) = P(x(t) = v)$$

Chapman Kolmogorov Theorem

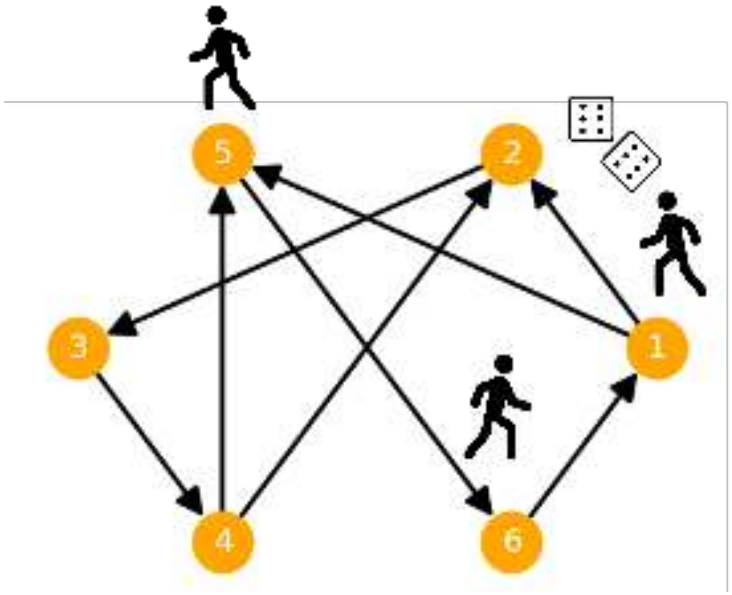
The probability of being in a certain state v (i.e. having a certain outcome) in an MC can be obtained from the probability of being in a state at the previous time instant as follows:

$$\pi_v(t) = \sum_{u=1}^{S} w_{vu}\pi_u(t-1) \quad \Longrightarrow \quad \boldsymbol{\pi}(t) = \boldsymbol{W}\boldsymbol{\pi}(t-1)$$

- Note that one may know the first outcome $x_0 = u$ and in that case $\boldsymbol{\pi}(0)$ has all entry equal to zero except for the $u$tn entry.

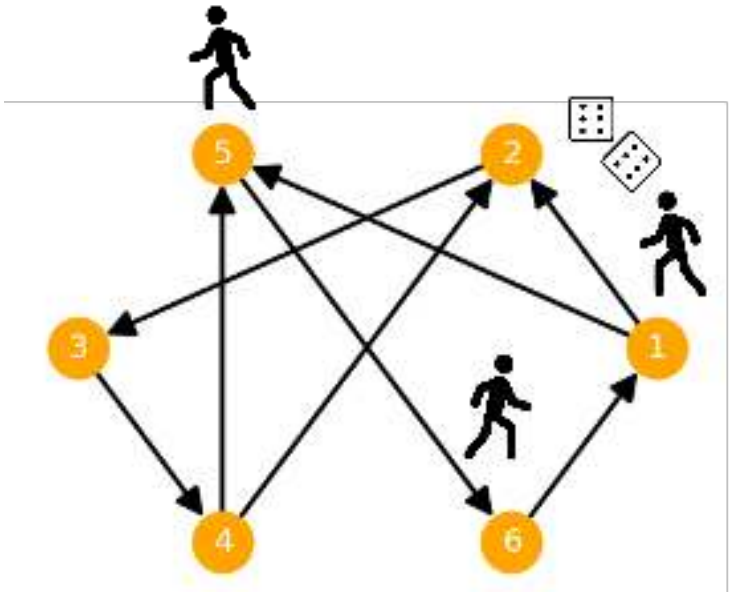
Random walk

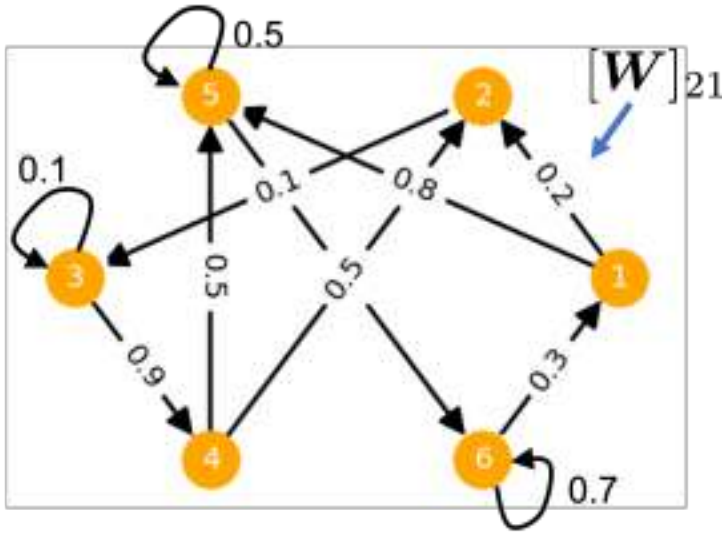
- The random walk has as outcomes $x(t) = v$ some node in $\mathcal{V}$.
- Given $x(t) = v$ the only possible new outcomes are neighboring nodes, i.e. the conditional experiment has sample space $\mathcal{N}_v^+$



Markov chain graph and random walk

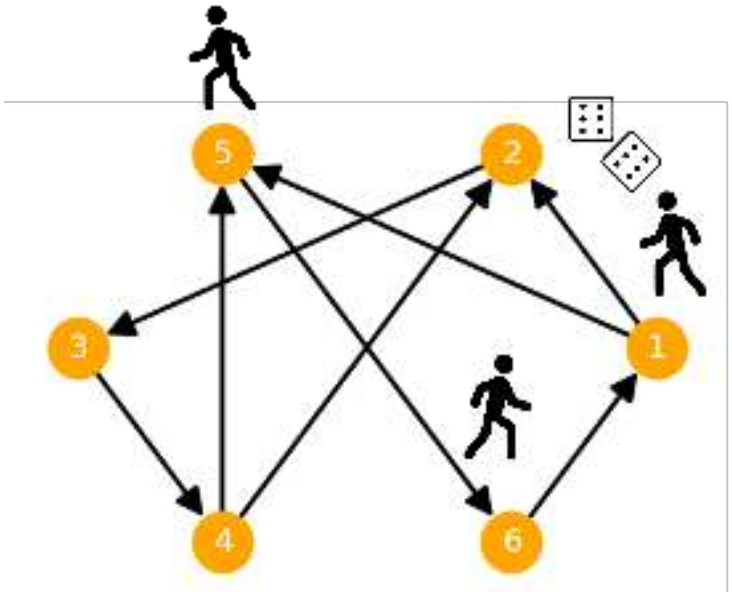
- The MC graph nodes are the outcomes of the stochastic experiment.
- The edges represent what are the possible new outcomes, given the present outcome. For these the probabilities $P(x(t) = v|x(t-1) = u)$ are non-zero.

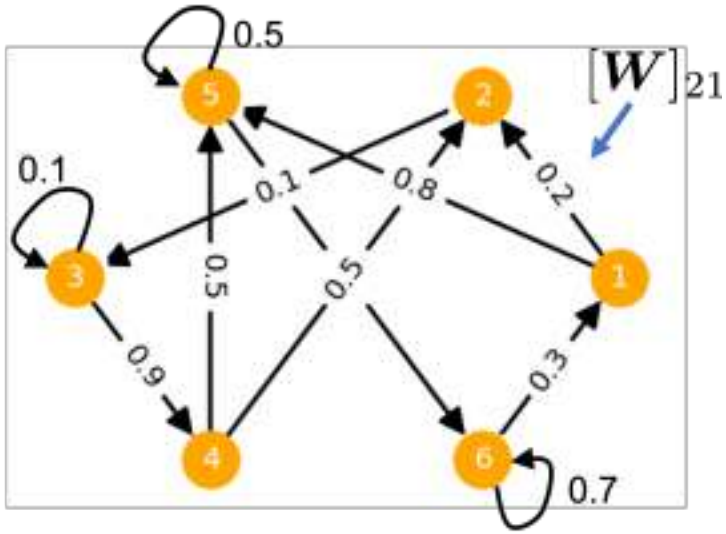




Markov chain graph and random walk

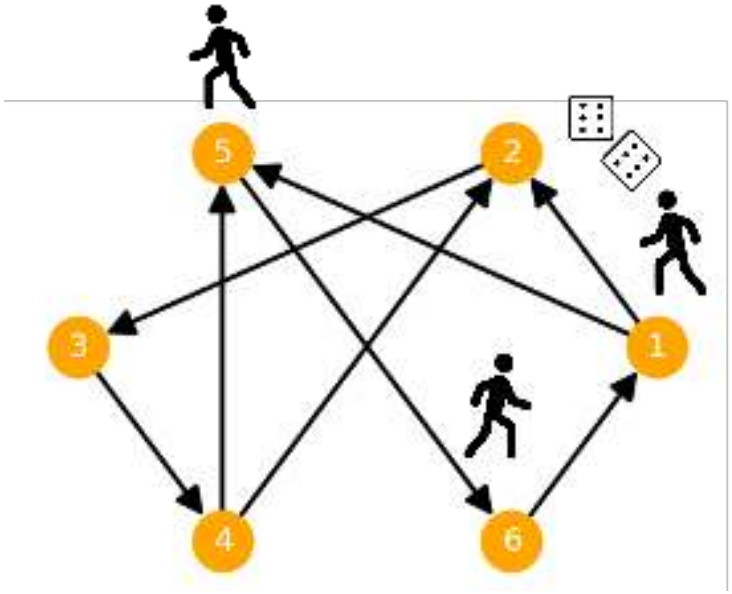
- The MC graph nodes are the outcomes of the stochastic experiment.
- The edges represent what are the possible new outcomes, given the present outcome. For these the probabilities $P(x(t) = v|x(t-1) = u)$ are non-zero.

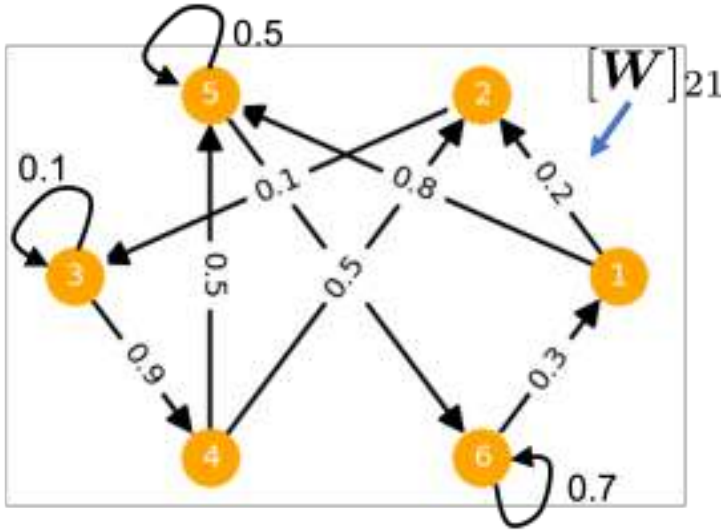




Markov chain graph and random walk

Cornell University





- The random walk graph and the MC graph are the same:
  1. a step can only land in one node $\rightarrow$ the MC sample space is $\mathcal{V}$
  2. the viable transitions are where the graph has an edge. The probabilities can be interpreted as edge weights.

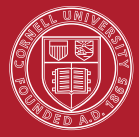
Probability of a walk

- Chapman Kolmogorov theorem only specifies the probability of landing somewhere in a single step is $\boldsymbol{\pi}(t) = \boldsymbol{W}\boldsymbol{\pi}(t-1)$:
- The probability mass function of landing at the various nodes after $t \geq$ steps is obtained (without loss of generality we start from $t = 0$) using the same law recursively:

$$\boldsymbol{\pi}(t) = \boldsymbol{W}^t \boldsymbol{\pi}(0)$$

- The probability of a t-long walk from u to v is $[\boldsymbol{W}^t]_{vu}$.

Probability of a walk

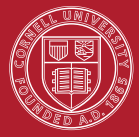


- Chapman Kolmogorov theorem only specifies the probability of landing somewhere in a single step is $\boldsymbol{\pi}(t) = \boldsymbol{W}\boldsymbol{\pi}(t-1)$:
- The probability mass function of landing at the various nodes after $t \geq$ steps is obtained (without loss of generality we start from $t = 0$) using the same law recursively:

$$\boldsymbol{\pi}(t) = \boldsymbol{W}^t \boldsymbol{\pi}(0)$$

- The probability of a t-long walk from u to v is $[\boldsymbol{W}^t]_{vu}$.

Probability of a walk

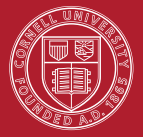


- Chapman Kolmogorov theorem only specifies the probability of landing somewhere in a single step is $\boldsymbol{\pi}(t) = \boldsymbol{W}\boldsymbol{\pi}(t-1)$:
- The probability mass function of landing at the various nodes after $t \geq$ steps is obtained (without loss of generality we start from $t = 0$) using the same law recursively:

$$\boldsymbol{\pi}(t) = \boldsymbol{W}^t \boldsymbol{\pi}(0)$$

- The probability of a t-long walk from u to v is $[\boldsymbol{W}^t]_{vu}$.

Uniform random walk



- Let $0 \leq \alpha \leq 1$. In this type of random walk the probability of choosing a neighboring node is uniform, that is:

$$w_{u,v} = \frac{\alpha}{d_u} \ , v \in \mathcal{N}_u^+, \ v \neq u,$$

$$w_{uu} = 1 - \sum_{v \in \mathcal{N}_u^+} w_{u,v} = 1 - d_u \frac{\alpha}{d_u} = (1-\alpha)$$

and for $\alpha = 1$ the walk cannot stay in the same place.

Uniform random walk

Cornell University

- Therefore, denote by $\text{diag}(\boldsymbol{d})$ a diagonal matrix with diagonal elements equal to the vector $\boldsymbol{d}$

$$\Rightarrow \quad \boldsymbol{W} = (1-\alpha)\boldsymbol{I} + \alpha\, \text{diag}(\boldsymbol{d})^{-1}\boldsymbol{A} = \boldsymbol{I} - \alpha \overbrace{(\boldsymbol{I} - \text{diag}(\boldsymbol{d})^{-1}\boldsymbol{A})}^{\boldsymbol{L}^{\text{rw}}}$$

where $\boldsymbol{A}$ is the unweighted *graph adjacency matrix* and $\boldsymbol{L}^{\text{rw}}$ is a normalized random walk Laplacian.

- Note that in general the network is not directed, and in this case the degree sequence vector d is actually the vector of outdegrees $d^{\text{out}} = A\mathbf{1}$

Uniform random walk

Cornell University

- Therefore, denote by $\text{diag}(\boldsymbol{d})$ a diagonal matrix with diagonal elements equal to the vector $\boldsymbol{d}$

$$\Rightarrow \quad \boldsymbol{W} = (1-\alpha)\boldsymbol{I} + \alpha\, \text{diag}(\boldsymbol{d})^{-1}\boldsymbol{A} = \boldsymbol{I} - \alpha \overbrace{(\boldsymbol{I} - \text{diag}(\boldsymbol{d})^{-1}\boldsymbol{A})}^{\boldsymbol{L}^{\text{rw}}}$$

  where $\boldsymbol{A}$ is the unweighted *graph adjacency matrix* and $\boldsymbol{L}^{\text{rw}}$ is a normalized random walk Laplacian.

- Note that in general the network is not directed, and in this case the degree sequence vector $\boldsymbol{d}$ is actually the vector of outdegrees $\boldsymbol{d}^{\text{out}} = \boldsymbol{A}\boldsymbol{1}$

Uniform random walk and the spectrum of L^{rw}

- Another interesting observation regarding the **uniform random walks** comes from the identity:

$$\boldsymbol{W} = \boldsymbol{I} - \alpha \boldsymbol{L}^{\text{rw}}$$

- In general, for any matrix $\boldsymbol{Q}$, with EVD $\boldsymbol{Q} = \boldsymbol{U}\text{diag}(\boldsymbol{\lambda})\boldsymbol{U}^{-1}$, the matrix $\boldsymbol{I} + \boldsymbol{Q}$ has an eigenvalue decomposition

$$\boldsymbol{Q} = \boldsymbol{U}\text{diag}(\boldsymbol{\lambda})\boldsymbol{U}^{-1} \quad \Rightarrow \quad \boldsymbol{I} + \boldsymbol{Q} = \boldsymbol{U}(\boldsymbol{I} + \text{diag}(\boldsymbol{\lambda}))\boldsymbol{U}^{-1}$$

- Hence, for a uniform random walk:
  1. the eigenvectors of $\boldsymbol{W}$, and $\boldsymbol{L}^{\text{rw}}$ are the same
  2. Let $\lambda_v(\boldsymbol{W})$ be the eigenvalues of $\boldsymbol{W}$ and $\lambda_v(\boldsymbol{L}^{\text{rw}})$, sorted in decreasting order. Then, $\lambda_v(\boldsymbol{W}) = 1 - \alpha\lambda_{N-v}(\boldsymbol{L}^{\text{rw}})$
  3. The eigenvectors are also the same, but in reverse order: $\boldsymbol{u}_{N-v}$ eigenvector for $\boldsymbol{L}^{\text{rw}}$ is $\boldsymbol{u}_v$ for $\boldsymbol{W}$.

Uniform random walk and the spectrum of L^{rw}

- Another interesting observation regarding the **uniform random walks** comes from the identity:

$$\boldsymbol{W} = \boldsymbol{I} - \alpha \boldsymbol{L}^{\text{rw}}$$

- In general, for any matrix $\boldsymbol{Q}$, with EVD $\boldsymbol{Q} = \boldsymbol{U}\text{diag}(\boldsymbol{\lambda})\boldsymbol{U}^{-1}$, the matrix $\boldsymbol{I} + \boldsymbol{Q}$ has an eigenvalue decomposition

$$\boldsymbol{Q} = \boldsymbol{U}\text{diag}(\boldsymbol{\lambda})\boldsymbol{U}^{-1} \quad \Rightarrow \quad \boldsymbol{I} + \boldsymbol{Q} = \boldsymbol{U}(\boldsymbol{I} + \text{diag}(\boldsymbol{\lambda}))\boldsymbol{U}^{-1}$$

- Hence, for a uniform random walk:
  1. the eigenvectors of $\boldsymbol{W}$, and $\boldsymbol{L}^{\text{rw}}$ are the same
  2. Let $\lambda_v(\boldsymbol{W})$ be the eigenvalues of $\boldsymbol{W}$ and $\lambda_v(\boldsymbol{L}^{\text{rw}})$, sorted in decreasting order. Then, $\lambda_v(\boldsymbol{W}) = 1 - \alpha\lambda_{N-v}(\boldsymbol{L}^{\text{rw}})$
  3. The eigenvectors are also the same, but in reverse order: $\boldsymbol{u}_{N-v}$ eigenvector for $\boldsymbol{L}^{\text{rw}}$ is $\boldsymbol{u}_v$ for $\boldsymbol{W}$.

Uniform random walk and the spectrum of L^{rw}

- Another interesting observation regarding the **uniform random walks** comes from the identity:

$$\boldsymbol{W} = \boldsymbol{I} - \alpha \boldsymbol{L}^{\text{rw}}$$

- In general, for any matrix $\boldsymbol{Q}$, with EVD $\boldsymbol{Q} = \boldsymbol{U}\text{diag}(\boldsymbol{\lambda})\boldsymbol{U}^{-1}$, the matrix $\boldsymbol{I} + \boldsymbol{Q}$ has an eigenvalue decomposition

$$\boldsymbol{Q} = \boldsymbol{U}\text{diag}(\boldsymbol{\lambda})\boldsymbol{U}^{-1} \quad \Rightarrow \quad \boldsymbol{I} + \boldsymbol{Q} = \boldsymbol{U}(\boldsymbol{I} + \text{diag}(\boldsymbol{\lambda}))\boldsymbol{U}^{-1}$$

- Hence, for a uniform random walk:
  1. the eigenvectors of $\boldsymbol{W}$, and $\boldsymbol{L}^{\text{rw}}$ are the same
  2. Let $\lambda_v(\boldsymbol{W})$ be the eigenvalues of $\boldsymbol{W}$ and $\lambda_v(\boldsymbol{L}^{\text{rw}})$, sorted in decreasting order. Then, $\lambda_v(\boldsymbol{W}) = 1 - \alpha\lambda_{N-v}(\boldsymbol{L}^{\text{rw}})$
  3. The eigenvectors are also the same, but in reverse order: $\boldsymbol{u}_{N-v}$ eigenvector for $\boldsymbol{L}^{\text{rw}}$ is $\boldsymbol{u}_v$ for $\boldsymbol{W}$.

Uniform random walk and the spectrum of L^{rw}

- Another interesting observation regarding the **uniform random walks** comes from the identity:

$$\boldsymbol{W} = \boldsymbol{I} - \alpha \boldsymbol{L}^{\text{rw}}$$

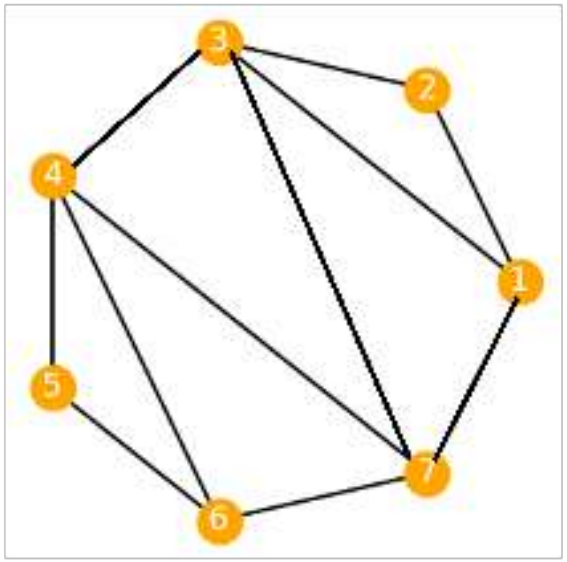
- In general, for any matrix $\boldsymbol{Q}$, with EVD $\boldsymbol{Q} = \boldsymbol{U}\text{diag}(\boldsymbol{\lambda})\boldsymbol{U}^{-1}$, the matrix $\boldsymbol{I} + \boldsymbol{Q}$ has an eigenvalue decomposition

$$\boldsymbol{Q} = \boldsymbol{U}\text{diag}(\boldsymbol{\lambda})\boldsymbol{U}^{-1} \quad \Rightarrow \quad \boldsymbol{I} + \boldsymbol{Q} = \boldsymbol{U}(\boldsymbol{I} + \text{diag}(\boldsymbol{\lambda}))\boldsymbol{U}^{-1}$$

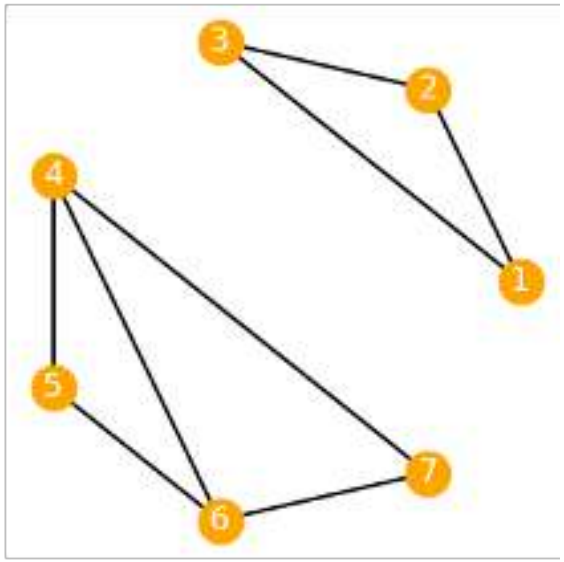
- Hence, for a uniform random walk:
  1. the eigenvectors of $\boldsymbol{W}$, and $\boldsymbol{L}^{\text{rw}}$ are the same
  2. Let $\lambda_v(\boldsymbol{W})$ be the eigenvalues of $\boldsymbol{W}$ and $\lambda_v(\boldsymbol{L}^{\text{rw}})$, sorted in decreasting order. Then, $\lambda_v(\boldsymbol{W}) = 1 - \alpha\lambda_{N-v}(\boldsymbol{L}^{\text{rw}})$
  3. The eigenvectors are also the same, but in reverse order: $\boldsymbol{u}_{N-v}$ eigenvector for $\boldsymbol{L}^{\text{rw}}$ is $\boldsymbol{u}_v$ for $\boldsymbol{W}$.

Uniform random walk and the spectrum of L^{rw}

- Recall that for the Laplacian $\lambda_N(\boldsymbol{L}^{\text{rw}}) = 0$ and the eigenvector is $\mathbf{1}$. It follows that $\lambda_1(\boldsymbol{W}) = 1 - \alpha\lambda_N(\boldsymbol{L}^{\text{rw}}) = 1$.
- Also, $\lambda_2(\boldsymbol{W}) = 1 - \lambda_{N-1}(\boldsymbol{L}^{\text{rw}})$.
- Also, recall $\lambda_{N-1}(\boldsymbol{L}^{\text{rw}})$ is the **algebraic connectivity of the graph** and it is $\lambda_{N-1}(\boldsymbol{L}^{\text{rw}}) > 0$ if and only if the graph is strongly connected.



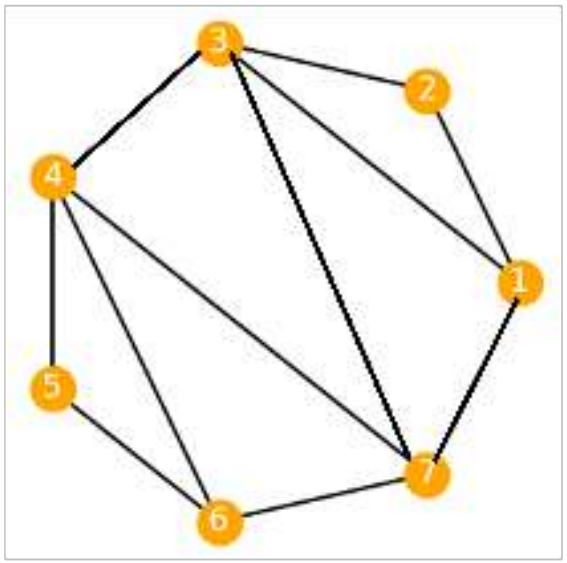
$\lambda_{N-1}(\boldsymbol{L}^{\text{rw}}) > 0$



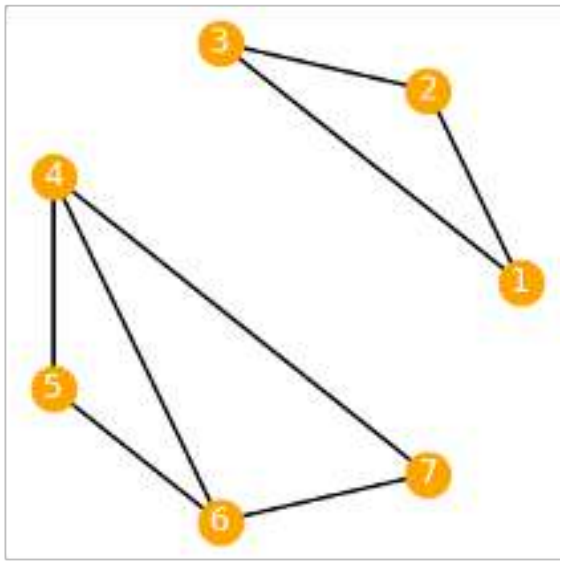
$\lambda_{N-1}(\boldsymbol{L}^{\text{rw}}) = 0$

Uniform random walk and the spectrum of L^{rw}

- Recall that for the Laplacian $\lambda_N(\boldsymbol{L}^{\text{rw}}) = 0$ and the eigenvector is $\mathbf{1}$. It follows that $\lambda_1(\boldsymbol{W}) = 1 - \alpha\lambda_N(\boldsymbol{L}^{\text{rw}}) = 1$.
- Also, $\lambda_2(\boldsymbol{W}) = 1 - \lambda_{N-1}(\boldsymbol{L}^{\text{rw}})$.
- Also, recall $\lambda_{N-1}(\boldsymbol{L}^{\text{rw}})$ is the **algebraic connectivity of the graph** and it is $\lambda_{N-1}(\boldsymbol{L}^{\text{rw}}) > 0$ if and only if the graph is strongly connected.



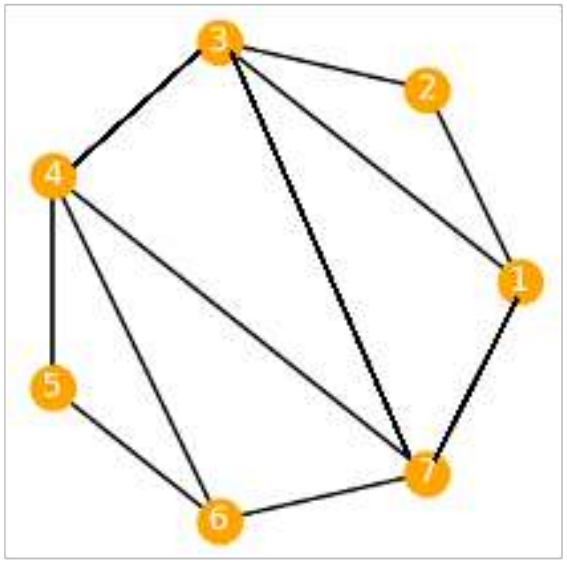
$\lambda_{N-1}(\boldsymbol{L}^{\text{rw}}) > 0$



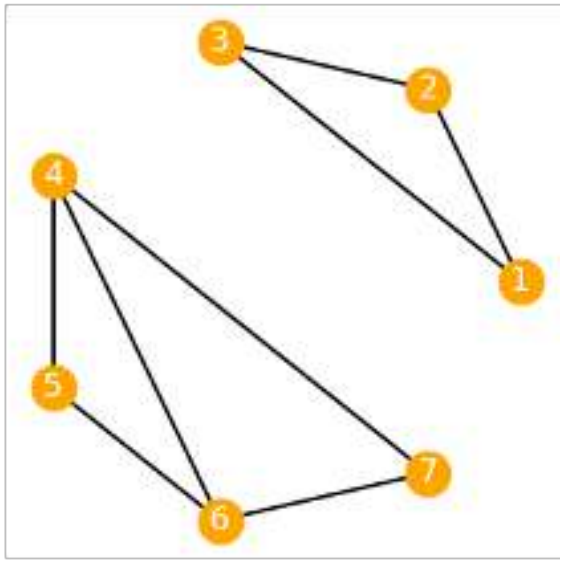
$\lambda_{N-1}(\boldsymbol{L}^{\text{rw}}) = 0$

Uniform random walk and the spectrum of L^{rw}

- Recall that for the Laplacian $\lambda_N(\boldsymbol{L}^{\text{rw}}) = 0$ and the eigenvector is $\mathbf{1}$. It follows that $\lambda_1(\boldsymbol{W}) = 1 - \alpha\lambda_N(\boldsymbol{L}^{\text{rw}}) = 1$.
- Also, $\lambda_2(\boldsymbol{W}) = 1 - \lambda_{N-1}(\boldsymbol{L}^{\text{rw}})$.
- Also, recall $\lambda_{N-1}(\boldsymbol{L}^{\text{rw}})$ is the **algebraic connectivity of the graph** and it is $\lambda_{N-1}(\boldsymbol{L}^{\text{rw}}) > 0$ if and only if the graph is strongly connected.



$\lambda_{N-1}(\boldsymbol{L}^{\text{rw}}) > 0$



$\lambda_{N-1}(\boldsymbol{L}^{\text{rw}}) = 0$

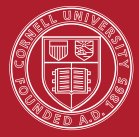
Uniform random walk and the spectrum of L^{rw}

- It follows that if and only if the graph is strongly connected $\lambda_{N-1}(\boldsymbol{L}^{\text{rw}}) > 0$, and it must be $\lambda_2(\boldsymbol{W}) < 1$
- A connected graph for the walk, means that the associated Markov Process has an **irreducible** transition probability matrix
- This is important because the eigenvalue decomposition of $\boldsymbol{W}$ helps understand the asymptotic properties of linear dynamics, such as $\boldsymbol{\pi}(t) = \boldsymbol{W}\boldsymbol{\pi}(t-1)$.

Uniform random walk and the spectrum of L^{rw}

- It follows that if and only if the graph is strongly connected $\lambda_{N-1}(\boldsymbol{L}^{\text{rw}}) > 0$, and it must be $\lambda_2(\boldsymbol{W}) < 1$
- A connected graph for the walk, means that the associated Markov Process has an **irreducible** transition probability matrix
- This is important because the eigenvalue decomposition of $\boldsymbol{W}$ helps understand the asymptotic properties of linear dynamics, such as $\boldsymbol{\pi}(t) = \boldsymbol{W}\boldsymbol{\pi}(t-1)$.

Some useful algebra



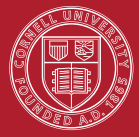
- In general, let consider a vector $\boldsymbol{x}(t)$ that represents the state of the network. **Linear dynamics** change the state dynamically with the following equations:

$$\boldsymbol{x}(t) = \boldsymbol{Q}\boldsymbol{x}(t-1) \quad \Rightarrow \quad \boldsymbol{x}(t) = \boldsymbol{Q}^t\boldsymbol{x}(0)$$

- In particular, let the eigenvalue decomposition of $\boldsymbol{Q} = \boldsymbol{U}\text{diag}(\boldsymbol{\lambda})\boldsymbol{U}^{-1}$, where $\boldsymbol{U}$ are the eigenvectors of $\boldsymbol{Q}$ and $\boldsymbol{\lambda}$ is the vector of eigenvalues, $(\lambda_1 \geq \lambda_2, \ldots, \lambda_N)$.
- Let us define $\boldsymbol{V}^\top := \boldsymbol{U}^{-1}$, then:

$$\boldsymbol{Q}^t = \boldsymbol{U}(\text{diag}(\boldsymbol{\lambda}))^t\boldsymbol{V}^\top \equiv \sum_{v=1}^{N} \lambda_v^t \boldsymbol{u}_v \boldsymbol{v}^\top$$

Some useful algebra



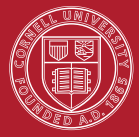
- In general, let consider a vector $\boldsymbol{x}(t)$ that represents the state of the network. **Linear dynamics** change the state dynamically with the following equations:

$$\boldsymbol{x}(t) = \boldsymbol{Q}\boldsymbol{x}(t-1) \quad \Rightarrow \quad \boldsymbol{x}(t) = \boldsymbol{Q}^t\boldsymbol{x}(0)$$

- In particular, let the eigenvalue decomposition of $\boldsymbol{Q} = \boldsymbol{U}\text{diag}(\boldsymbol{\lambda})\boldsymbol{U}^{-1}$, where $\boldsymbol{U}$ are the eigenvectors of $\boldsymbol{Q}$ and $\boldsymbol{\lambda}$ is the vector of eigenvalues, $(\lambda_1 \geq \lambda_2, \ldots, \lambda_N)$.
- Let us define $\boldsymbol{V}^\top := \boldsymbol{U}^{-1}$, then:

$$\boldsymbol{Q}^t = \boldsymbol{U}(\text{diag}(\boldsymbol{\lambda}))^t\boldsymbol{V}^\top \equiv \sum_{v=1}^{N} \lambda_v^t \boldsymbol{u}_v \boldsymbol{v}^\top$$

Some useful algebra



- In general, let consider a vector $\boldsymbol{x}(t)$ that represents the state of the network. **Linear dynamics** change the state dynamically with the following equations:

$$\boldsymbol{x}(t) = \boldsymbol{Q}\boldsymbol{x}(t-1) \quad \Rightarrow \quad \boldsymbol{x}(t) = \boldsymbol{Q}^t\boldsymbol{x}(0)$$

- In particular, let the eigenvalue decomposition of $\boldsymbol{Q} = \boldsymbol{U}\text{diag}(\boldsymbol{\lambda})\boldsymbol{U}^{-1}$, where $\boldsymbol{U}$ are the eigenvectors of $\boldsymbol{Q}$ and $\boldsymbol{\lambda}$ is the vector of eigenvalues, $(\lambda_1 \geq \lambda_2, \ldots, \lambda_N)$.
- Let us define $\boldsymbol{V}^\top := \boldsymbol{U}^{-1}$, then:

$$\boldsymbol{Q}^t = \boldsymbol{U}(\text{diag}(\boldsymbol{\lambda}))^t\boldsymbol{V}^\top \equiv \sum_{v=1}^{N} \lambda_v^t \boldsymbol{u}_v \boldsymbol{v}^\top$$

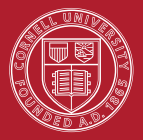
Asymptotic trends for linear dynamics

- As a consequence:

$$\Rightarrow \quad \boldsymbol{x}(t) = \boldsymbol{Q}^t \boldsymbol{x}(0) = \sum_{v=1}^{N} \lambda_v^t \boldsymbol{u}_v \overbrace{\boldsymbol{v}_v^\top \boldsymbol{x}(0)}^{\xi_v(0)} = \sum_{v=1}^{N} \lambda_v^t \xi_v(0) \boldsymbol{u}_v$$

- We can see three scenarios:
  1. some of the eigenvalues $\lambda_v > 1$ have $\lim_{t\to+\infty} |\lambda_v|^t = +\infty$; in this case the matrix state vector $\boldsymbol{x}(t)$ blows up
  2. all of the eigenvalues $|\lambda_v| < 1$ so all for of them $\lim_{t\to+\infty} |\lambda_v|^t = 0$; in this case the norm of state decreases exponentially
  3. There are some eigenvalues with magnitude equal to 1 and all the rest have magnitude less than 1. The *modes* $\boldsymbol{u}_v$ with $|\lambda_v| = 1$ persist.

Asymptotic trends for linear dynamics



- As a consequence:

$$\Rightarrow \quad \boldsymbol{x}(t) = \boldsymbol{Q}^t \boldsymbol{x}(0) = \sum_{v=1}^{N} \lambda_v^t \boldsymbol{u}_v \overbrace{\boldsymbol{v}_v^\top \boldsymbol{x}(0)}^{\xi_v(0)} = \sum_{v=1}^{N} \lambda_v^t \xi_v(0) \boldsymbol{u}_v$$

- We can see three scenarios:
  1. some of the eigenvalues $\lambda_v > 1$ have $\lim_{t\to+\infty} |\lambda_v|^t = +\infty$; in this case the matrix state vector $\boldsymbol{x}(t)$ blows up
  2. all of the eigenvalues $|\lambda_v| < 1$ so all for of them $\lim_{t\to+\infty} |\lambda_v|^t = 0$; in this case the norm of state decreases exponentially
  3. There are some eigenvalues with magnitude equal to 1 and all the rest have magnitude less than 1. The *modes* $\boldsymbol{u}_v$ with $|\lambda_v| = 1$ persist.

Asymptotic trends for linear dynamics

- As a consequence:

$$\Rightarrow \quad \boldsymbol{x}(t) = \boldsymbol{Q}^t \boldsymbol{x}(0) = \sum_{v=1}^{N} \lambda_v^t \boldsymbol{u}_v \overbrace{\boldsymbol{v}_v^\top \boldsymbol{x}(0)}^{\xi_v(0)} = \sum_{v=1}^{N} \lambda_v^t \xi_v(0) \boldsymbol{u}_v$$

- We can see three scenarios:
  1. some of the eigenvalues $\lambda_v > 1$ have $\lim_{t\to+\infty} |\lambda_v|^t = +\infty$; in this case the matrix state vector $\boldsymbol{x}(t)$ blows up
  2. all of the eigenvalues $|\lambda_v| < 1$ so all for of them $\lim_{t\to+\infty} |\lambda_v|^t = 0$; in this case the norm of state decreases exponentially
  3. There are some eigenvalues with magnitude equal to 1 and all the rest have magnitude less than 1. The *modes* $\boldsymbol{u}_v$ with $|\lambda_v| = 1$ persist.

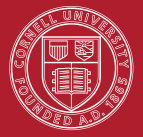
Asymptotic trends for linear dynamics

- As a consequence:

$$\Rightarrow \quad \boldsymbol{x}(t) = \boldsymbol{Q}^t \boldsymbol{x}(0) = \sum_{v=1}^{N} \lambda_v^t \boldsymbol{u}_v \overbrace{\boldsymbol{v}_v^\top \boldsymbol{x}(0)}^{\xi_v(0)} = \sum_{v=1}^{N} \lambda_v^t \xi_v(0) \boldsymbol{u}_v$$

- We can see three scenarios:
  1. some of the eigenvalues $\lambda_v > 1$ have $\lim_{t\to+\infty} |\lambda_v|^t = +\infty$; in this case the matrix state vector $\boldsymbol{x}(t)$ blows up
  2. all of the eigenvalues $|\lambda_v| < 1$ so all for of them $\lim_{t\to+\infty} |\lambda_v|^t = 0$; in this case the norm of state decreases exponentially
  3. There are some eigenvalues with magnitude equal to 1 and all the rest have magnitude less than 1. The *modes* $\boldsymbol{u}_v$ with $|\lambda_v| = 1$ persist.

Asymptotic trends for linear dynamics



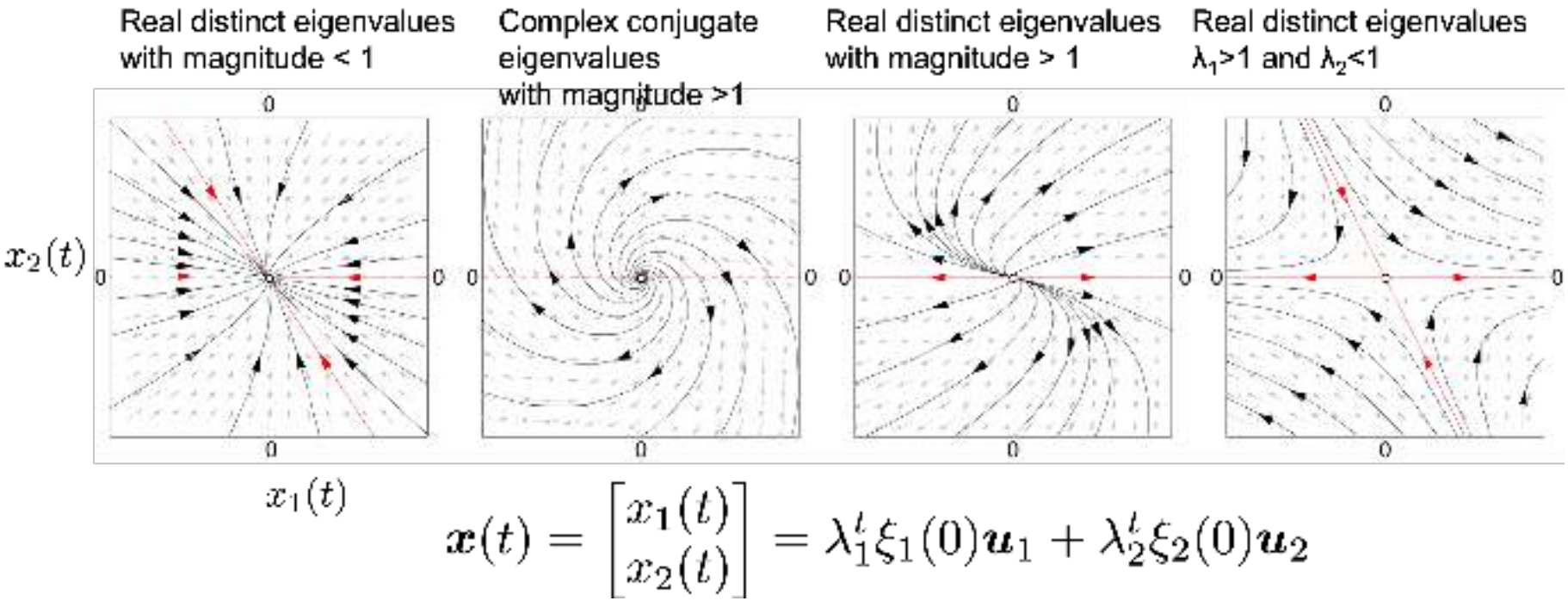
- Hence, there are only three possibilities for the limit:

$$\lim_{t\to+\infty} \boldsymbol{Q}^t = \begin{cases} +\boldsymbol{\infty} & \text{if } \exists\ |\lambda_v| > 1 \quad v = 1,\dots,N \\ \mathbf{0} & \text{if } |\lambda_1| < 1 \\ \sum_{v=1}^{M} \boldsymbol{u}_v \boldsymbol{v}_v^\top & \text{if } |\lambda_1| = \dots = |\lambda_M| = 1, M \leq N \end{cases}$$

Asymptotic trends for linear dynamics

- The state in the limit is:

$$\lim_{t\to+\infty} \boldsymbol{x}(t) = \begin{cases} +\infty & \text{if } \exists\ |\lambda_v| > 1 \quad v = 1,\dots,N \\ \mathbf{0} & \text{if } |\lambda_1| < 1 \\ \sum_{v=1}^{M} \xi_v(0) e^{jt\angle\lambda_v} \boldsymbol{u}_v & \text{if } |\lambda_1| = \dots = |\lambda_M| = 1, M \le N \end{cases}$$



$$\boldsymbol{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \lambda_1^t \xi_1(0)\boldsymbol{u}_1 + \lambda_2^t \xi_2(0)\boldsymbol{u}_2$$

Asymptotic properties for stochastic matrices

- Why to we care? The evolution of the state probability for the a Markov Chain in general, and for the random walk in particular is a linear dynamic, since $\boldsymbol{\pi}(t) = \boldsymbol{W}^t \boldsymbol{\pi}(0)$
- However, $\boldsymbol{W}$ does not have completely arbitrary entries, which limits what can happen:
  1. $w_{u,v} \geq 0$;
  2. $\sum_{u \in \mathcal{N}_v^{\text{out}}} w_{u,v} = 1$.
- These are the properties of a stochastic matrix.

Asymptotic properties for stochastic matrices

- Why to we care? The evolution of the state probability for the a Markov Chain in general, and for the random walk in particular is a linear dynamic, since $\boldsymbol{\pi}(t) = \boldsymbol{W}^t \boldsymbol{\pi}(0)$
- However, $\boldsymbol{W}$ does not have completely arbitrary entries, which limits what can happen:
  1. $w_{u,v} \geq 0$;
  2. $\sum_{u \in \mathcal{N}_v^{\text{out}}} w_{u,v} = 1$.
- These are the properties of a stochastic matrix.

Asymptotic properties for stochastic matrices

- Note that because $\boldsymbol{W}$ is a stochastic matrix (probabilities ought to add up to 1!) $\mathbf{1}^\top \boldsymbol{W} = \mathbf{1}^\top \quad \rightarrow \quad \mathbf{1}$ is a left eigenvector of $\boldsymbol{W}$ with eigenvalue equal to 1.
- Gershgorin circle theorem ensures that for a stochastic matrix $\lambda_v \leq 1$ and Perron Frobenious theorem that, if the MC form a strongly connected graph (it is called irreducible), there is only **one eigenvalue** $\lambda_1 = 1$.

Asymptotic properties for stochastic matrices

- Note that because $\boldsymbol{W}$ is a stochastic matrix (probabilities ought to add up to 1!) $\mathbf{1}^\top \boldsymbol{W} = \mathbf{1}^\top \quad \rightarrow \quad \mathbf{1}$ is a left eigenvector of $\boldsymbol{W}$ with eigenvalue equal to 1.
- Gershgorin circle theorem ensures that for a stochastic matrix $\lambda_v \leq 1$ and Perron Frobenious theorem that, if the MC form a strongly connected graph (it is called irreducible), there is only **one eigenvalue** $\lambda_1 = 1$.

Asymptotic properties for stochastic matrices

- Armed with this knowledge, we understand that for the random walk
  1. $\boldsymbol{W}^t = \boldsymbol{\pi}^* \mathbf{1}^\top + O(\lambda_2^t)$
  2. When the graph is connected $\lambda_2 < 1$ and so $O(\lambda_2^t) \to 0$ for $t \gg 1$
- Since $\mathbf{1}^\top \pi(0) = 1$:

$$\lim_{t \to +\infty} \boldsymbol{W}^t \boldsymbol{\pi}(0) = \boldsymbol{\pi}^* \overbrace{\mathbf{1}^\top \boldsymbol{\pi}(0)}^{=1} \equiv \boldsymbol{\pi}^* \quad \text{stationary state probability}$$

- The rate of convergence is exponential, i.e. $\|\boldsymbol{\pi}(t) - \boldsymbol{\pi}^*\| = O(\lambda_2^t)$.
- What happens if the graph is disconnected? The result depends on the graph where the walk starts and there is zero probability of crossing over

Asymptotic properties for stochastic matrices

- Armed with this knowledge, we understand that for the random walk
  1. $\boldsymbol{W}^t = \boldsymbol{\pi}^* \mathbf{1}^\top + O(\lambda_2^t)$
  2. When the graph is connected $\lambda_2 < 1$ and so $O(\lambda_2^t) \to 0$ for $t \gg 1$
- Since $\mathbf{1}^\top \pi(0) = 1$:

$$\lim_{t \to +\infty} \boldsymbol{W}^t \boldsymbol{\pi}(0) = \boldsymbol{\pi}^* \overbrace{\mathbf{1}^\top \boldsymbol{\pi}(0)}^{=1} \equiv \boldsymbol{\pi}^* \quad \text{stationary state probability}$$

- The rate of convergence is exponential, i.e. $\|\boldsymbol{\pi}(t) - \boldsymbol{\pi}^*\| = O(\lambda_2^t)$.
- What happens if the graph is disconnected? The result depends on the graph where the walk starts and there is zero probability of crossing over

Asymptotic properties for stochastic matrices

- Armed with this knowledge, we understand that for the random walk
  1. $\boldsymbol{W}^t = \boldsymbol{\pi}^* \mathbf{1}^\top + O(\lambda_2^t)$
  2. When the graph is connected $\lambda_2 < 1$ and so $O(\lambda_2^t) \to 0$ for $t \gg 1$
- Since $\mathbf{1}^\top \pi(0) = 1$:

$$\lim_{t \to +\infty} \boldsymbol{W}^t \boldsymbol{\pi}(0) = \boldsymbol{\pi}^* \overbrace{\mathbf{1}^\top \boldsymbol{\pi}(0)}^{=1} \equiv \boldsymbol{\pi}^*$$ stationary state probability

- The rate of convergence is exponential, i.e. $\|\boldsymbol{\pi}(t) - \boldsymbol{\pi}^*\| = O(\lambda_2^t)$.
- What happens if the graph is disconnected? The result depends on the graph where the walk starts and there is zero probability of crossing over

Asymptotic properties for stochastic matrices

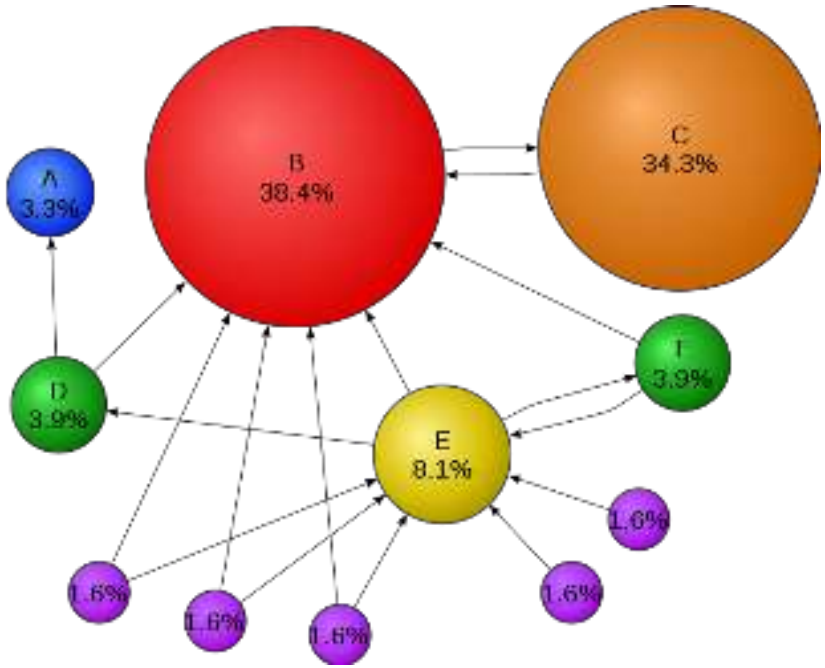
- Armed with this knowledge, we understand that for the random walk
  1. $\boldsymbol{W}^t = \boldsymbol{\pi}^* \mathbf{1}^\top + O(\lambda_2^t)$
  2. When the graph is connected $\lambda_2 < 1$ and so $O(\lambda_2^t) \to 0$ for $t \gg 1$
- Since $\mathbf{1}^\top \pi(0) = 1$:

$$\lim_{t \to +\infty} \boldsymbol{W}^t \boldsymbol{\pi}(0) = \boldsymbol{\pi}^* \overbrace{\mathbf{1}^\top \boldsymbol{\pi}(0)}^{=1} \equiv \boldsymbol{\pi}^*$$ stationary state probability

- The rate of convergence is exponential, i.e. $\|\boldsymbol{\pi}(t) - \boldsymbol{\pi}^*\| = O(\lambda_2^t)$.
- What happens if the graph is disconnected? The result depends on the graph where the walk starts and there is zero probability of crossing over

Random walk based algorithms

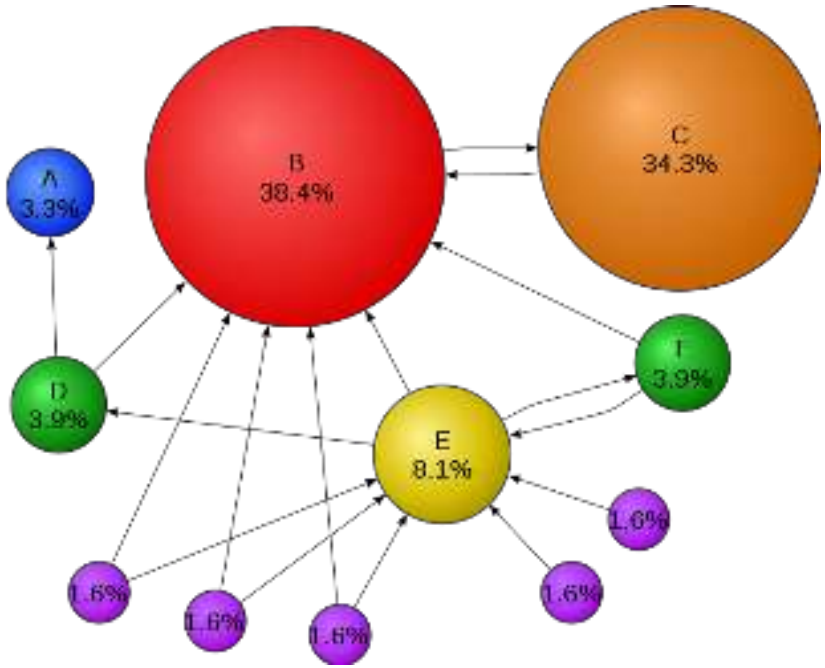
- The random walk model is ubiquitous in both modeling network phenomena as well as devising newtork algorithms
- It is a way to analyze the importance of certain nodes, since they are places the walk visits more often
- Using π^* was the original inspiration of the PageRank algorithm, to sort the importance of the subset of webpages that correspond to a query
- The graph is the webgraph, created by all World Wide Web pages as nodes and hyperlinks



Random walk based algorithms

- The random walk model is ubiquitous in both modeling network phenomena as well as devising newtork algorithms
- It is a way to analyze the importance of certain nodes, since they are places the walk visits more often

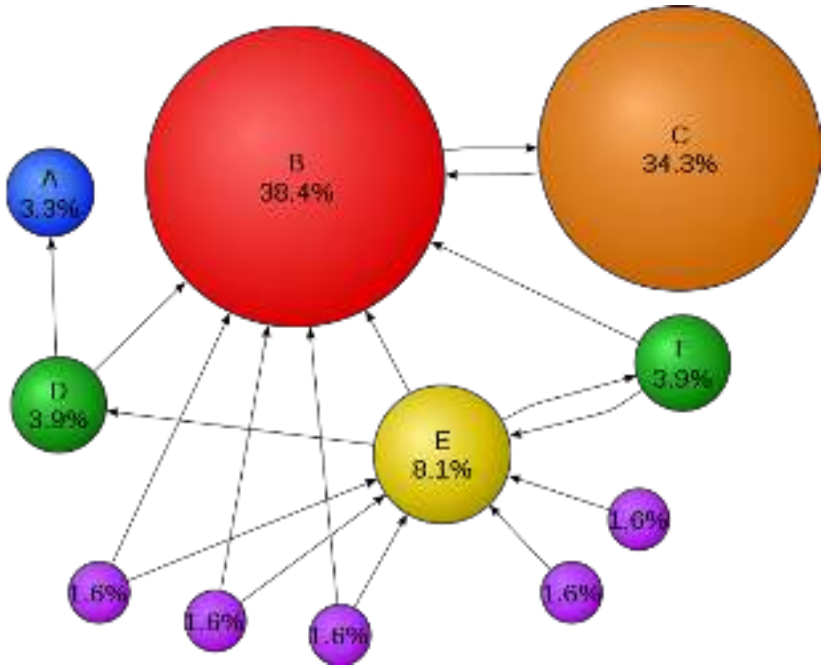
- Using π^* was the original inspiration of the PageRank algorithm, to sort the importance of the subset of webpages that correspond to a query
- The graph is the webgraph, created by all World Wide Web pages as nodes and hyperlinks



Random walk based algorithms

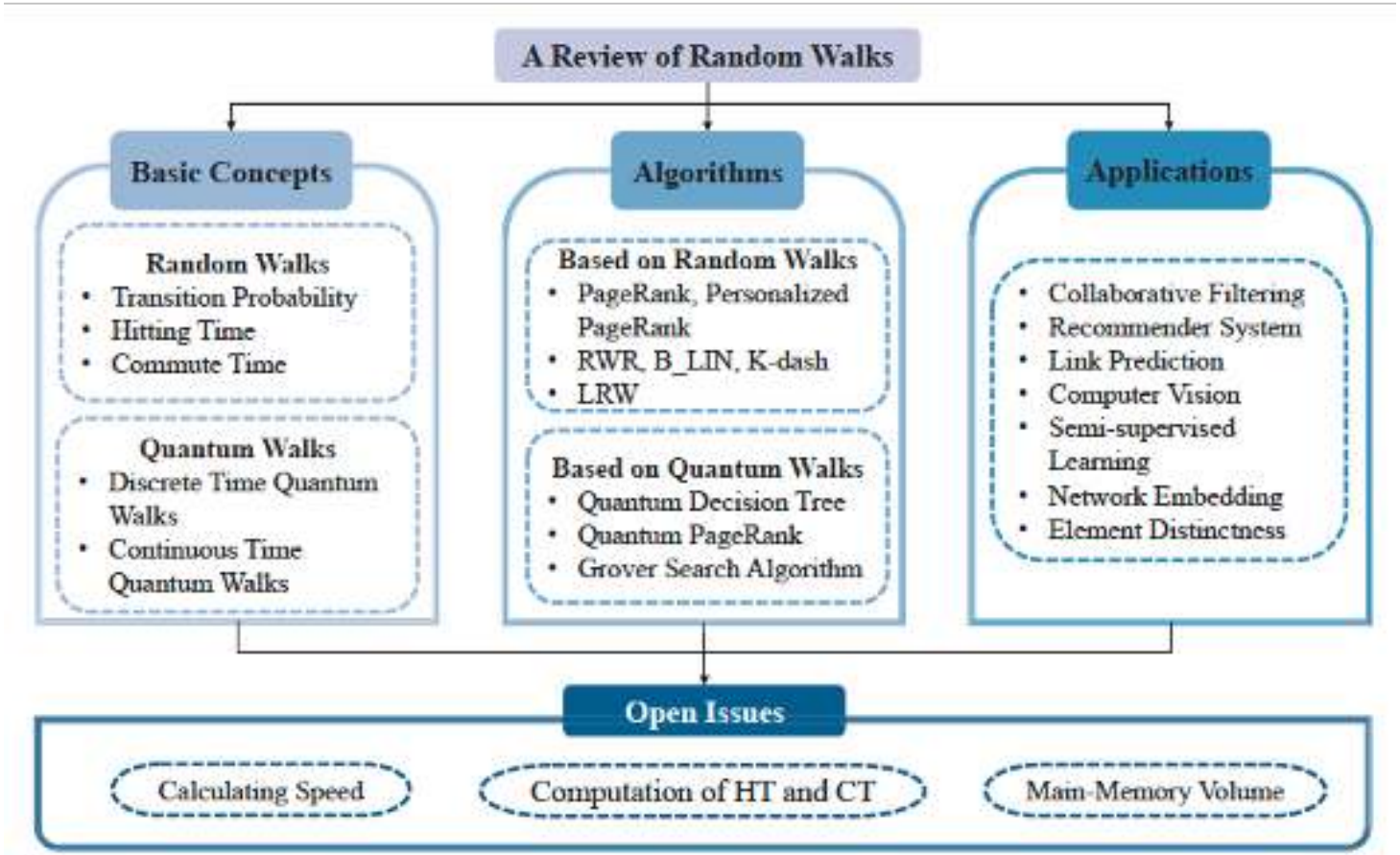
- The random walk model is ubiquitous in both modeling network phenomena as well as devising newtork algorithms
- It is a way to analyze the importance of certain nodes, since they are places the walk visits more often

- Using π^* was the original inspiration of the PageRank algorithm, to sort the importance of the subset of webpages that correspond to a query
- The graph is the webgraph, created by all World Wide Web pages as nodes and hyperlinks



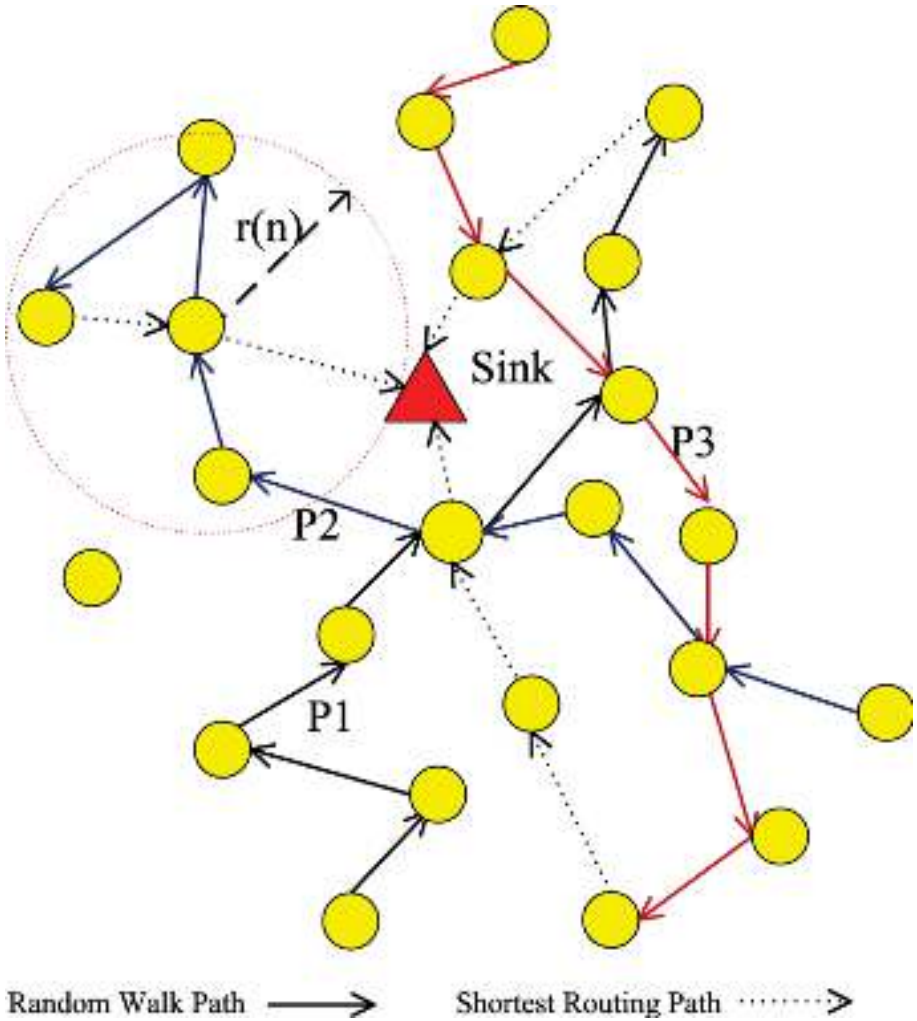
Random walk based algorithms

- A number of variations of PageRank exist, like the Random Walk with Restart (RWR) that we will discuss in the context of GSP or collaborative filtering



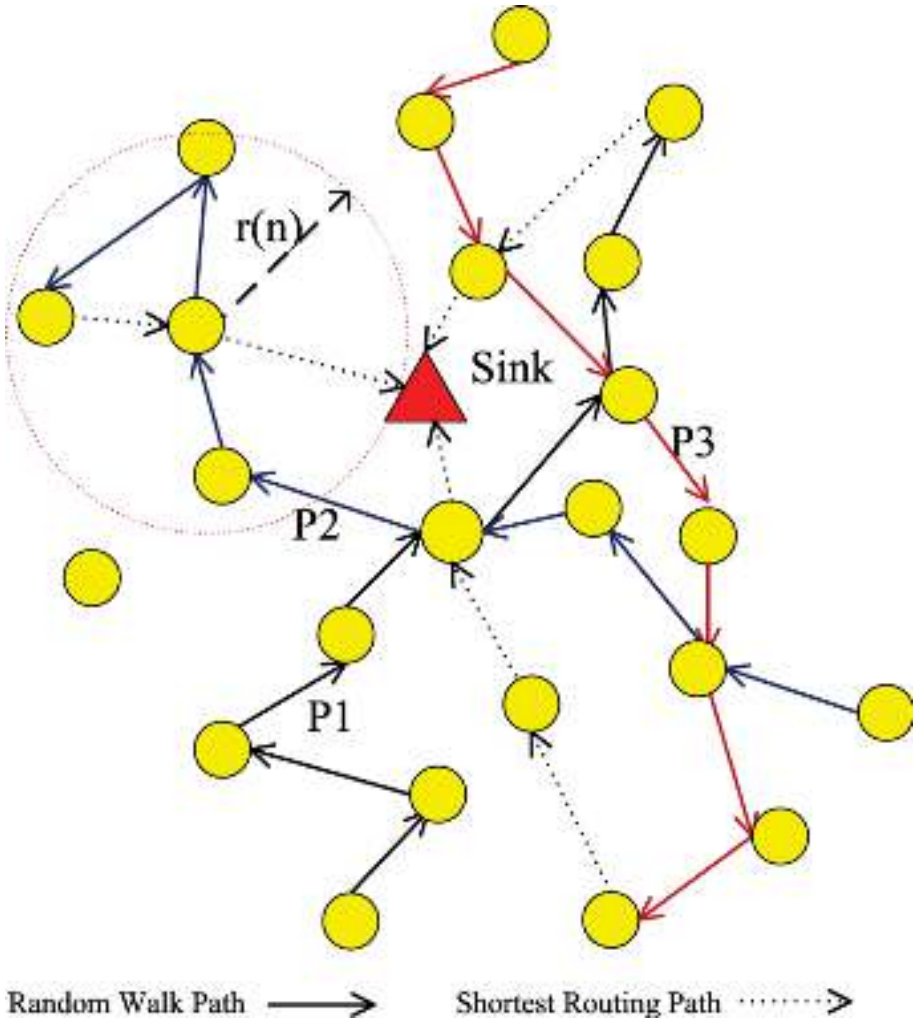
Random walk based algorithms

- Numerous wireless meshed sensor networks routing protocols leverage random walks as a simple routing mechanism to spread information, discover resources and nodes in a network
- It is not necessarily efficient, but it allows to avoid computing the route



Random walk based algorithms

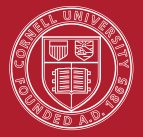
- Numerous wireless meshed sensor networks routing protocols leverage random walks as a simple routing mechanism to spread information, discover resources and nodes in a network
- It is not necessarily efficient, but it allows to avoid computing the route



Random walk models

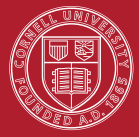
- They model the dynamics of many phenomena. For instance:
  - Epidemics, where what is moving randomly is the virus among people that enter in contact with each other
  - Opinion dynamics, where in a similar vein a belief spreads among people that are part of a social group.
- The analysis of the **consensus dynamics** we will see momentarily are inspired by the analysis of random walks

Random walk models



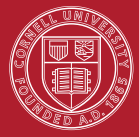
- They model the dynamics of many phenomena. For instance:
  - Epidemics, where what is moving randomly is the virus among people that enter in contact with each other
  - Opinion dynamics, where in a similar vein a belief spreads among people that are part of a social group.
- The analysis of the **consensus dynamics** we will see momentarily are inspired by the analysis of random walks

Outline



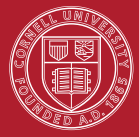
Cornell University

Outline



Cornell University

Average consensus dynamics



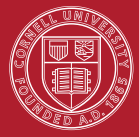
Cornell University

- Let us consider a strongly conencted network, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with Laplacian $\boldsymbol{L}$
- Each node has a certain initial value $x_v(0)$
- The goal is to compute the average:

$$x^{\text{ave}} = \frac{1}{N} \sum_{v=1}^{N} x_v(0)$$

- In other words, the nodes want to achieve what is called average consensus

Average consensus dynamics



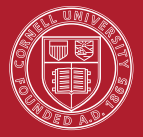
Cornell University

- Let us consider a strongly conencted network, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with Laplacian $\boldsymbol{L}$
- Each node has a certain initial value $x_v(0)$
- The goal is to compute the average:

$$x^{\text{ave}} = \frac{1}{N} \sum_{v=1}^{N} x_v(0)$$

- In other words, the nodes want to achieve what is called average consensus

Average consensus dynamics



Cornell University

- Let us consider a strongly conencted network, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with Laplacian $\boldsymbol{L}$
- Each node has a certain initial value $x_v(0)$
- The goal is to compute the average:

$$x^{\text{ave}} = \frac{1}{N} \sum_{v=1}^{N} x_v(0)$$

- In other words, the nodes want to achieve what is called average consensus

Average consensus dynamics computation

- This is a useful primitive to compute cooperatively something in this form:

$$\text{evaluate} \left(\sum_{v \in \mathcal{V}} \text{expression}_v \right)$$

where expression_v is some expression (whose value is $x_v(0)$) as a function of local (private) data at each node $v \in \mathcal{V}$

Laplacian matrix and consensus

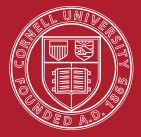
Cornell University

- We said that the Laplacian is the discrete version of the Laplacian for continuous fields ∇^2.
- The following quadratic form is a **measure of the variation of the graph signal** and is also referred to as Laplacian potential:

$$V(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x} = \sum_{uv \in \mathcal{E}} w_{uv}(x_u - x_v)^2 \geq 0$$

- The closer this value is to zero, the closer are the elements of $\boldsymbol{x}$ at adjacent vertices are. $V(\boldsymbol{x})$ is a measure of the *variation* of the graph signal $\boldsymbol{x}$.

Laplacian matrix and consensus



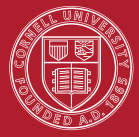
Cornell University

- We said that the Laplacian is the discrete version of the Laplacian for continuous fields ∇^2.
- The following quadratic form is a **measure of the variation of the graph signal** and is also referred to as Laplacian potential:

$$V(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x} = \sum_{uv \in \mathcal{E}} w_{uv}(x_u - x_v)^2 \geq 0$$

- The closer this value is to zero, the closer are the elements of $\boldsymbol{x}$ at adjacent vertices are. $V(\boldsymbol{x})$ is a measure of the *variation* of the graph signal $\boldsymbol{x}$.

Laplacian matrix and consensus



Cornell University

- We said that the Laplacian is the discrete version of the Laplacian for continuous fields ∇^2.
- The following quadratic form is a **measure of the variation of the graph signal** and is also referred to as Laplacian potential:

$$V(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x} = \sum_{uv \in \mathcal{E}} w_{uv}(x_u - x_v)^2 \geq 0$$

- The closer this value is to zero, the closer are the elements of $\boldsymbol{x}$ at adjacent vertices are. $V(\boldsymbol{x})$ is a measure of the *variation* of the graph signal $\boldsymbol{x}$.

Laplacian matrix and consensus

Cornell University

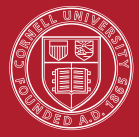
- We know that if $\boldsymbol{x} = \mathbf{1}$:

$$\mathbf{1}^\top \boldsymbol{L}\mathbf{1} = \mathbf{1}^\top = \mathbf{0} = 0 \quad \Rightarrow \quad \min_{\boldsymbol{x}} \boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x} = 0$$

- For each connected graph every $\boldsymbol{x}$ that is not proportional to $\mathbf{1}$ will have $\boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x} > 0$.
- A collaborative minimization of the Laplacian potential should therefore be able to lead to consensus:

$$\min_{\boldsymbol{x}} V(\boldsymbol{x}) = \min_{\boldsymbol{x}} V(\boldsymbol{x})$$

Laplacian matrix and consensus



Cornell University

- We know that if $\boldsymbol{x} = \mathbf{1}$:

$$\mathbf{1}^\top \boldsymbol{L}\mathbf{1} = \mathbf{1}^\top \mathbf{0} = 0 \quad \Rightarrow \quad \min_{\boldsymbol{x}} \boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x} = 0$$

- For each connected graph every $\boldsymbol{x}$ that is not proportional to $\mathbf{1}$ will have $\boldsymbol{x}^\top \boldsymbol{L}\boldsymbol{x} > 0$.
- A collaborative minimization of the Laplacian potential should therefore be able to lead to consensus:

$$\min_{\boldsymbol{x}} V(\boldsymbol{x}) = \min_{\boldsymbol{x}} V(\boldsymbol{x})$$

Laplacian matrix and consensus

Cornell University

- The gradient descent (with step size $0 < \alpha < 1$) on this cost is:

$$\boldsymbol{x}(t+1) = \boldsymbol{x}(t) - \alpha \overbrace{\boldsymbol{L}\boldsymbol{x}(t)}^{\nabla_{\boldsymbol{x}} V(\boldsymbol{x})} = \overbrace{(\boldsymbol{I} - \alpha \boldsymbol{L})}^{\boldsymbol{W}} \boldsymbol{x}(t)$$

$$x_v(t+1) = \sum_{u \in \mathcal{N}_v} [\boldsymbol{W}]_{vu} x_u(t) = \sum_{u \in \mathcal{N}_v} w_{vu} x_u(t)$$

- In the computation of $x_v(t+1) = \sum_{u \in \mathcal{N}_v} w_{vu} x_u(t)$ node v only needs to communicate to its neighbors in $\mathcal{N}_v$
- How can we ensure that the algorithm converges to the average?

Laplacian matrix and consensus

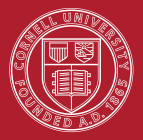
- The gradient descent (with step size $0 < \alpha < 1$) on this cost is:

$$\boldsymbol{x}(t+1) = \boldsymbol{x}(t) - \alpha \overbrace{\boldsymbol{L}\boldsymbol{x}(t)}^{\nabla_{\boldsymbol{x}} V(\boldsymbol{x})} = \overbrace{(\boldsymbol{I} - \alpha \boldsymbol{L})}^{\boldsymbol{W}} \boldsymbol{x}(t)$$

$$x_v(t+1) = \sum_{u \in \mathcal{N}_v} [\boldsymbol{W}]_{vu} x_u(t) = \sum_{u \in \mathcal{N}_v} w_{vu} x_u(t)$$

- In the computation of $x_v(t+1) = \sum_{u \in \mathcal{N}_v} w_{vu} x_u(t)$ node v only needs to communicate to its neighbors in $\mathcal{N}_v$
- How can we ensure that the algorithm converges to the average?

Laplacian matrix and consensus



Cornell University

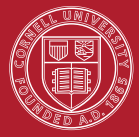
- The gradient descent (with step size $0 < \alpha < 1$) on this cost is:

$$\boldsymbol{x}(t+1) = \boldsymbol{x}(t) - \alpha \overbrace{\boldsymbol{L}\boldsymbol{x}(t)}^{\nabla_{\boldsymbol{x}} V(\boldsymbol{x})} = \overbrace{(\boldsymbol{I} - \alpha \boldsymbol{L})}^{\boldsymbol{W}} \boldsymbol{x}(t)$$

$$x_v(t+1) = \sum_{u \in \mathcal{N}_v} [\boldsymbol{W}]_{vu} x_u(t) = \sum_{u \in \mathcal{N}_v} w_{vu} x_u(t)$$

- In the computation of $x_v(t+1) = \sum_{u \in \mathcal{N}_v} w_{vu} x_u(t)$ node v only needs to communicate to its neighbors in $\mathcal{N}_v$
- How can we ensure that the algorithm converges to the average?

Laplacian matrix and consensus



Cornell University

- We can leverage the theory we developed for the random walk
- If $\boldsymbol{W} = (\boldsymbol{I} - \alpha \boldsymbol{L})$ is:
  1. Doubly stochastic: $\boldsymbol{W} = \boldsymbol{W}^\top$, and $\boldsymbol{W}\mathbf{1} = \mathbf{1}$
  2. Spectral radius: $0 < \lambda_2(\boldsymbol{W}) < 1$
- The first is true because

$$\boldsymbol{L}^\top = \boldsymbol{L}, \quad \text{and} \quad \boldsymbol{W}\mathbf{1} = (\boldsymbol{I} - \alpha \boldsymbol{L})\mathbf{1} = \mathbf{1} - \alpha \overbrace{\boldsymbol{L}\mathbf{1}}^{=0} = \mathbf{1}$$

- We know that a positive spectral radius. In fact:
  - Since $\lambda_2(\boldsymbol{W}) = 1 - \alpha \lambda_{N-1}$ the condition $0 < \lambda_2(\boldsymbol{W}) < 1$ implies $0 < \lambda_{N-1}(\boldsymbol{L}) < 1/\alpha$.
- Note that the algebraic connectivity $\lambda^{\text{conn}} = \lambda_{N-1}(\boldsymbol{L}) > 0$ if and only if the graph is strongly connected.

Laplacian matrix and consensus

- We can leverage the theory we developed for the random walk
- If $\boldsymbol{W} = (\boldsymbol{I} - \alpha \boldsymbol{L})$ is:
  1. Doubly stochastic: $\boldsymbol{W} = \boldsymbol{W}^\top$, and $\boldsymbol{W}\mathbf{1} = \mathbf{1}$
  2. Spectral radius: $0 < \lambda_2(\boldsymbol{W}) < 1$
- The first is true because

$$\boldsymbol{L}^\top = \boldsymbol{L}, \quad \text{and} \quad \boldsymbol{W}\mathbf{1} = (\boldsymbol{I} - \alpha \boldsymbol{L})\mathbf{1} = \mathbf{1} - \alpha \overbrace{\boldsymbol{L}\mathbf{1}}^{=0} = \mathbf{1}$$

- We know that a positive spectral radius. In fact:
  - Since $\lambda_2(\boldsymbol{W}) = 1 - \alpha \lambda_{N-1}$ the condition $0 < \lambda_2(\boldsymbol{W}) < 1$ implies $0 < \lambda_{N-1}(\boldsymbol{L}) < 1/\alpha$.
- Note that the algebraic connectivity $\lambda^{\text{conn}} = \lambda_{N-1}(\boldsymbol{L}) > 0$ if and only if the graph is strongly connected.

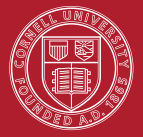
Laplacian matrix and consensus

- We can leverage the theory we developed for the random walk
- If $\boldsymbol{W} = (\boldsymbol{I} - \alpha \boldsymbol{L})$ is:
  1. Doubly stochastic: $\boldsymbol{W} = \boldsymbol{W}^\top$, and $\boldsymbol{W}\mathbf{1} = \mathbf{1}$
  2. Spectral radius: $0 < \lambda_2(\boldsymbol{W}) < 1$
- The first is true because

$$\boldsymbol{L}^\top = \boldsymbol{L}, \quad \text{and} \quad \boldsymbol{W}\mathbf{1} = (\boldsymbol{I} - \alpha \boldsymbol{L})\mathbf{1} = \mathbf{1} - \alpha \overbrace{\boldsymbol{L}\mathbf{1}}^{=0} = \mathbf{1}$$

- We know that a positive spectral radius. In fact:
  - Since $\lambda_2(\boldsymbol{W}) = 1 - \alpha\lambda_{N-1}$ the condition $0 < \lambda_2(\boldsymbol{W}) < 1$ implies $0 < \lambda_{N-1}(\boldsymbol{L}) < 1/\alpha$.
- Note that the algebraic connectivity $\lambda^{\text{conn}} = \lambda_{N-1}(\boldsymbol{L}) > 0$ if and only if the graph is strongly connected.

Laplacian matrix and consensus



Cornell University

- We can leverage the theory we developed for the random walk
- If $\boldsymbol{W} = (\boldsymbol{I} - \alpha \boldsymbol{L})$ is:
  1. Doubly stochastic: $\boldsymbol{W} = \boldsymbol{W}^\top$, and $\boldsymbol{W}\mathbf{1} = \mathbf{1}$
  2. Spectral radius: $0 < \lambda_2(\boldsymbol{W}) < 1$
- The first is true because

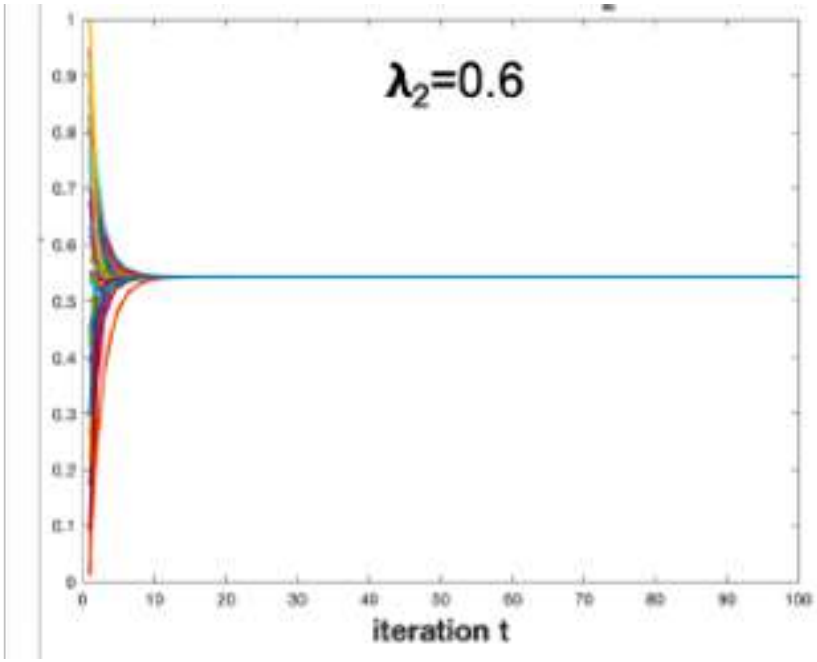
$$\boldsymbol{L}^\top = \boldsymbol{L}, \quad \text{and} \quad \boldsymbol{W}\mathbf{1} = (\boldsymbol{I} - \alpha \boldsymbol{L})\mathbf{1} = \mathbf{1} - \alpha \overbrace{\boldsymbol{L}\mathbf{1}}^{=0} = \mathbf{1}$$

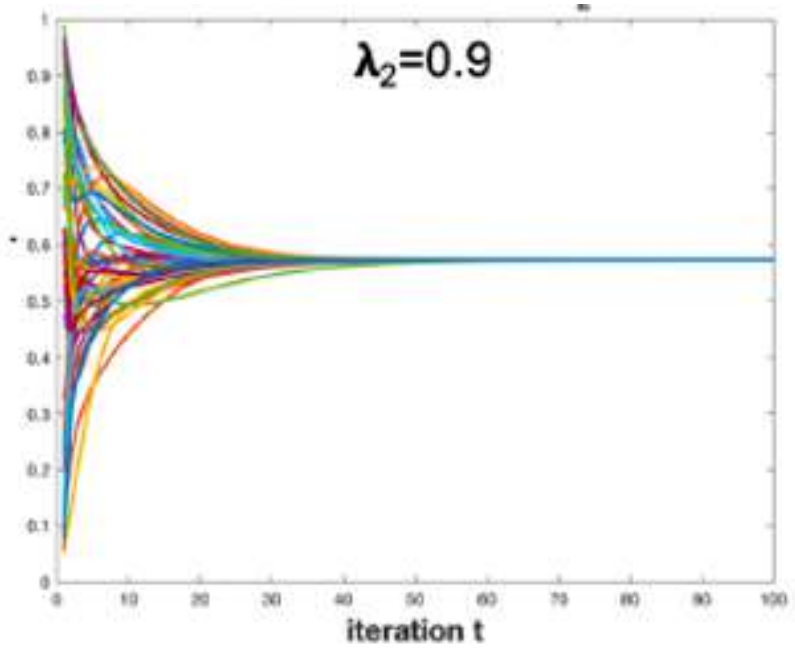
- We know that a positive spectral radius. In fact:
  - Since $\lambda_2(\boldsymbol{W}) = 1 - \alpha\lambda_{N-1}$ the condition $0 < \lambda_2(\boldsymbol{W}) < 1$ implies $0 < \lambda_{N-1}(\boldsymbol{L}) < 1/\alpha$.
- Note that the algebraic connectivity $\lambda^{\text{conn}} = \lambda_{N-1}(\boldsymbol{L}) > 0$ if and only if the graph is strongly connected.

Consensus dynamics

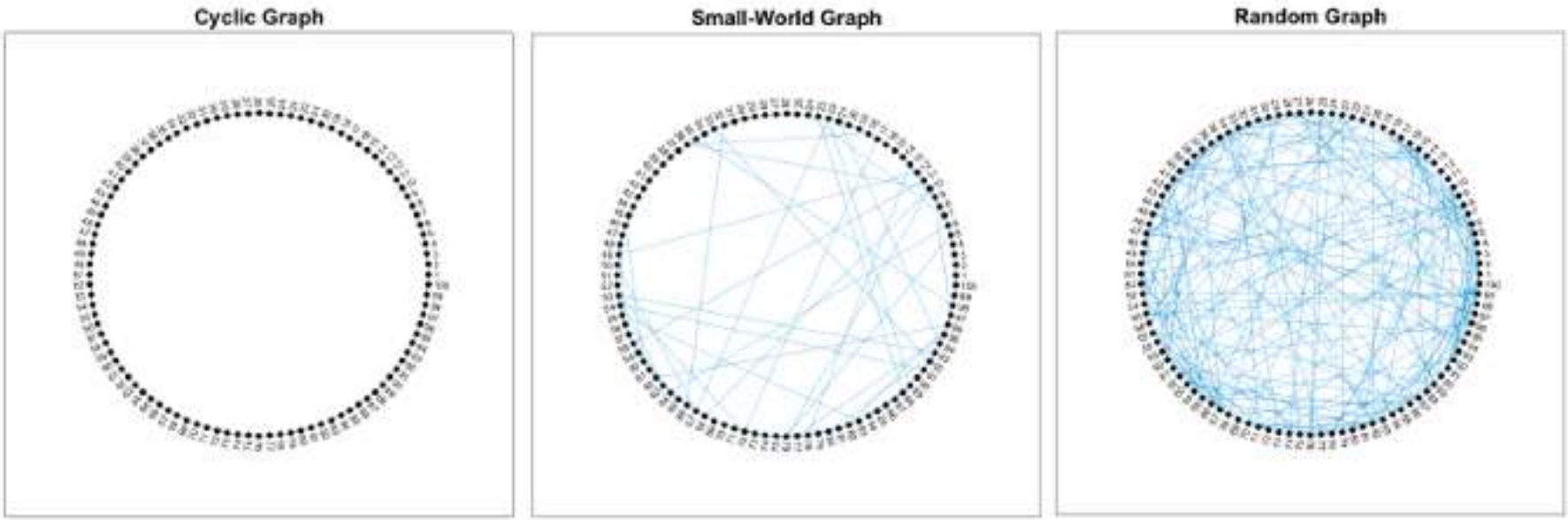
Cornell University

- The convergence rate is exponential w.r.t. $\lambda_2(\boldsymbol{W}) = 1 - \alpha\lambda_{\text{conn.}}$, that is $\|x(t) - x^{\text{ave}}\| \leq O(\lambda_2^t(\boldsymbol{W}))$





Consensus and connectivity

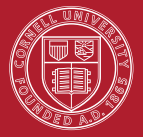


Examples of graph topology

Left: a cyclic graph with $S = 100$ nodes and node degree 4; Middle: a small-world graph with $S = 100$, average degree 4 and re-wiring probability 0.2. Right: a random graph with $S = 100$ nodes and node degree 4.

Guess which one has the largest $\lambda_{\text{conn.}}$?

Consensus dynamics and random walks



Cornell University

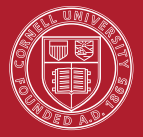
- It should not go unnoticed that the type of equations used to study both are analogous:

$$\text{Random walk state probabilities} \quad \boldsymbol{\pi}(t+1) = \boldsymbol{W}\boldsymbol{\pi}(t)$$

$$\text{Consensus dynamics} \quad \boldsymbol{x}(t+1) = \boldsymbol{W}\boldsymbol{x}(t)$$

- In both cases, $\boldsymbol{W}$ are stochastic matrices and analogous properties, except for average consensus we want $\boldsymbol{W}^\top = \boldsymbol{W}$ (doubly stochastic matrix)

Consensus dynamics and random walks



Cornell University

- It should not go unnoticed that the type of equations used to study both are analogous:

$$\text{Random walk state probabilities} \quad \boldsymbol{\pi}(t+1) = \boldsymbol{W}\boldsymbol{\pi}(t)$$

$$\text{Consensus dynamics} \quad \boldsymbol{x}(t+1) = \boldsymbol{W}\boldsymbol{x}(t)$$

- In both cases, $\boldsymbol{W}$ are stochastic matrices and analogous properties, except for average consensus we want $\boldsymbol{W}^\top = \boldsymbol{W}$ (doubly stochastic matrix)

Random walk as a graph signal

Cornell University

- We can associate to the random walk a graph signal which indicates what node is currently visited:

$$[\boldsymbol{x}^{\text{rw}}(t)]_v = \begin{cases} 1 & \text{if the node visited is } v \\ 0 & \text{else} \end{cases}$$

- Note that the consensus dynamics are analogous to those of the expected values of the random walk signal

$$\boldsymbol{\pi}(t) = \mathbb{E}[\boldsymbol{x}^{\text{rw}}(t)], \quad \boldsymbol{\pi}(t+1) = \mathbb{E}[\boldsymbol{x}^{\text{rw}}(t+1)] = \boldsymbol{W} \overbrace{\mathbb{E}[\boldsymbol{x}^{\text{rw}}(t)]}^{\boldsymbol{\pi}(t)}$$

Random walk as a graph signal

Cornell University

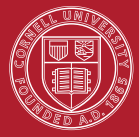
- We can associate to the random walk a graph signal which indicates what node is currently visited:

$$[\boldsymbol{x}^{\text{rw}}(t)]_v = \begin{cases} 1 & \text{if the node visited is } v \\ 0 & \text{else} \end{cases}$$

- Note that the consensus dynamics are analogous to those of the expected values of the random walk signal

$$\boldsymbol{\pi}(t) = \mathbb{E}[\boldsymbol{x}^{\text{rw}}(t)], \quad \boldsymbol{\pi}(t+1) = \mathbb{E}[\boldsymbol{x}^{\text{rw}}(t+1)] = \boldsymbol{W} \overbrace{\mathbb{E}[\boldsymbol{x}^{\text{rw}}(t)]}^{\boldsymbol{\pi}(t)}$$

Outline



Cornell University

Randomized interaction

- There is another type of randomness that occurs in networks models: the node interaction is random
- This is captured throgh a random matrix $\boldsymbol{W}(t)$
- Although it can be more general, $\boldsymbol{W}(t)$ often represents the random on and off switching of links
- The process goes also by the name of *random gossiping*:

$$\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{x}(t)$$

- Assuming the network starts at $\boldsymbol{x}(0)$, the state at time $t+1$ is:

$$\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{W}(t-1)\dots\boldsymbol{W}(1)\boldsymbol{W}(0)\boldsymbol{x}(0)$$

Randomized interaction

Cornell University

- There is another type of randomness that occurs in networks models: the node interaction is random
- This is captured throgh a random matrix $\boldsymbol{W}(t)$
- Although it can be more general, $\boldsymbol{W}(t)$ often represents the random on and off switching of links
- The process goes also by the name of *random gossiping*:

$$\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{x}(t)$$

- Assuming the network starts at $\boldsymbol{x}(0)$, the state at time $t+1$ is:

$$\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{W}(t-1)\dots\boldsymbol{W}(1)\boldsymbol{W}(0)\boldsymbol{x}(0)$$

Randomized interaction

- There is another type of randomness that occurs in networks models: the node interaction is random
- This is captured throgh a random matrix $\boldsymbol{W}(t)$
- Although it can be more general, $\boldsymbol{W}(t)$ often represents the random on and off switching of links
- The process goes also by the name of *random gossiping*:

$$\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{x}(t)$$

- Assuming the network starts at $\boldsymbol{x}(0)$, the state at time $t+1$ is:

$$\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{W}(t-1)\dots\boldsymbol{W}(1)\boldsymbol{W}(0)\boldsymbol{x}(0)$$

Randomized interaction

Cornell University

- There is another type of randomness that occurs in networks models: the node interaction is random
- This is captured throgh a random matrix $\boldsymbol{W}(t)$
- Although it can be more general, $\boldsymbol{W}(t)$ often represents the random on and off switching of links
- The process goes also by the name of *random gossiping*:

$$\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{x}(t)$$

- Assuming the network starts at $\boldsymbol{x}(0)$, the state at time $t+1$ is:

$$\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{W}(t-1)\dots\boldsymbol{W}(1)\boldsymbol{W}(0)\boldsymbol{x}(0)$$

Randomized interaction

- There is another type of randomness that occurs in networks models: the node interaction is random
- This is captured throgh a random matrix $\boldsymbol{W}(t)$
- Although it can be more general, $\boldsymbol{W}(t)$ often represents the random on and off switching of links
- The process goes also by the name of *random gossiping*:

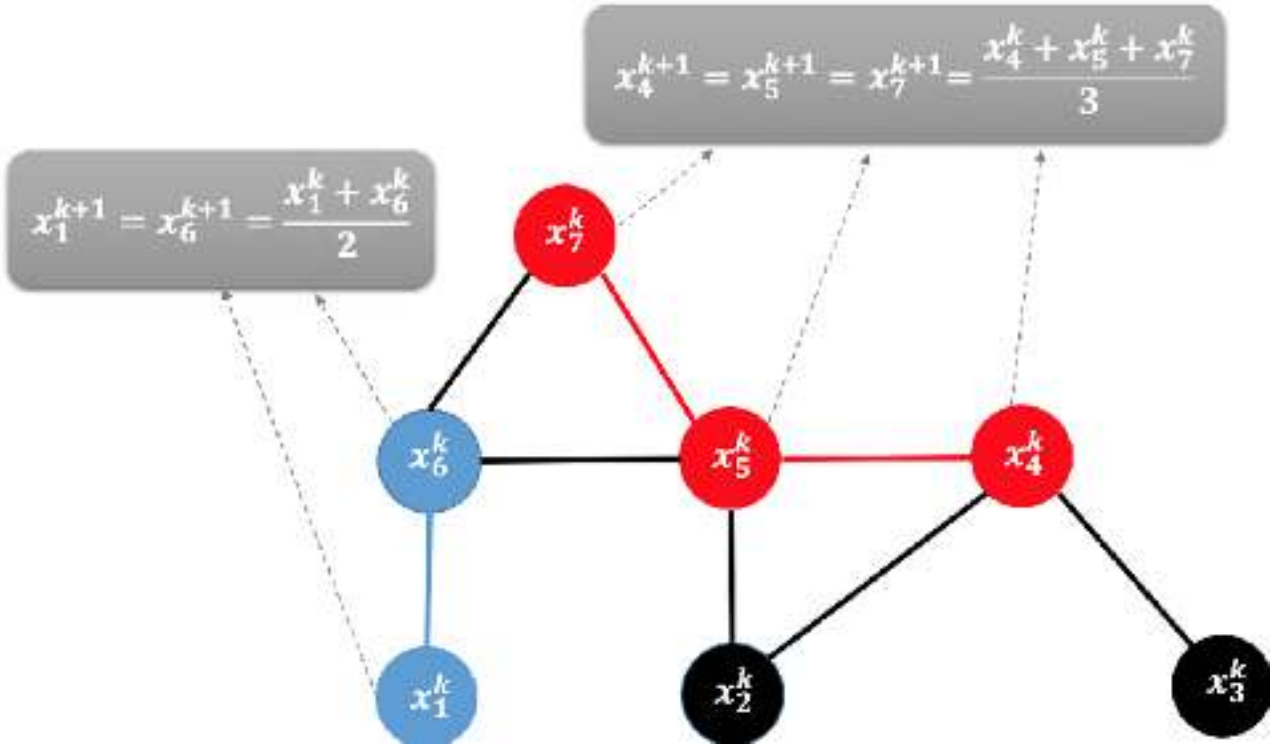
$$\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{x}(t)$$

- Assuming the network starts at $\boldsymbol{x}(0)$, the state at time $t+1$ is:

$$\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{W}(t-1)\dots\boldsymbol{W}(1)\boldsymbol{W}(0)\boldsymbol{x}(0)$$

Randomized "Gossiping" protocol

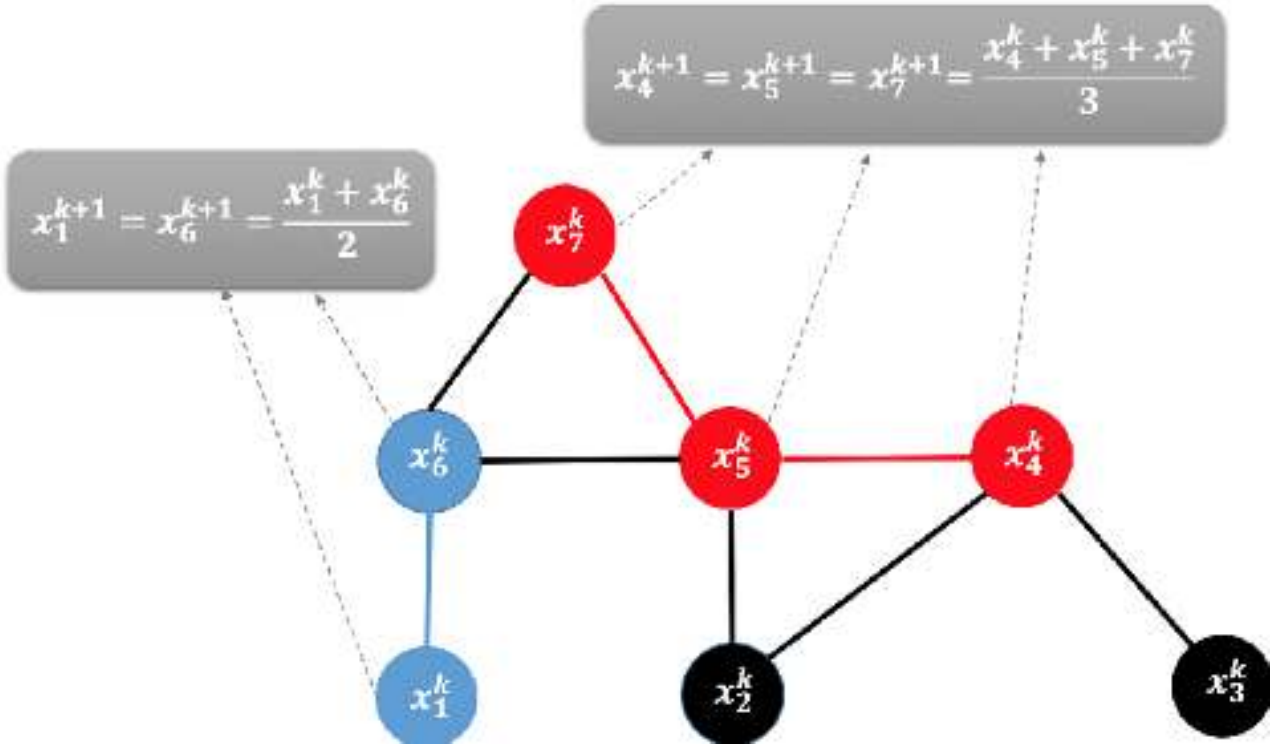
- The assumption is that there is an underlying graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the set $\mathcal{E}$ of *possible edges*, i.e. $w_{uv}(t)$ can be $w_{uv}(t) > 0$ if and only if $(u, v) \in \mathcal{V}$



- The $\boldsymbol{W}(t) = \boldsymbol{I} - \alpha \boldsymbol{L}(t)$ are very sparse, but for every t it holds $\mathbf{1}^\top \boldsymbol{W}(t) = \mathbf{1}^\top \Rightarrow$ the average of the network is conserved $\forall t \; x^{\text{ave}} = \frac{1}{N} \mathbf{1}^\top \boldsymbol{x}(t)$

Randomized "Gossiping" protocol

- The assumption is that there is an underlying graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the set $\mathcal{E}$ of *possible edges*, i.e. $w_{uv}(t)$ can be $w_{uv}(t) > 0$ if and only if $(u, v) \in \mathcal{V}$



- The $\boldsymbol{W}(t) = \boldsymbol{I} - \alpha \boldsymbol{L}(t)$ are very sparse, but for every t it holds $\mathbf{1}^\top \boldsymbol{W}(t) = \mathbf{1}^\top \Rightarrow$ the average of the network is conserved $\forall t \; x^{\text{ave}} = \frac{1}{N} \mathbf{1}^\top \boldsymbol{x}(t)$

Randomized "Gossiping" protocol

Cornell University

- An example of random interaction is that a pair of nodes (i_t, j_t) *wakes up* at time t and averages their values:

$$[\boldsymbol{x}(t+1)]_v = \begin{cases} \frac{x_{i_t}(t)+x_{j_t}(t)}{2} & i_t = v \text{ or } j_t = v \\ x_v(t) & \text{else} \end{cases}$$

- Let $\boldsymbol{e}_i$ be the ith coordinate vector. In this case each $\boldsymbol{W}(t)$ is a doubly stochastic matrix, and if (i, j) is the node pair that is activated at time t then equal to:

$$\boldsymbol{W}(t) = \boldsymbol{I} - \overbrace{\frac{1}{2}\left((\boldsymbol{e}_{j_t} - \boldsymbol{e}_{i_t})\boldsymbol{e}_{j_t}^\top + (\boldsymbol{e}_{i_t} - \boldsymbol{e}_{j_t})\boldsymbol{e}_{i_t}^\top\right)}^{\frac{1}{2}\boldsymbol{L}(t)}$$

- This can be generalized to the case where, for an $0 < \alpha < 1$

$$\boldsymbol{W}(t) = \boldsymbol{I} - \left(\alpha(\boldsymbol{e}_{j_t} - \boldsymbol{e}_{i_t})\boldsymbol{e}_{j_t}^\top + (1-\alpha)(\boldsymbol{e}_{i_t} - \boldsymbol{e}_{j_t})\boldsymbol{e}_{i_t}^\top\right)$$

Randomized "Gossiping" protocol

- An example of random interaction is that a pair of nodes (i_t, j_t) *wakes up* at time t and averages their values:

$$[\boldsymbol{x}(t+1)]_v = \begin{cases} \frac{x_{i_t}(t)+x_{j_t}(t)}{2} & i_t = v \text{ or } j_t = v \\ x_v(t) & \text{else} \end{cases}$$

- Let $\boldsymbol{e}_i$ be the ith coordinate vector. In this case each $\boldsymbol{W}(t)$ is a doubly stochastic matrix, and if (i, j) is the node pair that is activated at time t then equal to:

$$\boldsymbol{W}(t) = \boldsymbol{I} - \overbrace{\frac{1}{2}\left((\boldsymbol{e}_{j_t} - \boldsymbol{e}_{i_t})\boldsymbol{e}_{j_t}^\top + (\boldsymbol{e}_{i_t} - \boldsymbol{e}_{j_t})\boldsymbol{e}_{i_t}^\top\right)}^{\frac{1}{2}\boldsymbol{L}(t)}$$

- This can be generalized to the case where, for an $0 < \alpha < 1$

$$\boldsymbol{W}(t) = \boldsymbol{I} - \left(\alpha(\boldsymbol{e}_{j_t} - \boldsymbol{e}_{i_t})\boldsymbol{e}_{j_t}^\top + (1-\alpha)(\boldsymbol{e}_{i_t} - \boldsymbol{e}_{j_t})\boldsymbol{e}_{i_t}^\top\right)$$

Randomized "Gossiping" protocol

- An example of random interaction is that a pair of nodes (i_t, j_t) *wakes up* at time t and averages their values:

$$[\boldsymbol{x}(t+1)]_v = \begin{cases} \frac{x_{i_t}(t)+x_{j_t}(t)}{2} & i_t = v \text{ or } j_t = v \\ x_v(t) & \text{else} \end{cases}$$

- Let $\boldsymbol{e}_i$ be the ith coordinate vector. In this case each $\boldsymbol{W}(t)$ is a doubly stochastic matrix, and if (i, j) is the node pair that is activated at time t then equal to:

$$\boldsymbol{W}(t) = \boldsymbol{I} - \overbrace{\frac{1}{2}\left((\boldsymbol{e}_{j_t} - \boldsymbol{e}_{i_t})\boldsymbol{e}_{j_t}^\top + (\boldsymbol{e}_{i_t} - \boldsymbol{e}_{j_t})\boldsymbol{e}_{i_t}^\top\right)}^{\frac{1}{2}\boldsymbol{L}(t)}$$

- This can be generalized to the case where, for an $0 < \alpha < 1$

$$\boldsymbol{W}(t) = \boldsymbol{I} - \left(\alpha(\boldsymbol{e}_{j_t} - \boldsymbol{e}_{i_t})\boldsymbol{e}_{j_t}^\top + (1-\alpha)(\boldsymbol{e}_{i_t} - \boldsymbol{e}_{j_t})\boldsymbol{e}_{i_t}^\top\right)$$

The update index t versus time

Cornell University

- Note that t represents an **update index** which is not necessarily a **time** value
- In fact, the time elapsing between updates can be anything, does not have to be regular
- There can be an underlying process of *inter-update* times τ_t between any group that averages, so that the tth update occurs at time $\sum_{k=0}^{t} \tau_k$
- The average time it takes to complete t updates is therefore $\sum_{k=0}^{t} \mathbb{E}[\tau_k]$

The update index t versus time

Cornell University

- Note that t represents an **update index** which is not necessarily a **time** value
- In fact, the time elapsing between updates can be anything, does not have to be regular
- There can be an underlying process of *inter-update* times τ_t between any group that averages, so that the tth update occurs at time $\sum_{k=0}^{t} \tau_k$
- The average time it takes to complete t updates is therefore $\sum_{k=0}^{t} \mathbb{E}[\tau_k]$

The update index t versus time

Cornell University

- Note that t represents an **update index** which is not necessarily a **time** value
- In fact, the time elapsing between updates can be anything, does not have to be regular
- There can be an underlying process of *inter-update* times τ_t between any group that averages, so that the tth update occurs at time $\sum_{k=0}^{t} \tau_k$
- The average time it takes to complete t updates is therefore $\sum_{k=0}^{t} \mathbb{E}[\tau_k]$

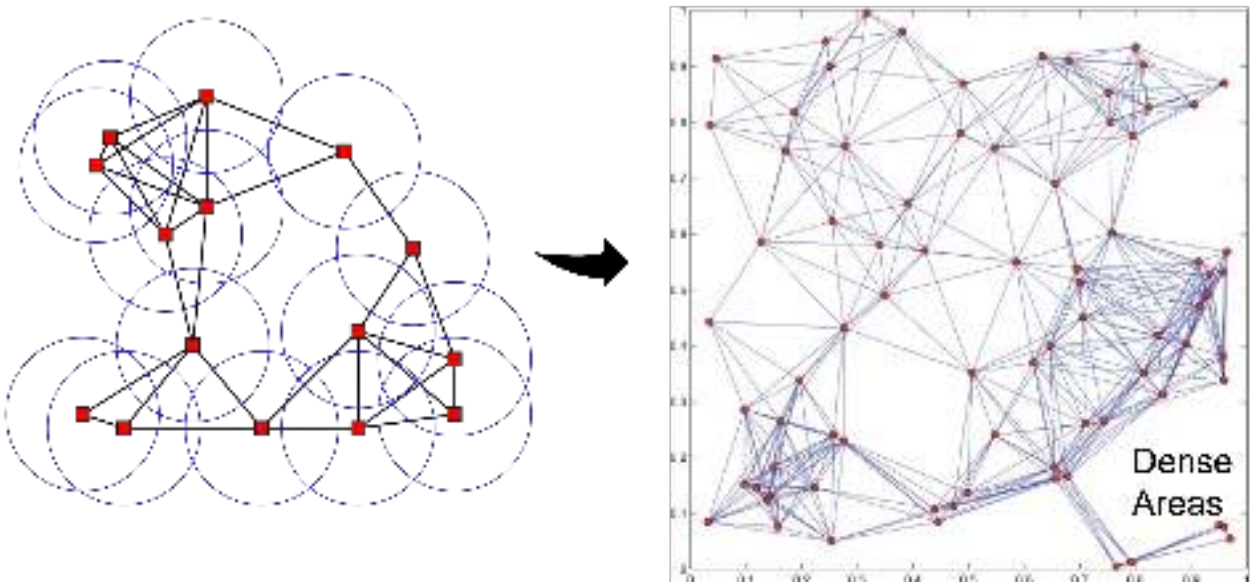
The update index t versus time

Cornell University

- Note that t represents an **update index** which is not necessarily a **time** value
- In fact, the time elapsing between updates can be anything, does not have to be regular
- There can be an underlying process of *inter-update* times τ_t between any group that averages, so that the tth update occurs at time $\sum_{k=0}^{t} \tau_k$
- The average time it takes to complete t updates is therefore $\sum_{k=0}^{t} \mathbb{E}[\tau_k]$

Wireless pairwise random gossiping

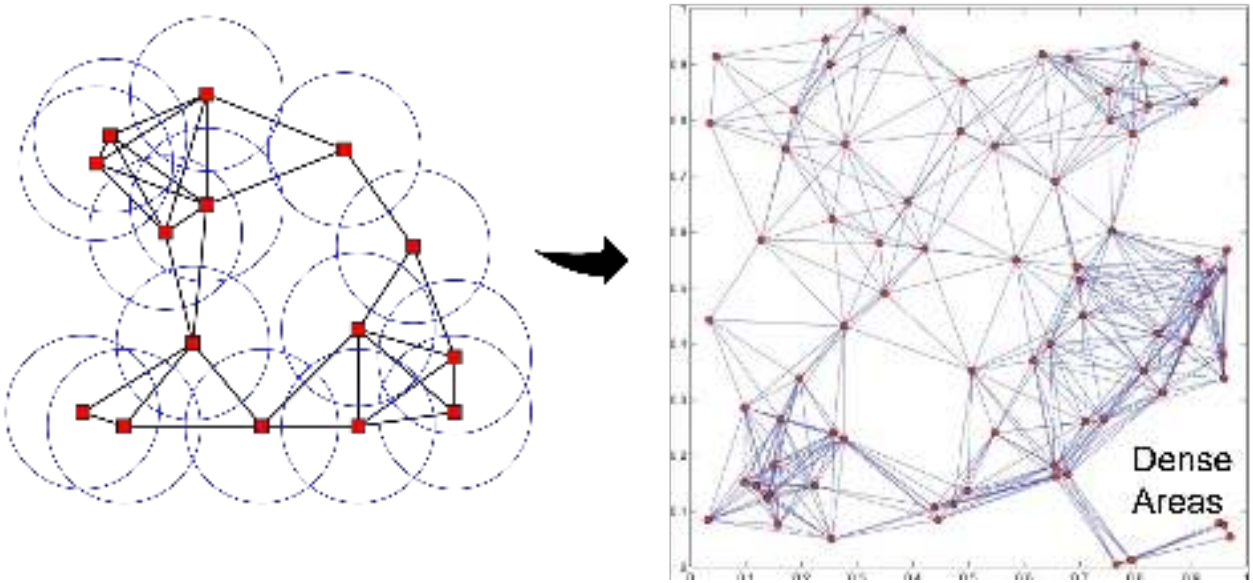
- The pairwise gossip algorithm is particularly convenient in wireless networks (e.g. synchronizing and coordinating robots swarms)
- The nodes can use a random access protocol called Carrier Sensing Multiple Access, sending a Request to Send and receiving a Clear to Send message to perform the update



Wireless pairwise random gossiping

Cornell University

- The pairwise gossip algorithm is particularly convenient in wireless networks (e.g. synchronizing and coordinating robots swarms)
- The nodes can use a random access protocol called Carrier Sensing Multiple Access, sending a Request to Send and receiving a Clear to Send message to perform the update



Dynamics

- If we consider the case of i.i.d. $\boldsymbol{W}(t)$ with mean $\overline{\boldsymbol{W}} = \mathbb{E}[\boldsymbol{W}(t)]$ we can take the expectation with respect to the process $\boldsymbol{W}(t)$

$$\mathbb{E}[\boldsymbol{x}(t+1)|\boldsymbol{x}(t)] = \mathbb{E}[\boldsymbol{W}(t)\boldsymbol{x}(t)|\boldsymbol{x}(t)] = \overbrace{\mathbb{E}[\boldsymbol{W}(t)]}^{\overline{\boldsymbol{W}}}\boldsymbol{x}(t)$$

the convergence properties depend on $\overline{\boldsymbol{W}} = \mathbb{E}[\boldsymbol{W}(t)]$

Convergence of random gossiping

If $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is strongly connected, and as long as

1. each $\boldsymbol{W}(t)$ is doubly stochastic, so that for $\mathbf{1}^\top \boldsymbol{W}(t) = \mathbf{1}^\top$,
2. $\lambda_2(\overline{\boldsymbol{W}}) < 1$

the network will reach consensus.

Dynamics

- If we consider the case of i.i.d. $\boldsymbol{W}(t)$ with mean $\overline{\boldsymbol{W}} = \mathbb{E}[\boldsymbol{W}(t)]$ we can take the expectation with respect to the process $\boldsymbol{W}(t)$

$$\mathbb{E}[\boldsymbol{x}(t+1)|\boldsymbol{x}(t)] = \mathbb{E}[\boldsymbol{W}(t)\boldsymbol{x}(t)|\boldsymbol{x}(t)] = \overbrace{\mathbb{E}[\boldsymbol{W}(t)]}^{\overline{\boldsymbol{W}}}\boldsymbol{x}(t)$$

the convergence properties depend on $\overline{\boldsymbol{W}} = \mathbb{E}[\boldsymbol{W}(t)]$

Convergence of random gossiping

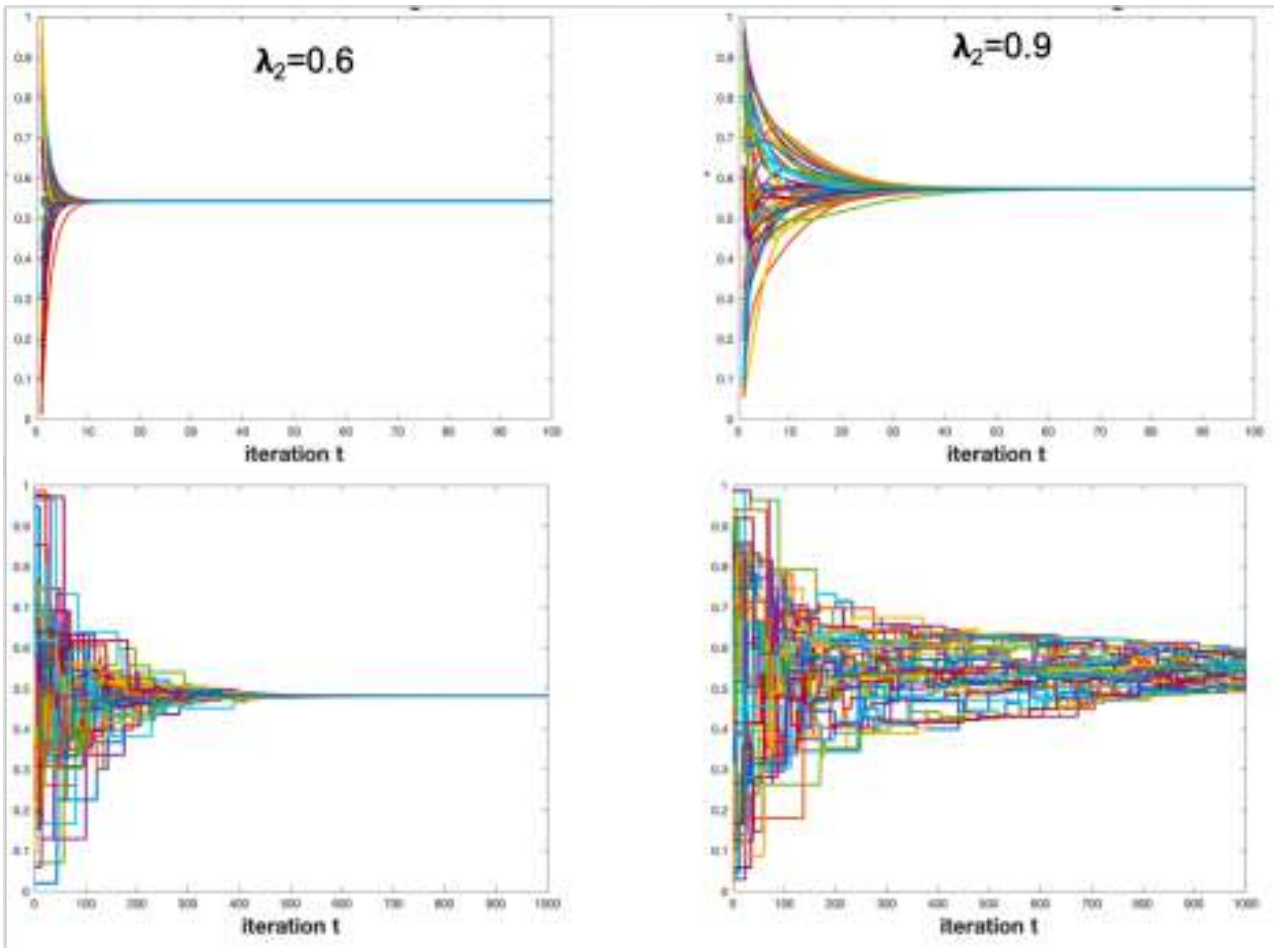
If $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is strongly connected, and as long as

1. each $\boldsymbol{W}(t)$ is doubly stochastic, so that for $\mathbf{1}^\top \boldsymbol{W}(t) = \mathbf{1}^\top$,
2. $\lambda_2(\overline{\boldsymbol{W}}) < 1$

the network will reach consensus.

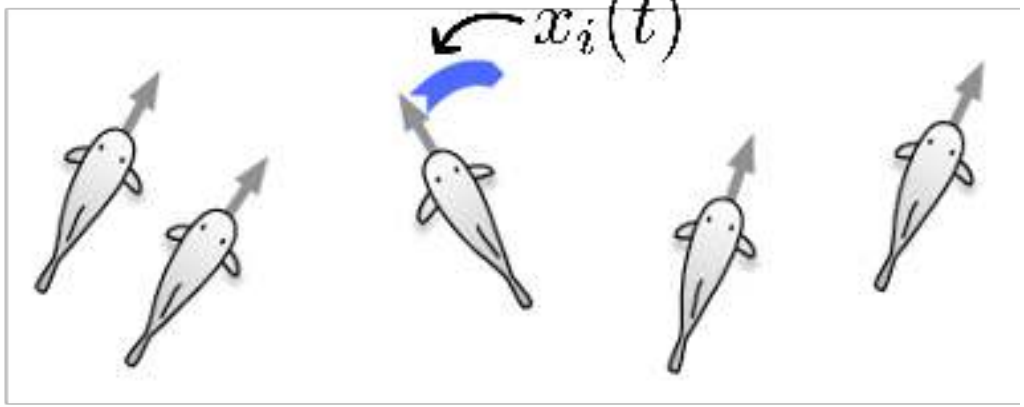
Dynamics simulation

- Comparing the difference between the dynamics $\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{x}(t)$ and $\boldsymbol{x}(t+1) = \overline{\boldsymbol{W}}\boldsymbol{x}(t)$, the deterministic dynamics with the average $\overline{\boldsymbol{W}}$



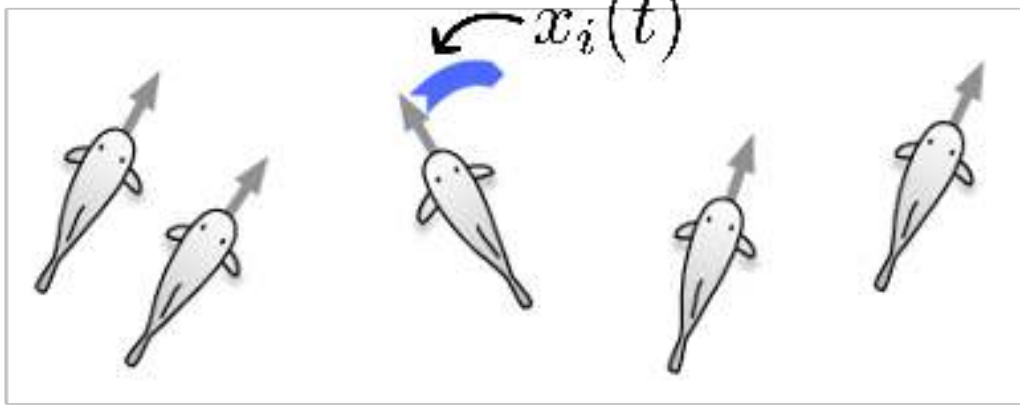
Randomized interaction model

- In addition to be used to define multi-agent computation protocols that require minimum overhead, randomized gossip algorithms are used to model **flocking behavior** in nature
- For instance, how flocks of fish align their angle, depening on the angle they perceive from their immediate neighbors



Randomized interaction model

- In addition to be used to define multi-agent computation protocols that require minimum overhead, randomized gossip algorithms are used to model **flocking behavior** in nature
- For instance, how flocks of fish align their angle, depening on the angle they perceive from their immediate neighbors



Summary of network processes discussed

- **Random walk**: At each t:
  1. $\boldsymbol{x}^{\text{rw}}(t) = \boldsymbol{e}_{v_t}(t)$, where v_t is the vertex chosen at random and $\boldsymbol{e}_v(t)$ is the $v$thcoordinate vector
  2. The probability mass of the vertex being chosen is the vector such that $\boldsymbol{\pi}(t+1) = \boldsymbol{W}\boldsymbol{\pi}(t)$, where $\boldsymbol{W}$ is the transition probability matrix ($\boldsymbol{W} = \boldsymbol{I} - \boldsymbol{L}^{\text{rw}}$ for the uniform random walk)
  3. $\lambda_2(\boldsymbol{W}) < 1$ implies that $\boldsymbol{\pi}(t) = \boldsymbol{\pi}^* + O(\lambda_2^t(\boldsymbol{W}))$
- **Consensus Gossiping**: At each t:
  1. The state is updated **synchrounously** $\boldsymbol{x}(t+1) = \boldsymbol{W}\boldsymbol{x}(t)$,
  2. The $\|\boldsymbol{x}(t) - x^{\text{ave}}\mathbf{1}\| = O(\lambda_2^t(\boldsymbol{W}))$
  3. The state is updated **asynchrounously** $\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{x}(t)$,
  4. The $\mathbb{E}[\|\boldsymbol{x}(t) - x^{\text{ave}}\mathbf{1}\|] = O(\lambda_2^t(\overline{\boldsymbol{W}}))$

Summary of network processes discussed

- **Random walk**: At each t:
  1. $\boldsymbol{x}^{\text{rw}}(t) = \boldsymbol{e}_{v_t}(t)$, where v_t is the vertex chosen at random and $\boldsymbol{e}_v(t)$ is the $v$thcoordinate vector
  2. The probability mass of the vertex being chosen is the vector such that $\boldsymbol{\pi}(t+1) = \boldsymbol{W}\boldsymbol{\pi}(t)$, where $\boldsymbol{W}$ is the transition probability matrix ($\boldsymbol{W} = \boldsymbol{I} - \boldsymbol{L}^{\text{rw}}$ for the uniform random walk)
  3. $\lambda_2(\boldsymbol{W}) < 1$ implies that $\boldsymbol{\pi}(t) = \boldsymbol{\pi}^* + O(\lambda_2^t(\boldsymbol{W}))$
- **Consensus Gossiping**: At each t:
  1. The state is updated **synchrounously** $\boldsymbol{x}(t+1) = \boldsymbol{W}\boldsymbol{x}(t)$,
  2. The $\|\boldsymbol{x}(t) - x^{\text{ave}}\mathbf{1}\| = O(\lambda_2^t(\boldsymbol{W}))$
  3. The state is updated **asynchrounously** $\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{x}(t)$,
  4. The $\mathbb{E}[\|\boldsymbol{x}(t) - x^{\text{ave}}\mathbf{1}\|] = O(\lambda_2^t(\overline{\boldsymbol{W}}))$

Summary of network processes discussed

- **Random walk**: At each t:
  1. $\boldsymbol{x}^{\text{rw}}(t) = \boldsymbol{e}_{v_t}(t)$, where v_t is the vertex chosen at random and $\boldsymbol{e}_v(t)$ is the $v$thcoordinate vector
  2. The probability mass of the vertex being chosen is the vector such that $\boldsymbol{\pi}(t+1) = \boldsymbol{W}\boldsymbol{\pi}(t)$, where $\boldsymbol{W}$ is the transition probability matrix ($\boldsymbol{W} = \boldsymbol{I} - \boldsymbol{L}^{\text{rw}}$ for the uniform random walk)
  3. $\lambda_2(\boldsymbol{W}) < 1$ implies that $\boldsymbol{\pi}(t) = \boldsymbol{\pi}^* + O(\lambda_2^t(\boldsymbol{W}))$
- **Consensus Gossiping**: At each t:
  1. The state is updated **synchrounously** $\boldsymbol{x}(t+1) = \boldsymbol{W}\boldsymbol{x}(t)$,
  2. The $\|\boldsymbol{x}(t) - x^{\text{ave}}\mathbf{1}\| = O(\lambda_2^t(\boldsymbol{W}))$
  3. The state is updated **asynchrounously** $\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{x}(t)$,
  4. The $\mathbb{E}[\|\boldsymbol{x}(t) - x^{\text{ave}}\mathbf{1}\|] = O(\lambda_2^t(\overline{\boldsymbol{W}}))$

Summary of network processes discussed

- **Random walk**: At each t:
  1. $\boldsymbol{x}^{\text{rw}}(t) = \boldsymbol{e}_{v_t}(t)$, where v_t is the vertex chosen at random and $\boldsymbol{e}_v(t)$ is the $v$thcoordinate vector
  2. The probability mass of the vertex being chosen is the vector such that $\boldsymbol{\pi}(t+1) = \boldsymbol{W}\boldsymbol{\pi}(t)$, where $\boldsymbol{W}$ is the transition probability matrix ($\boldsymbol{W} = \boldsymbol{I} - \boldsymbol{L}^{\text{rw}}$ for the uniform random walk)
  3. $\lambda_2(\boldsymbol{W}) < 1$ implies that $\boldsymbol{\pi}(t) = \boldsymbol{\pi}^* + O(\lambda_2^t(\boldsymbol{W}))$
- **Consensus Gossiping**: At each t:
  1. The state is updated **synchrounously** $\boldsymbol{x}(t+1) = \boldsymbol{W}\boldsymbol{x}(t)$,
  2. The $\|\boldsymbol{x}(t) - x^{\text{ave}}\mathbf{1}\| = O(\lambda_2^t(\boldsymbol{W}))$
  3. The state is updated **asynchrounously** $\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{x}(t)$,
  4. The $\mathbb{E}[\|\boldsymbol{x}(t) - x^{\text{ave}}\mathbf{1}\|] = O(\lambda_2^t(\overline{\boldsymbol{W}}))$

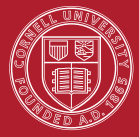
Summary of network processes discussed

- **Random walk**: At each t:
  1. $\boldsymbol{x}^{\text{rw}}(t) = \boldsymbol{e}_{v_t}(t)$, where v_t is the vertex chosen at random and $\boldsymbol{e}_v(t)$ is the $v$thcoordinate vector
  2. The probability mass of the vertex being chosen is the vector such that $\boldsymbol{\pi}(t+1) = \boldsymbol{W}\boldsymbol{\pi}(t)$, where $\boldsymbol{W}$ is the transition probability matrix ($\boldsymbol{W} = \boldsymbol{I} - \boldsymbol{L}^{\text{rw}}$ for the uniform random walk)
  3. $\lambda_2(\boldsymbol{W}) < 1$ implies that $\boldsymbol{\pi}(t) = \boldsymbol{\pi}^* + O(\lambda_2^t(\boldsymbol{W}))$
- **Consensus Gossiping**: At each t:
  1. The state is updated **synchrounously** $\boldsymbol{x}(t+1) = \boldsymbol{W}\boldsymbol{x}(t)$,
  2. The $\|\boldsymbol{x}(t) - x^{\text{ave}}\mathbf{1}\| = O(\lambda_2^t(\boldsymbol{W}))$
  3. The state is updated **asynchrounously** $\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{x}(t)$,
  4. The $\mathbb{E}[\|\boldsymbol{x}(t) - x^{\text{ave}}\mathbf{1}\|] = O(\lambda_2^t(\overline{\boldsymbol{W}}))$

Summary of network processes discussed

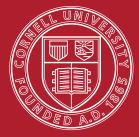
- **Random walk**: At each t:
  1. $\boldsymbol{x}^{\text{rw}}(t) = \boldsymbol{e}_{v_t}(t)$, where v_t is the vertex chosen at random and $\boldsymbol{e}_v(t)$ is the $v$thcoordinate vector
  2. The probability mass of the vertex being chosen is the vector such that $\boldsymbol{\pi}(t+1) = \boldsymbol{W}\boldsymbol{\pi}(t)$, where $\boldsymbol{W}$ is the transition probability matrix ($\boldsymbol{W} = \boldsymbol{I} - \boldsymbol{L}^{\text{rw}}$ for the uniform random walk)
  3. $\lambda_2(\boldsymbol{W}) < 1$ implies that $\boldsymbol{\pi}(t) = \boldsymbol{\pi}^* + O(\lambda_2^t(\boldsymbol{W}))$
- **Consensus Gossiping**: At each t:
  1. The state is updated **synchrounously** $\boldsymbol{x}(t+1) = \boldsymbol{W}\boldsymbol{x}(t)$,
  2. The $\|\boldsymbol{x}(t) - x^{\text{ave}}\mathbf{1}\| = O(\lambda_2^t(\boldsymbol{W}))$
  3. The state is updated **asynchrounously** $\boldsymbol{x}(t+1) = \boldsymbol{W}(t)\boldsymbol{x}(t)$,
  4. The $\mathbb{E}[\|\boldsymbol{x}(t) - x^{\text{ave}}\mathbf{1}\|] = O(\lambda_2^t(\overline{\boldsymbol{W}}))$

Outline



Cornell University

Concluding remarks



- In this lecture we introduced two network dynamics:
  - The random walk
  - The average consensus dynamics
- Next time we will introduce non-linear dynamical models in general, and then introduce epidemic models