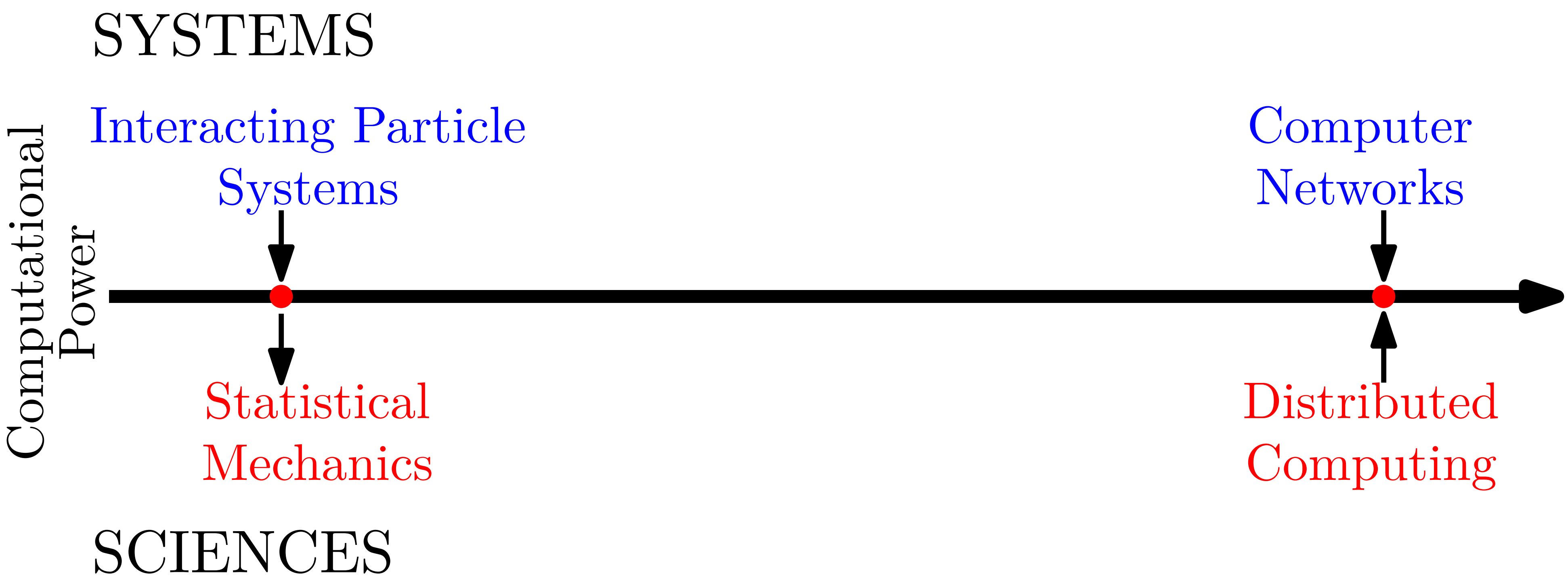
On the Computational Power of Simple Dynamics

Emanuele Natale

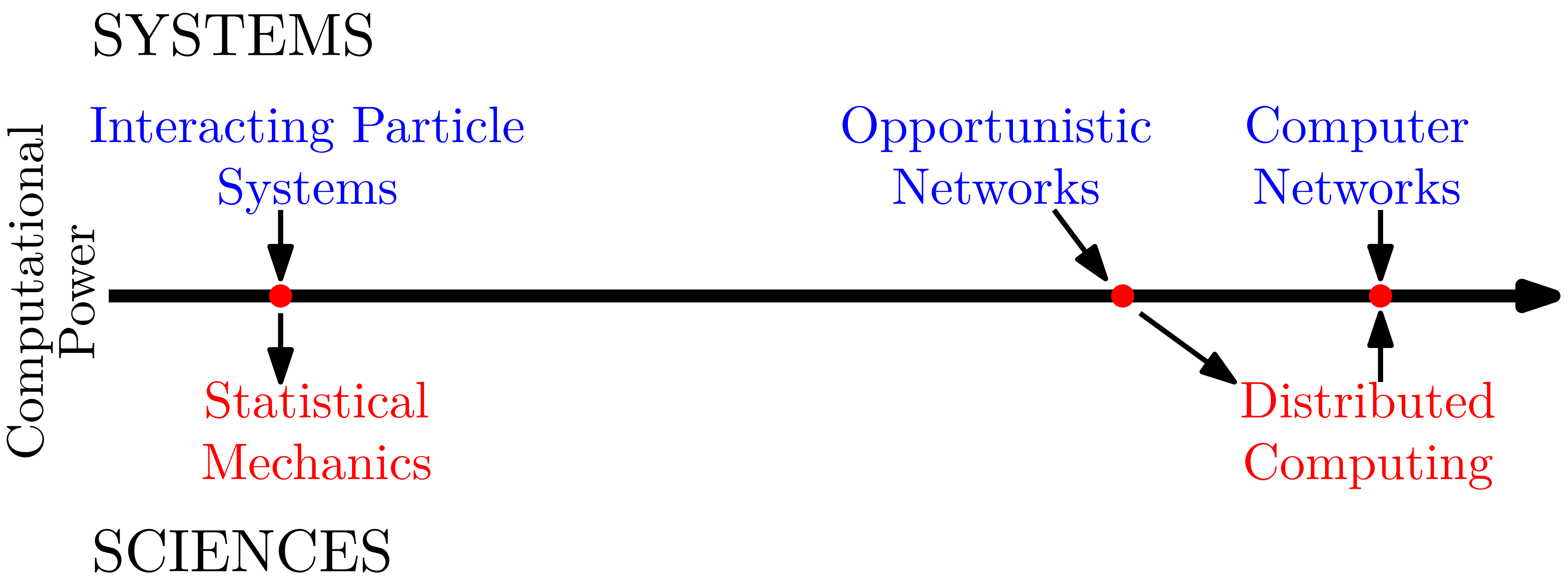
Supervisors: Andrea Clementi, Riccardo Silvestri

Sapienza University of Rome
February 13, 2017

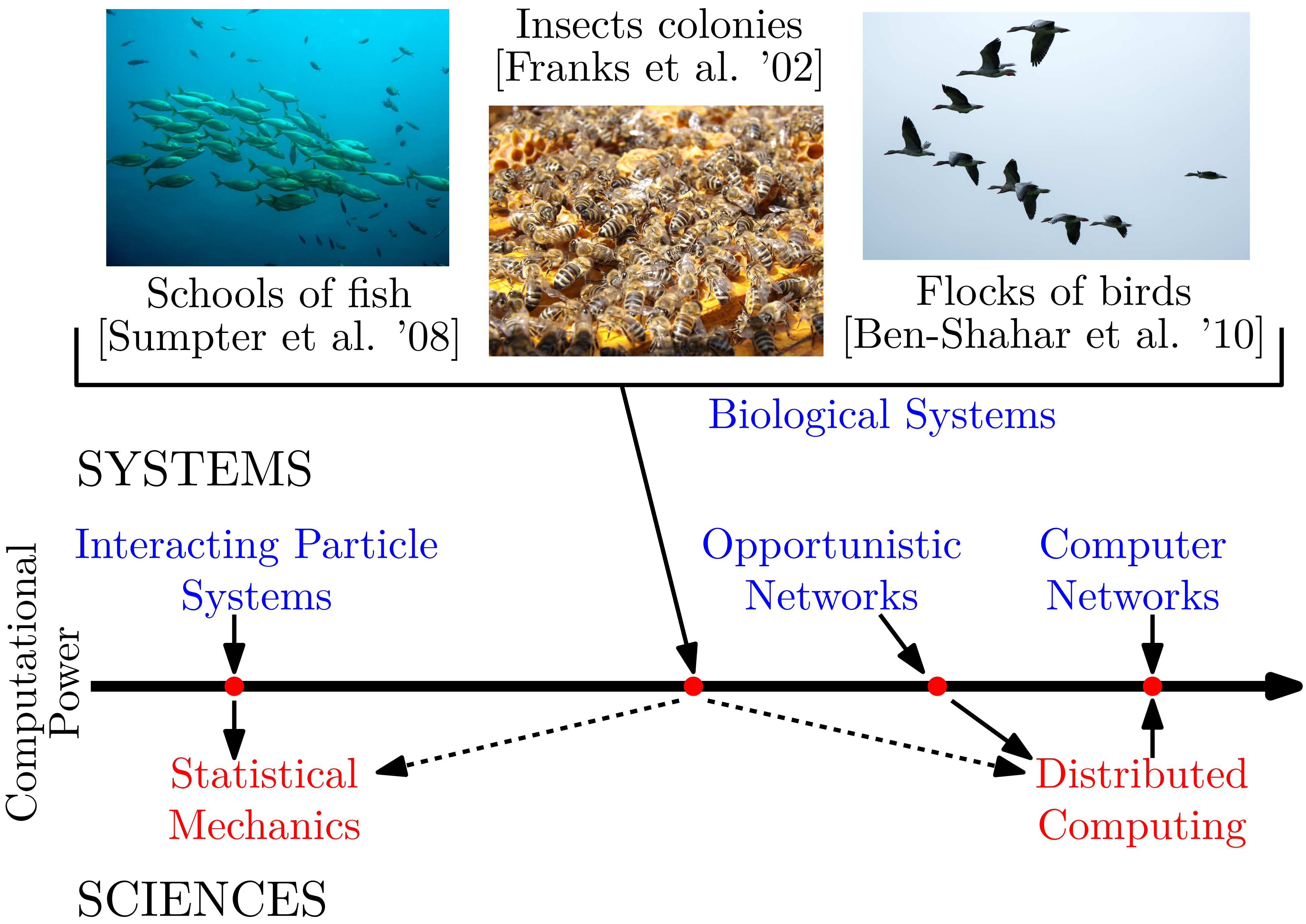
Communication in *Simple* Systems



Communication in *Simple* Systems



Communication in *Simple* Systems



Communication Model

Animal communication:

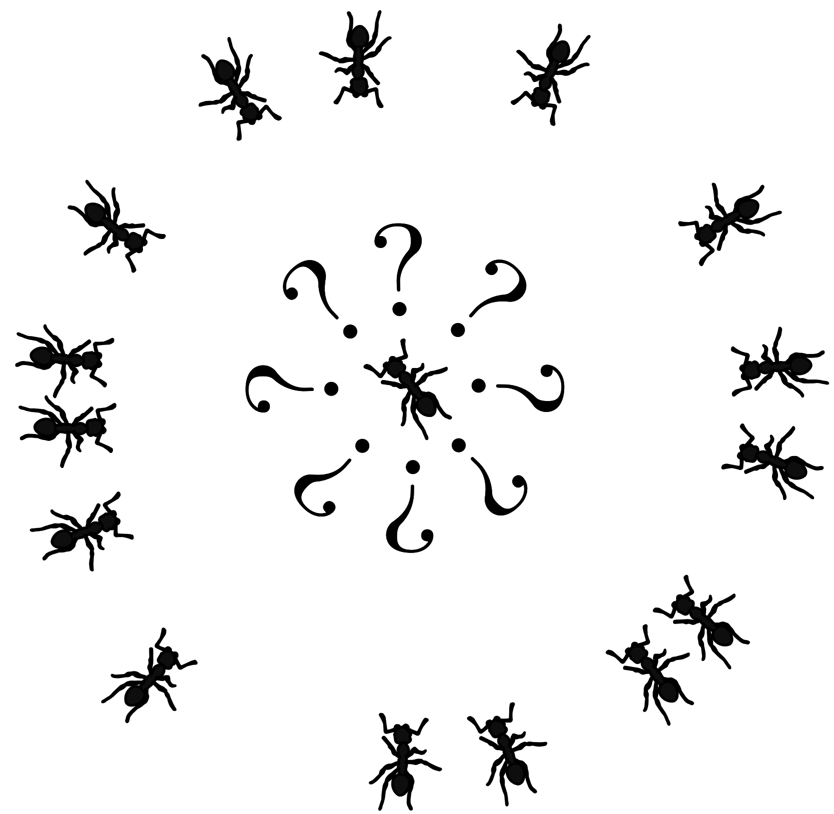
- Chaotic
- Anonymous
- Parsimonious
- Uni-directional (Passive)
- Noisy

Communication Model

Animal communication:

- Chaotic
- Anonymous
- Parsimonious
- Uni-directional (Passive)
- Noisy

$\textcolor{blue}{\mathcal{PULL}(h,\ell)}$ **model** [**Demers '88**]: at each round each agent can *observe* h other agents chosen independently and uniformly at random, and *shows* ℓ *bits* to her observers.



Communication Model

Animal communication:

- Chaotic
- Anonymous
- Parsimonious
- Uni-directional (Passive)
- Noisy

$\mathcal{PULL}(h,\ell)$ **model** [**Demers '88**]: at each round each agent can *observe* h other agents chosen independently and uniformly at random, and *shows* ℓ *bits* to her observers.

Dynamics

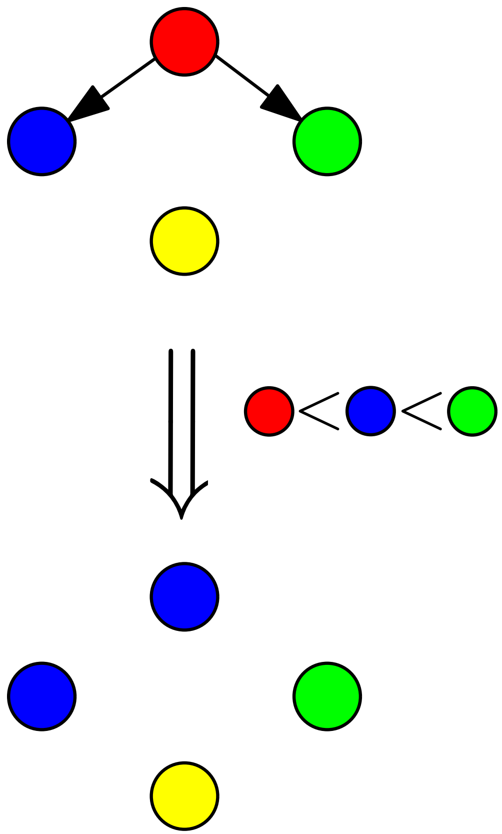
(informal) *Very simple* distributed algorithms: For every graph, agent and round, states are updated according to fixed rule of current state and symmetric function of states of neighbors.

Dynamics

(informal) *Very simple* distributed algorithms: For every graph, agent and round, states are updated according to fixed rule of current state and symmetric function of states of neighbors.

Examples of Dynamics

- 3-Median dynamics

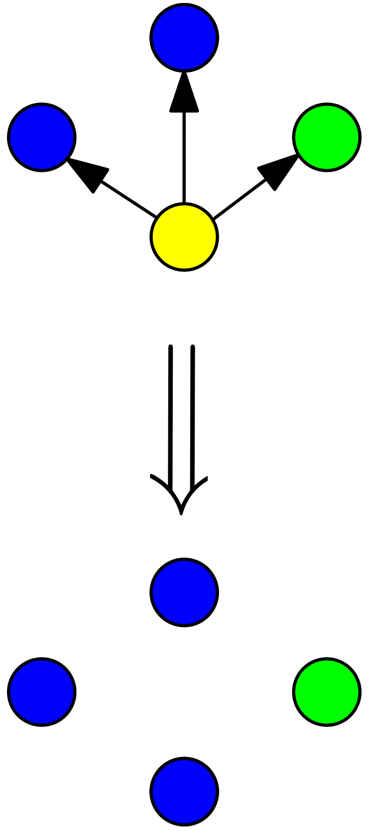


Dynamics

(informal) *Very simple* distributed algorithms: For every graph, agent and round, states are updated according to fixed rule of current state and symmetric function of states of neighbors.

Examples of Dynamics

- 3-Median dynamics
- 3-Majority dynamics

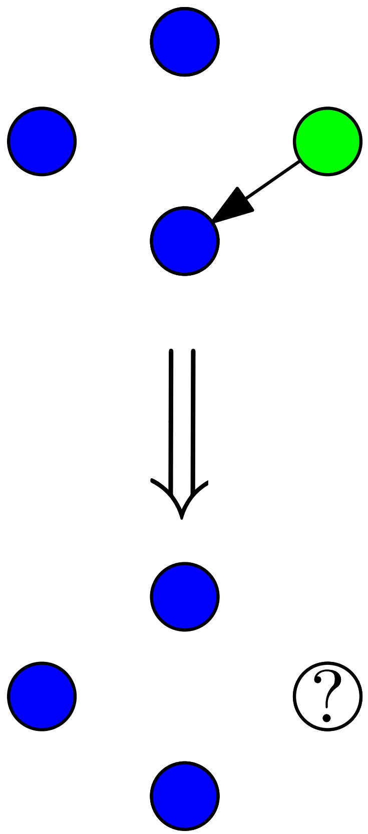


Dynamics

(informal) *Very simple* distributed algorithms: For every graph, agent and round, states are updated according to fixed rule of current state and symmetric function of states of neighbors.

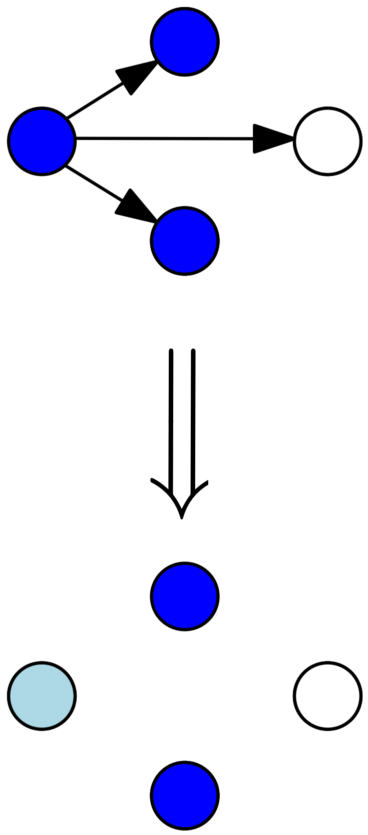
Examples of Dynamics

- 3-Median dynamics
- 3-Majority dynamics
- Undecided-state dynamics



Dynamics

(informal) *Very simple* distributed algorithms: For every graph, agent and round, states are updated according to fixed rule of current state and symmetric function of states of neighbors.



Examples of Dynamics

- 3-Median dynamics
- 3-Majority dynamics
- Undecided-state dynamics
- Averaging dynamics

The Power of Dynamics: Plurality Consensus

Computing the Median

3-Median dynamics [Doerr et al. '11]. Converge to $\mathcal{O}(\sqrt{n \log n})$ approximation of median of system in $\mathcal{O}(\log n)$ rounds w.h.p., even if $\mathcal{O}(\sqrt{n})$ states are arbitrarily changed at each round ($\mathcal{O}(\sqrt{n})$-bounded adversary).

The Power of Dynamics: Plurality Consensus

Computing the Median

3-Median dynamics [Doerr et al. '11]. Converge to $\mathcal{O}(\sqrt{n \log n})$ approximation of median of system in $\mathcal{O}(\log n)$ rounds w.h.p., even if $\mathcal{O}(\sqrt{n})$ states are arbitrarily changed at each round ($\mathcal{O}(\sqrt{n})$-bounded adversary).

Computing the Majority

3-Majority dynamics [SPAA '14, SODA '16]. If plurality has **bias** $\mathcal{O}(\sqrt{kn \log n})$, converges to it in $\mathcal{O}(k \log n)$ rounds w.h.p., even against $o(\sqrt{n/k})$-bounded adversary. Without bias, converges in $\text{poly}(k)$. h-majority converges in $\Omega(k/h^2)$.

The Power of Dynamics: Plurality Consensus

Computing the Median

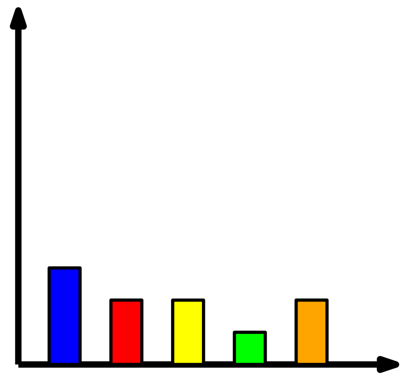
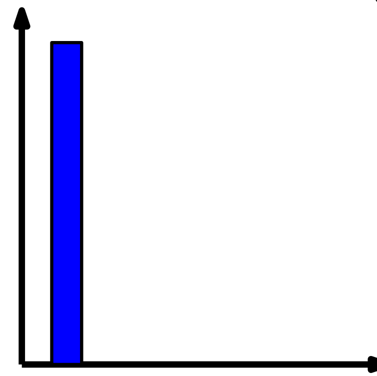
3-Median dynamics [Doerr et al. '11]. Converge to $\mathcal{O}(\sqrt{n \log n})$ approximation of median of system in $\mathcal{O}(\log n)$ rounds w.h.p., even if $\mathcal{O}(\sqrt{n})$ states are arbitrarily changed at each round ($\mathcal{O}(\sqrt{n})$-bounded adversary).

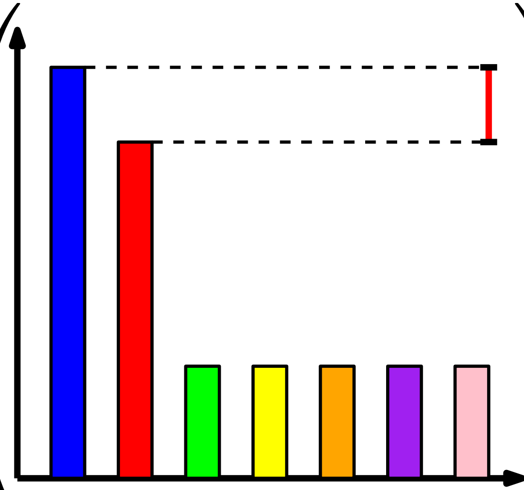
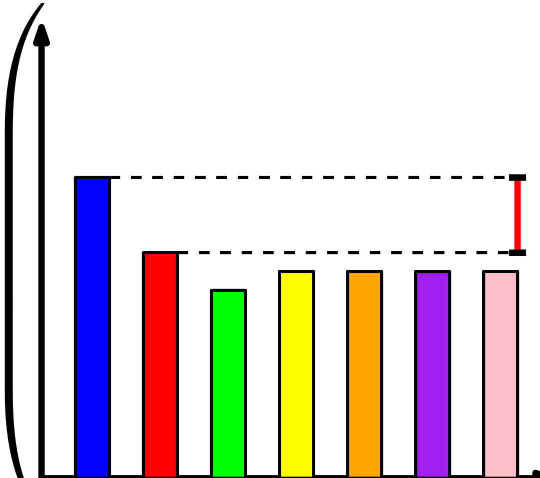
Computing the Majority

3-Majority dynamics [SPAA '14, SODA '16]. If plurality has **bias** $\mathcal{O}(\sqrt{kn \log n})$, converges to it in $\mathcal{O}(k \log n)$ rounds w.h.p., even against $o(\sqrt{n/k})$-bounded adversary. Without bias, converges in $\text{poly}(k)$. h-majority converges in $\Omega(k/h^2)$.

Undecided-State dynamics [SODA '15]. If majority/second-majority ($c_{maj}/c_{2^{nd}maj}$) is at least $1 + \epsilon$, system converges to plurality within $\tilde{\Theta}(\text{md}(\mathbf{c}))$ rounds w.h.p.

A Global Measure of Bias

$$\mathrm{md}(\mathbf{c}^{(\mathbf{0})}) := \sum_{i=1}^{k} \left(\frac{c_i^{(0)}}{c_{maj}^{(0)}} \right)^2 = 1 + \mathcal{D}\Bigg($$  ,  $$\Bigg)$$

$$1 \leq \mathrm{md}\Bigg($$  $$\Bigg) \ll \mathrm{md}\Bigg($$  $$\Bigg) \leq k$$

Undecided-State dynamics [SODA '15]. If majority/second-majority ($c_{maj}/c_{2^{nd}maj}$) is at least $1+\epsilon$, system converges to plurality within $\tilde{\Theta}(\mathrm{md}(\mathbf{c}))$ rounds w.h.p.

The Median, the Mode and... the Mean

Dynamics can solve Consensus, Median, Majority, in robust and fault tolerant ways, but this is trivial in centralized setting.

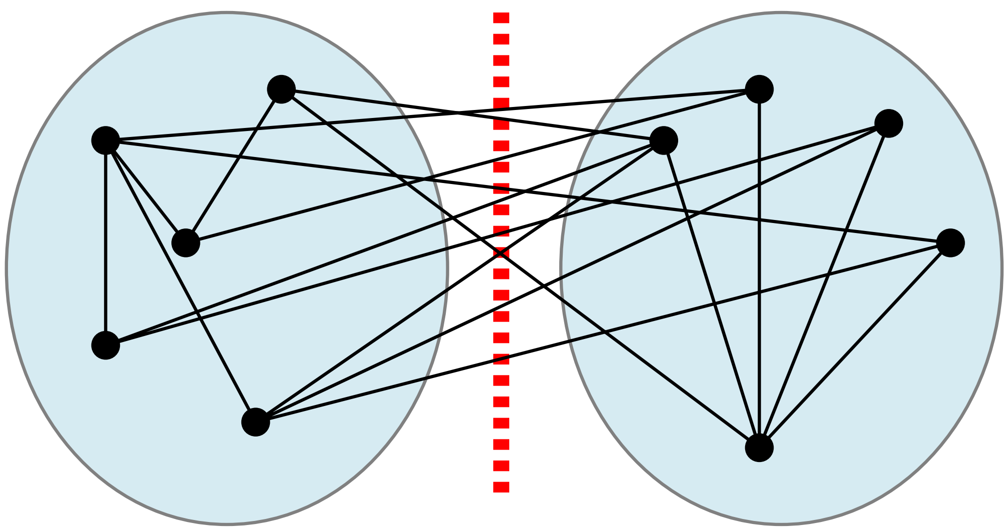
The Median, the Mode and... the Mean

Dynamics can solve Consensus, Median, Majority, in robust and fault tolerant ways, but this is trivial in centralized setting.

Can dynamics solve a problem non-trivial in centralized setting?

Community Detection as Minimum Bisection

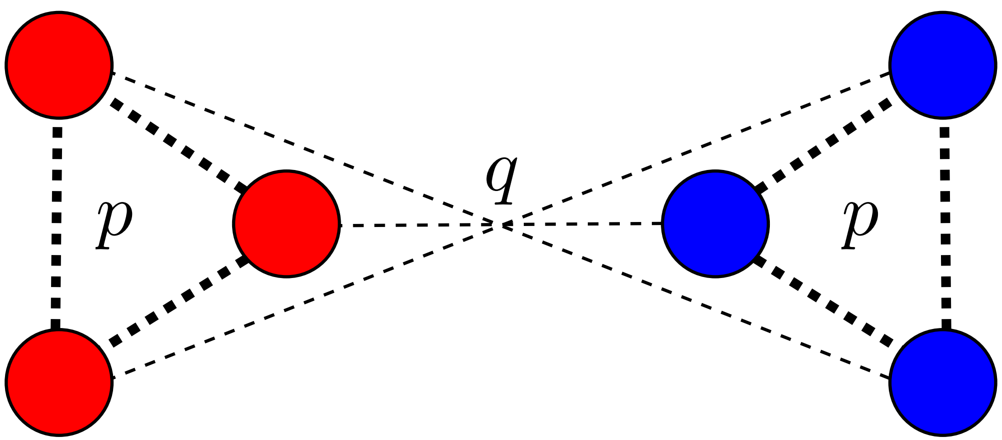
Minimum Bisection Problem.
Input: a graph G with $2n$ nodes.
Output: $S = \arg\min_{\substack{S \subset V \\ |S| = n}} E(S, V - S)$.



[Garey, Johnson, Stockmeyer '76]:
Min-Bisection is *NP-Complete*.

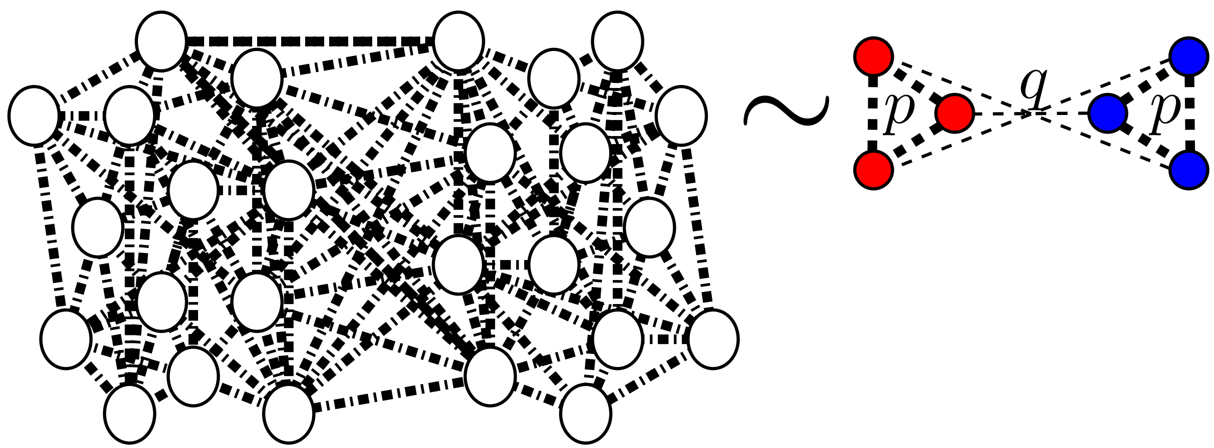
The Stochastic Block Model

Stochastic Block Model (SBM). Two "communities" of equal size V_1 and V_2, each edge inside a community included with probability $p = \frac{a}{n}$, each edge across communities included with probability $q = \frac{b}{n} < p$.



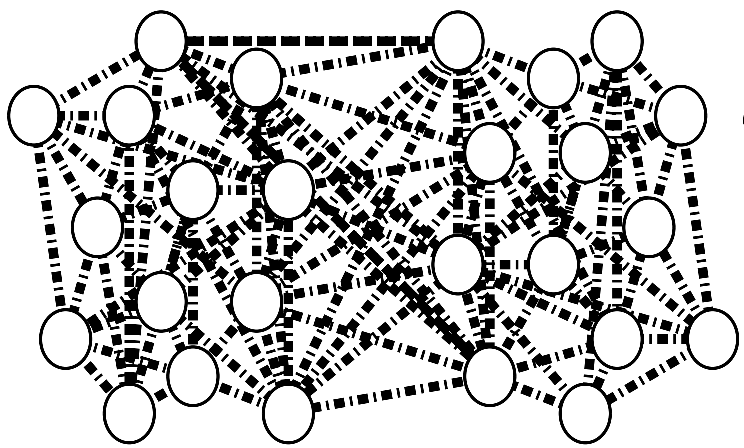
The Stochastic Block Model

Reconstruction problem. Given graph generated by SBM, find original partition.

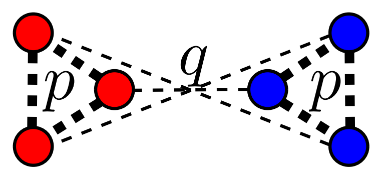


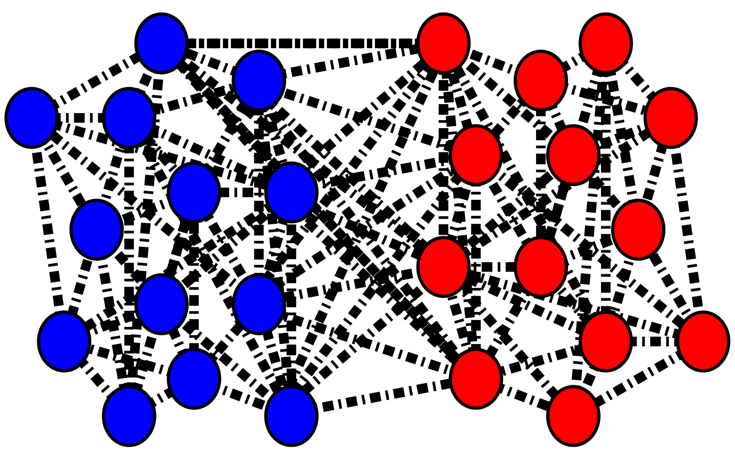
The Stochastic Block Model

Reconstruction problem. Given graph generated by SBM, find original partition.



$\sim$

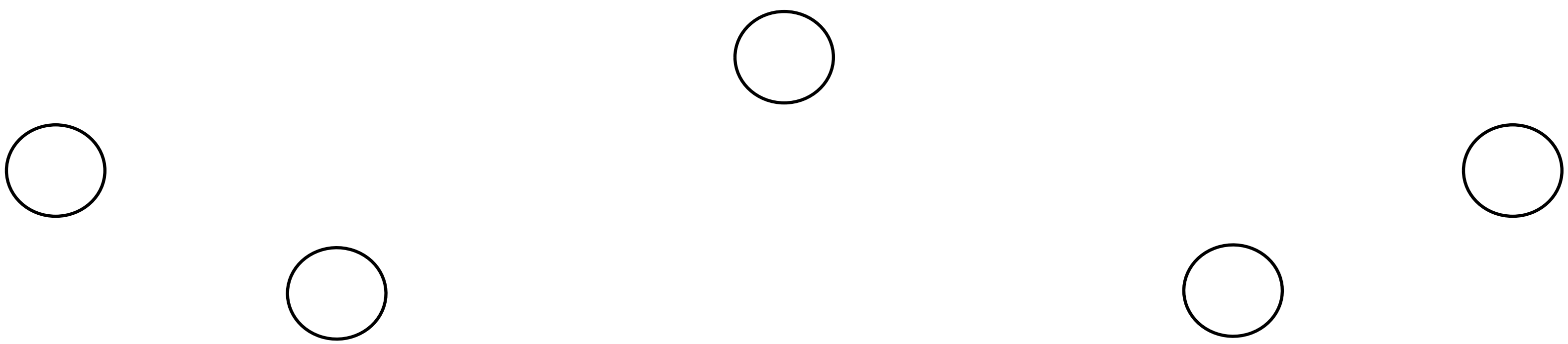




The Averaging Dynamics

Al nodes at the same time:

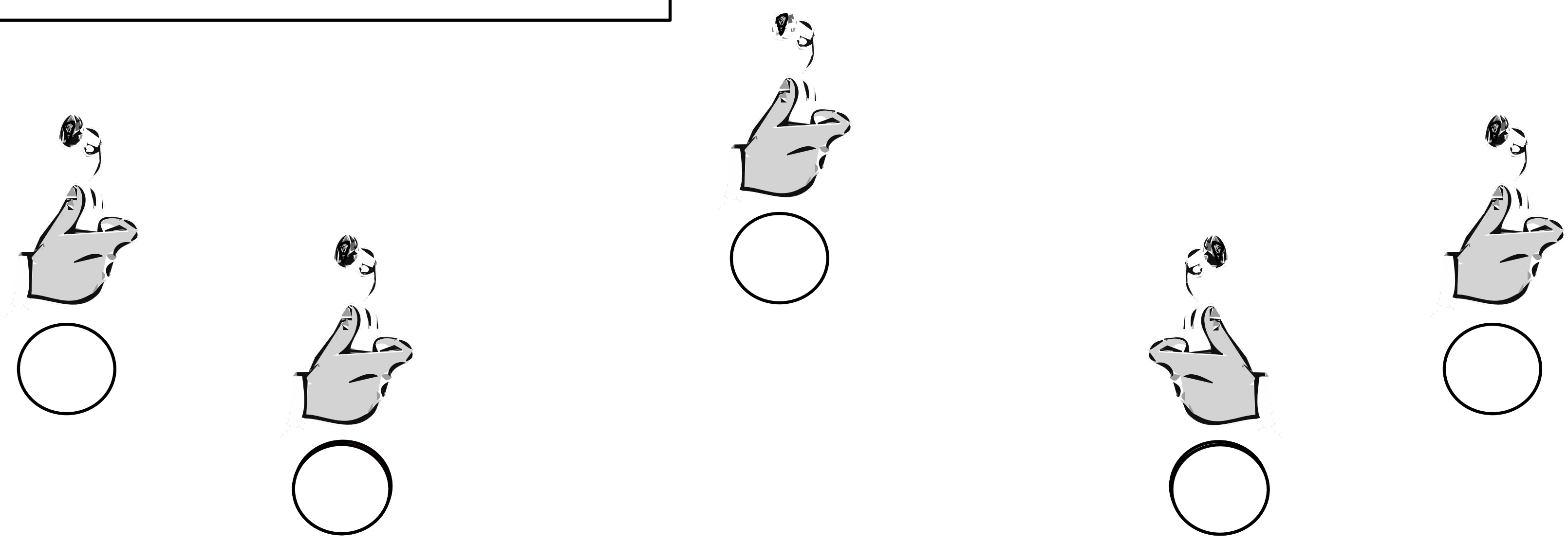
- At $t = 0$, randomly pick value $x^{(t)} \in \{\mathbf{+1}, \mathbf{-1}\}$.
- Then, at each round
  1. Set value $x^{(t)}$ to average of neighbors,
  2. Set label to **blue** if $x^{(t)} < x^{(t-1)}$, **red** otherwise.



The Averaging Dynamics

Al nodes at the same time:

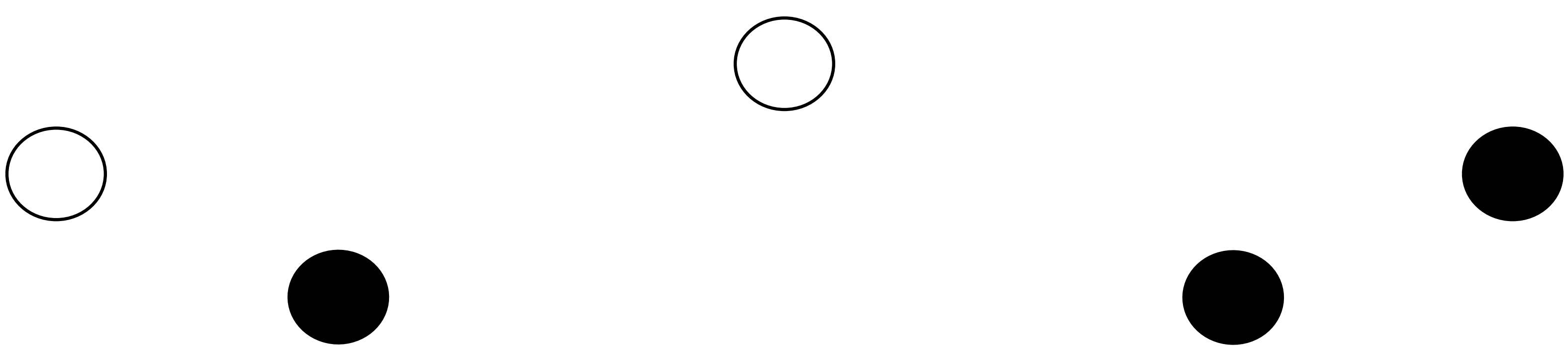
- At $t = 0$, randomly pick value $x^{(t)} \in \{\mathbf{+1}, \mathbf{\text{-}1}\}$.
- Then, at each round
  1. Set value $x^{(t)}$ to average of neighbors,
  2. Set label to **blue** if $x^{(t)} < x^{(t-1)}$, **red** otherwise.



The Averaging Dynamics

Al nodes at the same time:

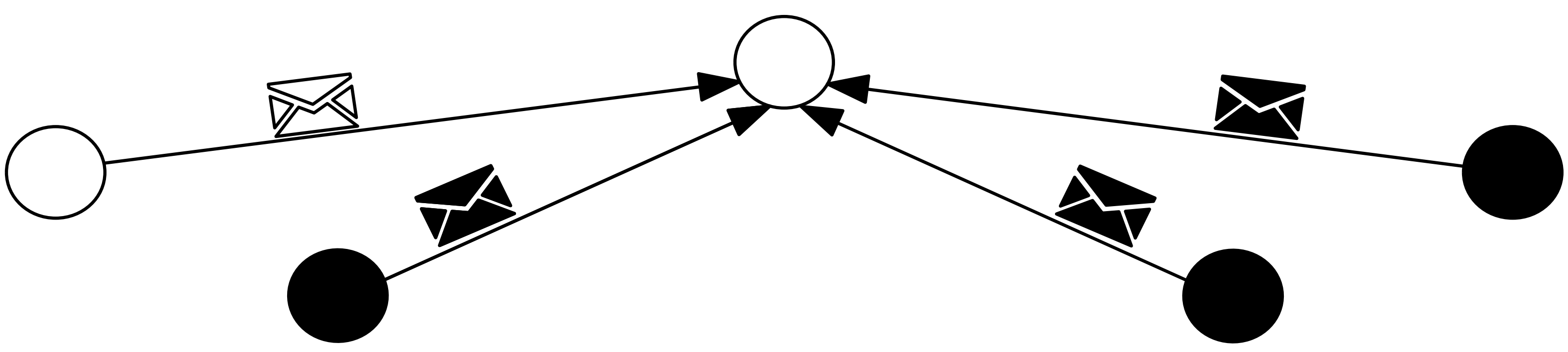
- At $t = 0$, randomly pick value $x^{(t)} \in \{\mathbf{+1}, \mathbf{\text{-}1}\}$.
- Then, at each round
  1. Set value $x^{(t)}$ to average of neighbors,
  2. Set label to **blue** if $x^{(t)} < x^{(t-1)}$, **red** otherwise.



The Averaging Dynamics

Al nodes at the same time:

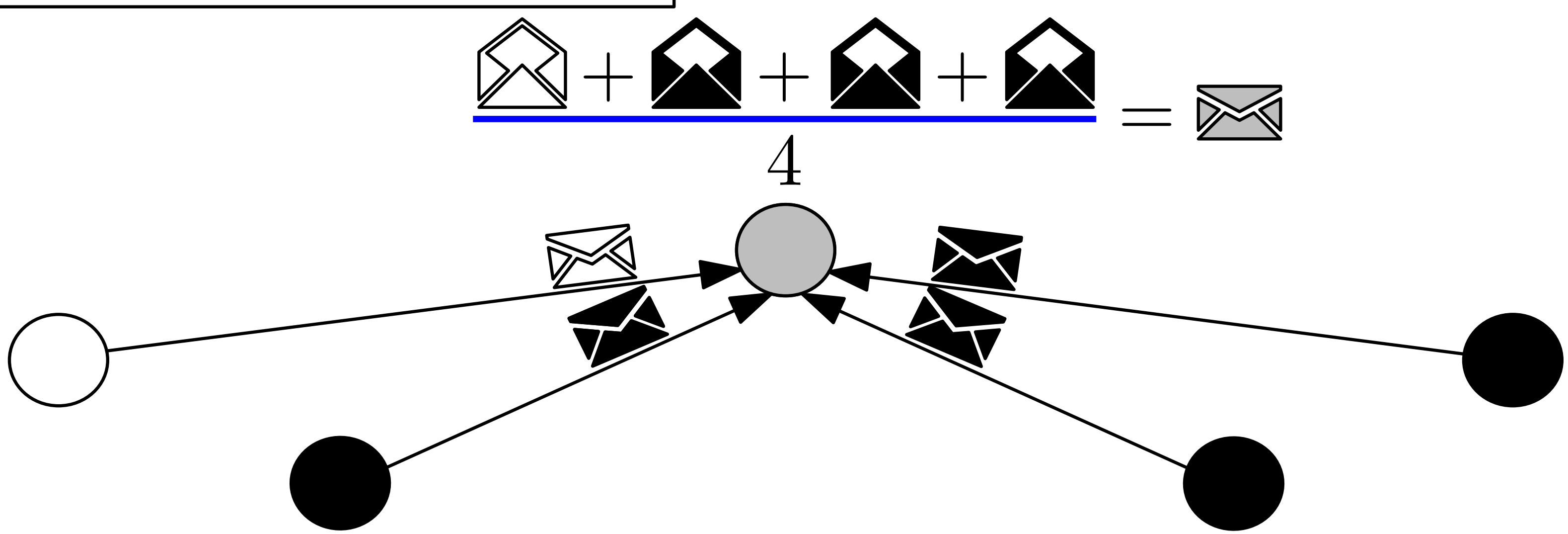
- At $t = 0$, randomly pick value $x^{(t)} \in \{\mathbf{+1}, \mathbf{-1}\}$.
- Then, at each round
  1. Set value $x^{(t)}$ to average of neighbors,
  2. Set label to **blue** if $x^{(t)} < x^{(t-1)}$, **red** otherwise.



The Averaging Dynamics

Al nodes at the same time:

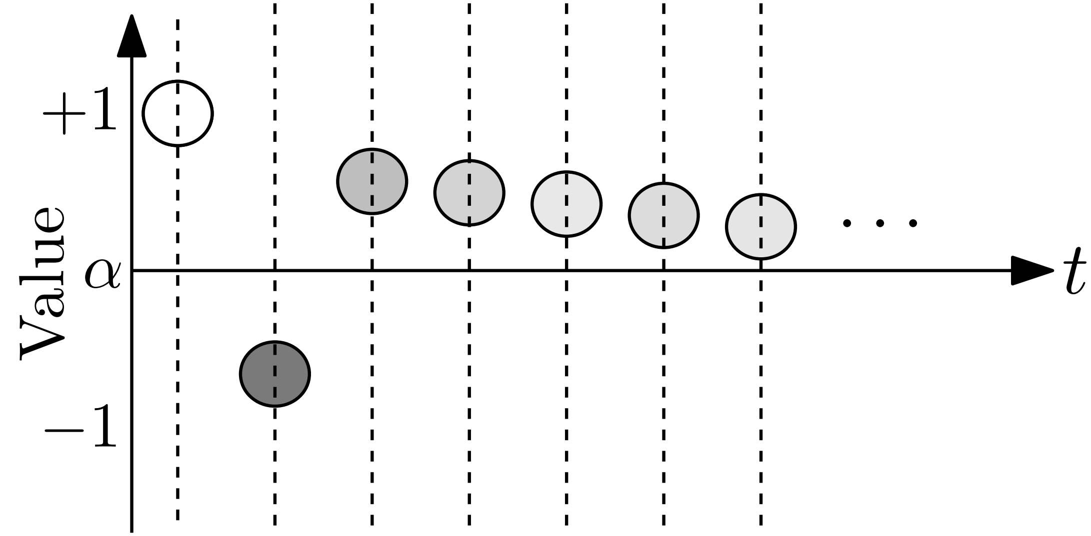
- At $t = 0$, randomly pick value $x^{(t)} \in \{\mathbf{+1}, \mathbf{-1}\}$.
- Then, at each round
  1. Set value $x^{(t)}$ to average of neighbors,
  2. Set label to **blue** if $x^{(t)} < x^{(t-1)}$, **red** otherwise.

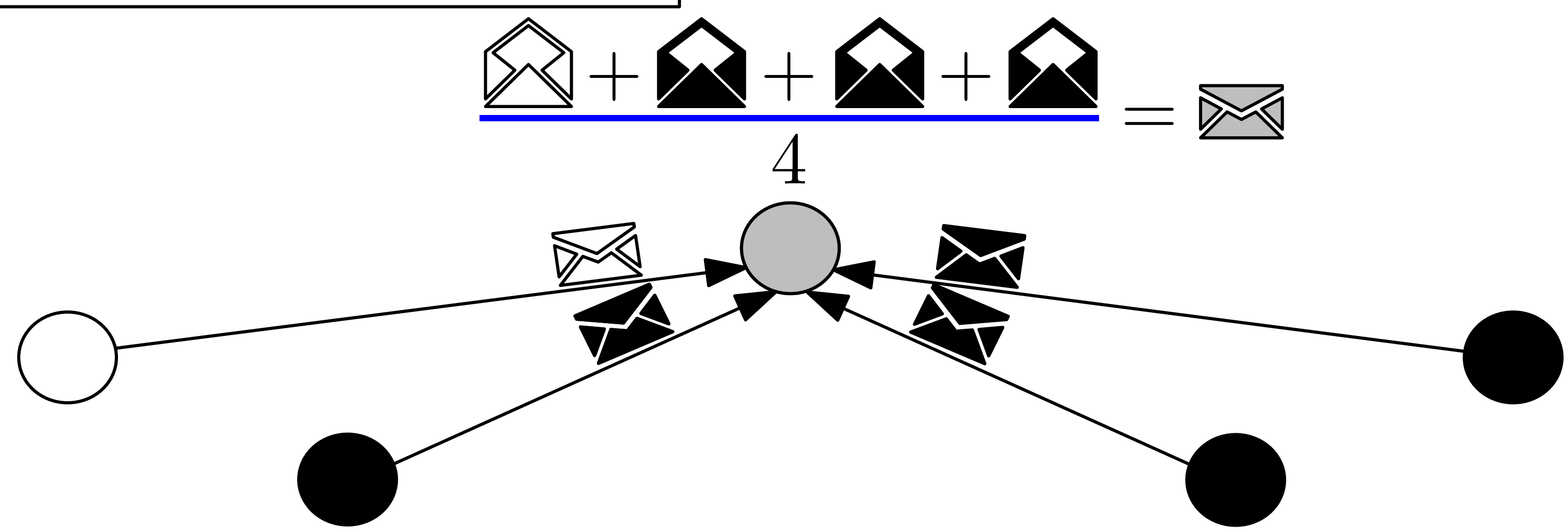


The Averaging Dynamics

Al nodes at the same time:

- At $t = 0$, randomly pick value $x^{(t)} \in \{\mathbf{+1}, \mathbf{-1}\}$.
- Then, at each round
  1. Set value $x^{(t)}$ to average of neighbors,
  2. Set label to **blue** if $x^{(t)} < x^{(t-1)}$, **red** otherwise.

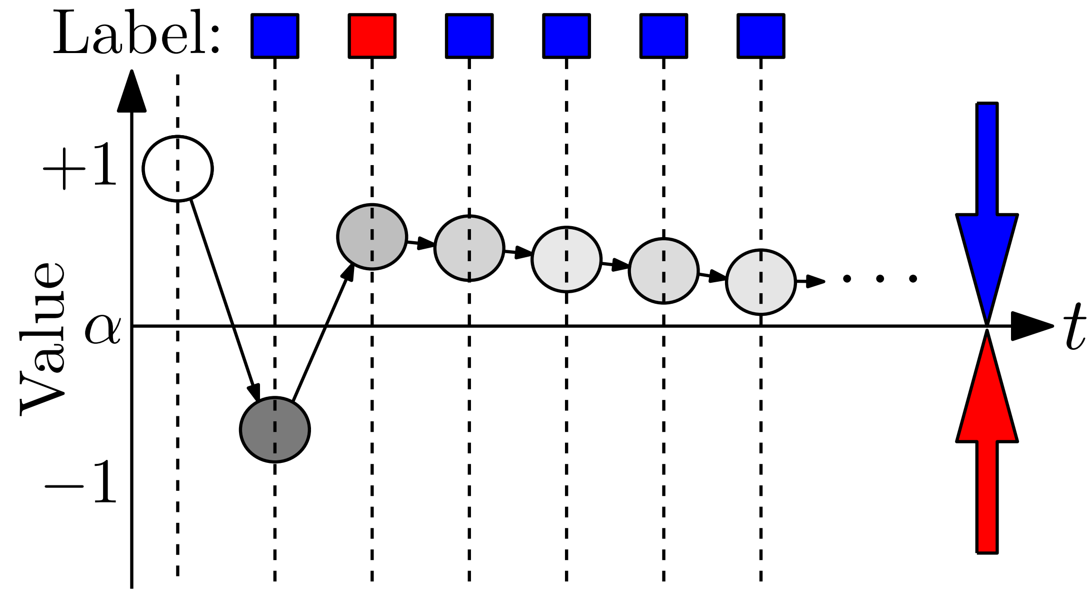


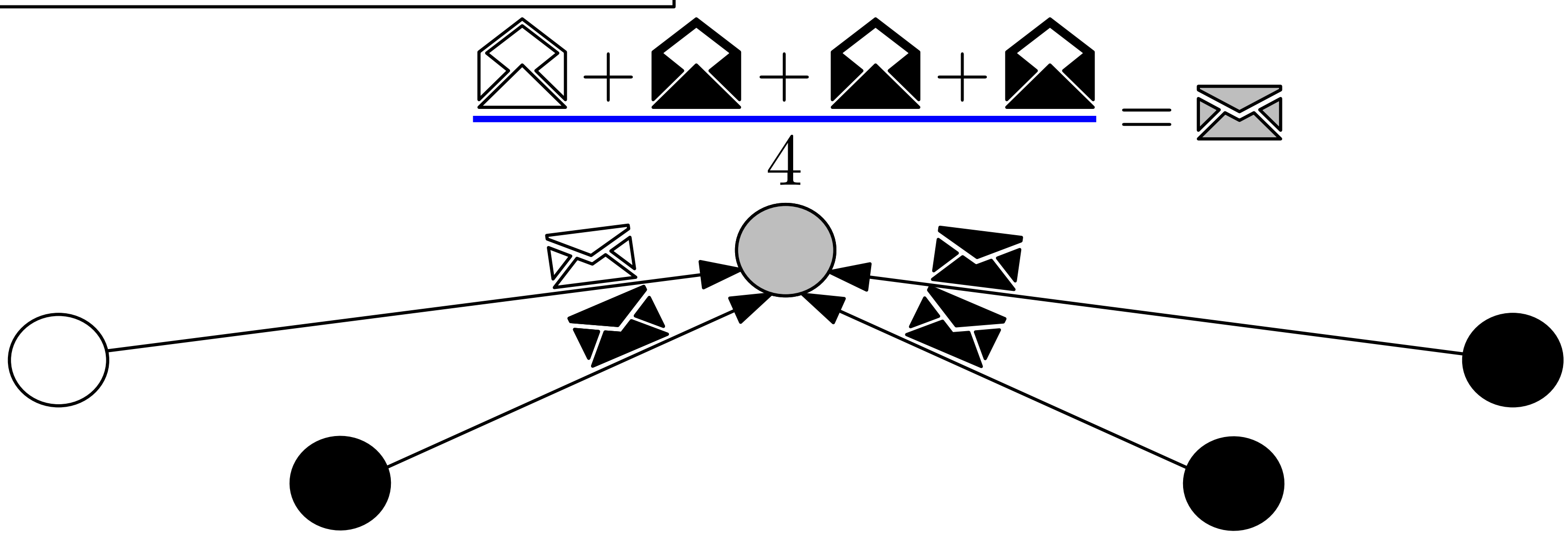


The Averaging Dynamics

Al nodes at the same time:

- At $t = 0$, randomly pick value $x^{(t)} \in \{\mathbf{+1}, \mathbf{-1}\}$.
- Then, at each round
  1. Set value $x^{(t)}$ to average of neighbors,
  2. Set label to **blue** if $x^{(t)} < x^{(t-1)}$, **red** otherwise.





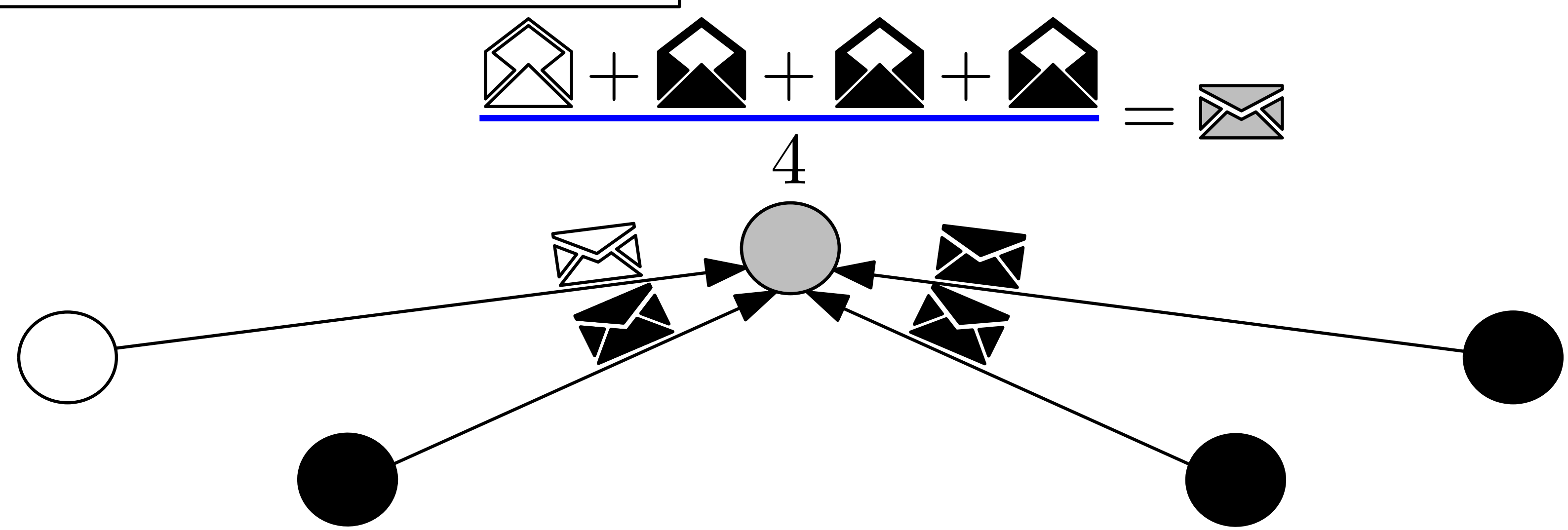
The Averaging Dynamics

Al nodes at the same time:

- At $t = 0$, randomly pick value $x^{(t)} \in \{\mathbf{+1}, \mathbf{-1}\}$.
- Then, at each round
  1. Set value $x^{(t)}$ to average of neighbors,
  2. Set label to **blue** if $x^{(t)} < x^{(t-1)}$, **red** otherwise.

Well studied process [Shah '09]:

- Converges to (weighted) global average of initial values,
- Convergence time = mixing time of G,
- Important applications in fault-tolerant self-stabilizing consensus.



The Averaging Dynamics

Al nodes at the same time:

- At $t = 0$, randomly pick value $x^{(t)} \in \{\mathbf{+1}, \mathbf{-1}\}$.
- Then, at each round
  1. Set value $x^{(t)}$ to average of neighbors,
  2. Set label to **blue** if $x^{(t)} < x^{(t-1)}$, **red** otherwise.

Well studied process [Shah '09]:

- Converges to (weighted) global average of initial values,
- Convergence time = mixing time of G,
- Important applications in fault-tolerant self-stabilizing consensus.

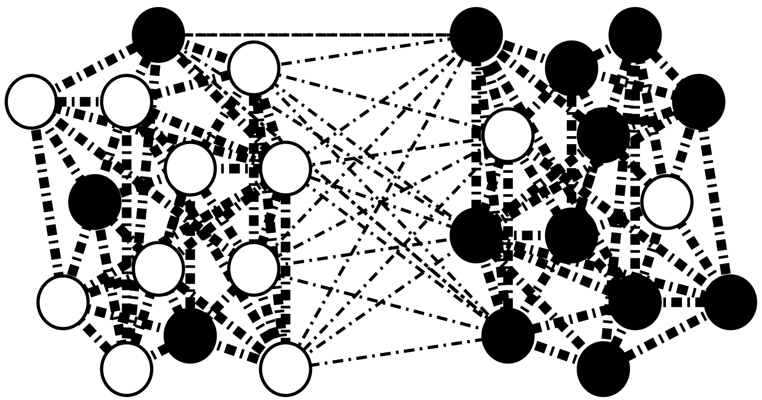
Averaging is a **linear** dynamics

$$\mathbf{x}^{(t)} = \begin{pmatrix} \circ \\ \bullet \\ \circ \\ \bullet \\ \bullet \end{pmatrix}$$

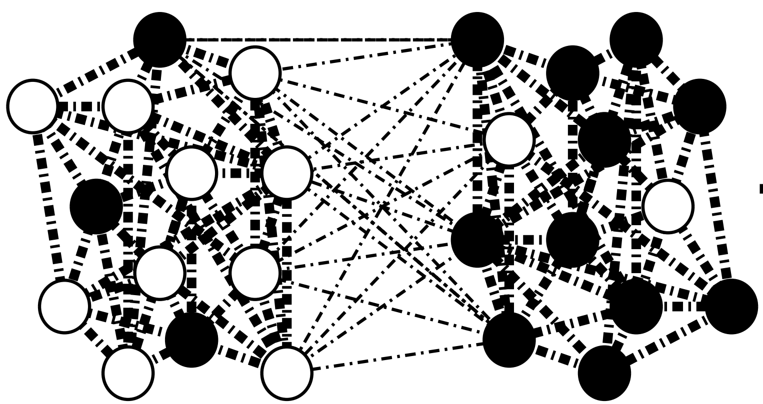
$$\boxed{\mathbf{x}^{(t)} = P \cdot \mathbf{x}^{(t-1)} = P^t \cdot \mathbf{x}^{(0)}}$$

P transition matrix of random walk

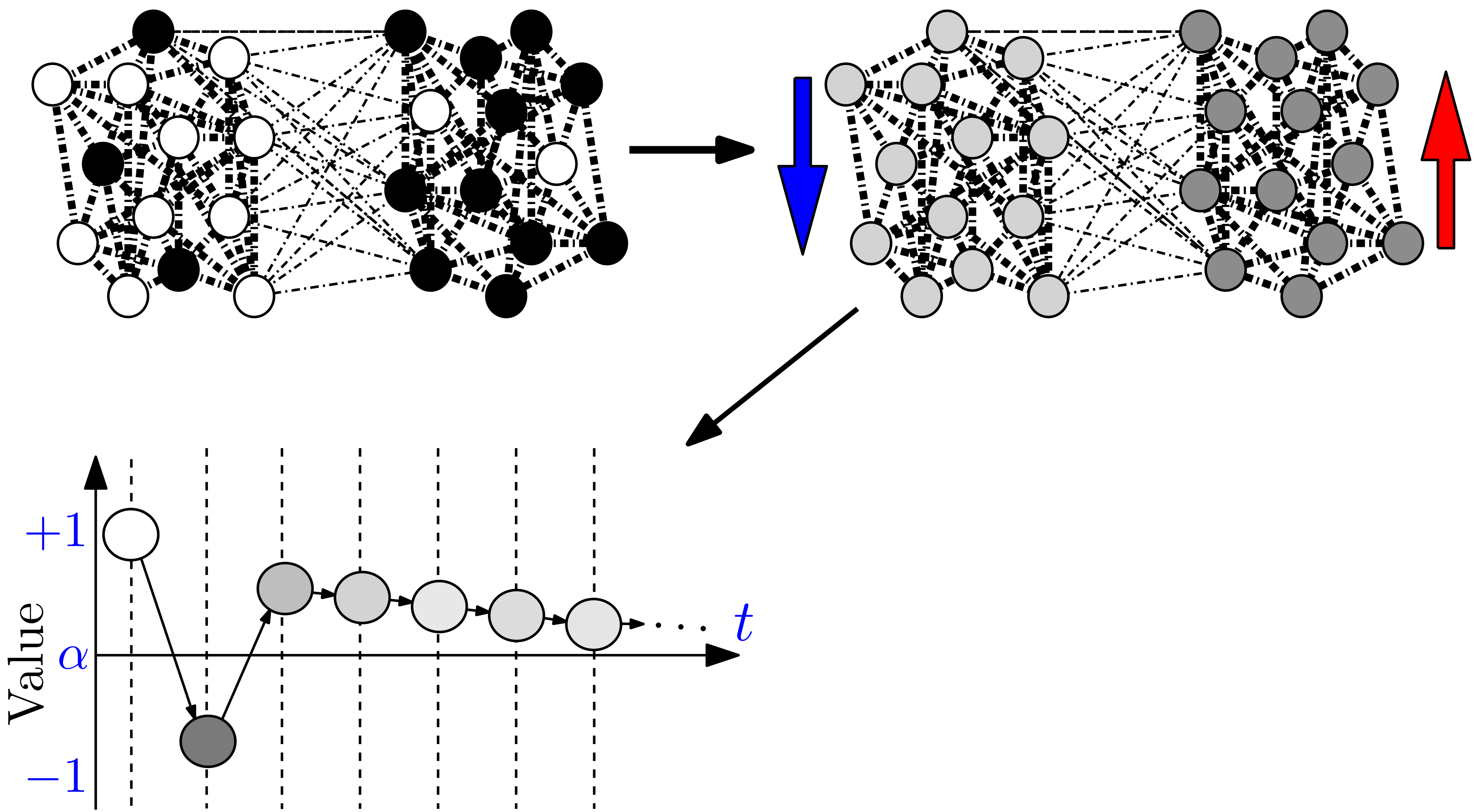
Community Detection via Averaging Dynamics



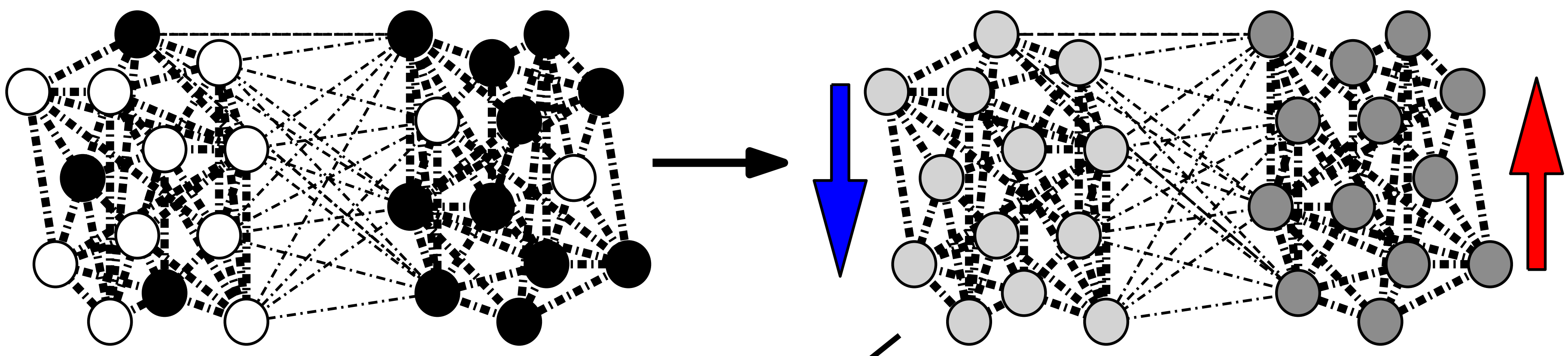
Community Detection via Averaging Dynamics

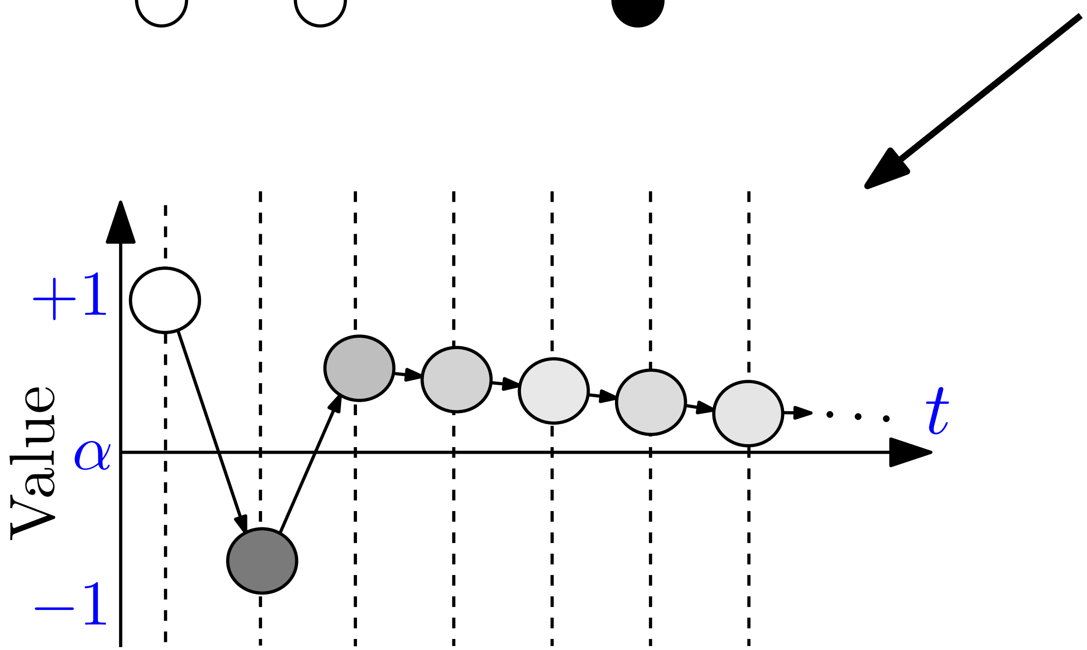


Community Detection via Averaging Dynamics



Community Detection via Averaging Dynamics





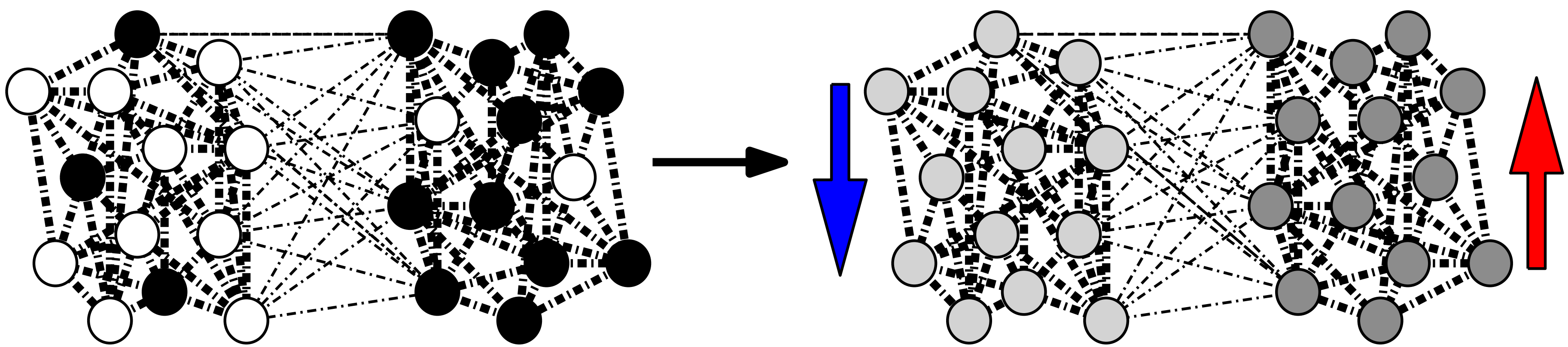
$v_1, \ldots, v_n$ eigenvectors of random walk matrix P:

$$v_1 = \mathbb{1} = (1, \ldots, 1)$$

$$v_2 \approx \chi = (1, \ldots, 1, -1, \ldots, -1)$$

“nice” graph

Community Detection via Averaging Dynamics



[SODA '17](Informal). $G = (V_1 \dot{\bigcup} V_2, E)$ s.t.
i) $\chi = \mathbf{1}_{V_1} - \mathbf{1}_{V_2}$ close to right-eigenvector of eigenvalue λ_2 of transition matrix of G, and
ii) gap between λ_2 and $\lambda = \max\{\lambda_3, |\lambda_n|\}$ sufficiently large,
then Averaging (approximately) identifies (V_1, V_2).

Thank you!

Analysis on Regular SBM

$$P \longrightarrow$$ symmetric $\implies$ orthonormal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and real eigenvalues $\lambda_1, \ldots, \lambda_n$.

Analysis on Regular SBM

$P \longrightarrow$ symmetric $\implies$ orthonormal eigenvectors $\mathbf{v}_1, ..., \mathbf{v}_n$ and real eigenvalues $\lambda_1, ..., \lambda_n$.

$$\mathbf{x}^{(t)} = P^t \cdot \mathbf{x}^{(0)} = \sum_i \lambda_i^t (\mathbf{v}_i^\intercal \mathbf{x}^{(0)}) \mathbf{v}_i$$

Analysis on Regular SBM

$P \longrightarrow$ symmetric $\Longrightarrow$ orthonormal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and real eigenvalues $\lambda_1, \ldots, \lambda_n$.

$$\mathbf{x}^{(t)} = P^t \cdot \mathbf{x}^{(0)} = \sum_i \lambda_i^t (\mathbf{v}_i^\intercal \mathbf{x}^{(0)}) \mathbf{v}_i$$

$$\mathbf{v}_1 = \frac{1}{\sqrt{n}} \mathbf{1}$$

Regular SBM $\Longrightarrow$ $P\chi = \left(\frac{a-b}{a+b}\right) \cdot \chi$

Analysis on Regular SBM

$P \longrightarrow$ symmetric $\Longrightarrow$ orthonormal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and real eigenvalues $\lambda_1, \ldots, \lambda_n$.

$$\mathbf{x}^{(t)} = P^t \cdot \mathbf{x}^{(0)} = \sum_i \lambda_i^t (\mathbf{v}_i^\intercal \mathbf{x}^{(0)}) \mathbf{v}_i$$

$$\mathbf{v}_1 = \frac{1}{\sqrt{n}} \mathbf{1}$$

Regular SBM $\Longrightarrow$ $P\chi = (\frac{a-b}{a+b}) \cdot \chi$

$$\frac{1}{a+b} \left(\begin{array}{c:c} \cdots\cdots\cdots & \cdots\cdots\cdots \\ \cdots a \text{ “1”s} \cdots & \cdots b \text{ “1”s} \cdots \\ \cdots\cdots\cdots & \cdots\cdots\cdots \\ \hdashline \cdots\cdots\cdots & \cdots\cdots\cdots \\ \cdots b \text{ “1”s} \cdots & \cdots a \text{ “1”s} \cdots \\ \cdots\cdots\cdots & \cdots\cdots\cdots \end{array} \right) \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \\ -1 \\ \vdots \\ -1 \end{pmatrix} = \frac{a-b}{a+b} \begin{pmatrix} 1 \\ \vdots \\ 1 \\ -1 \\ \vdots \\ -1 \end{pmatrix}$$

Analysis on Regular SBM

$P \longrightarrow$ symmetric $\Longrightarrow$ orthonormal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and real eigenvalues $\lambda_1, \ldots, \lambda_n$.

$\mathbf{x}^{(t)} = P^t \cdot \mathbf{x}^{(0)} = \sum_i \lambda_i^t (\mathbf{v}_i^\intercal \mathbf{x}^{(0)}) \mathbf{v}_i$

$\mathbf{v}_1 = \frac{1}{\sqrt{n}} \mathbf{1}$

Regular SBM $\Longrightarrow$ $P\chi = (\frac{a-b}{a+b}) \cdot \chi$

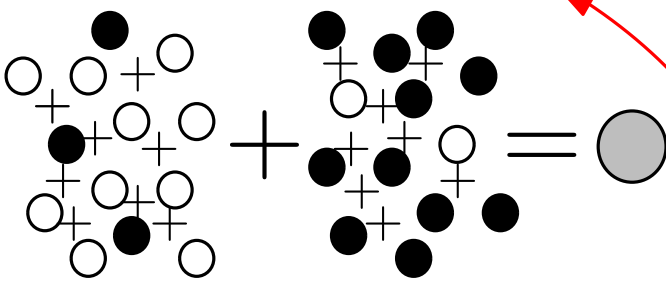
W.h.p. $\max\{\lambda_3, |\lambda_n|\}(1+\delta) < \frac{a-b}{a+b} = \lambda_2$, then

$$\mathbf{x}^{(t+1)} = \frac{1}{n}(\mathbf{1}^\intercal \mathbf{x}^{(0)})\mathbf{1} + \lambda_2^t \frac{1}{n}(\chi^\intercal \mathbf{x}^{(0)})\chi + \mathbf{e}^{(t)}$$

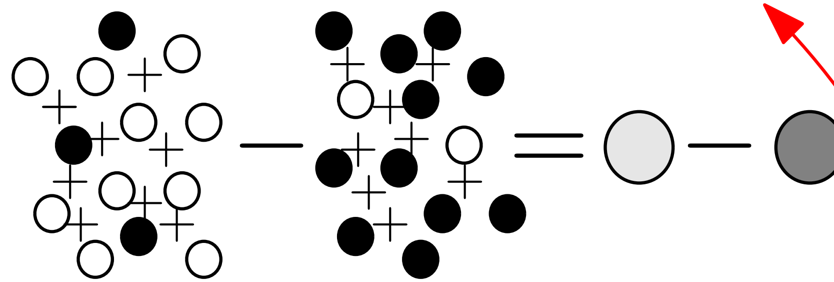
with $\|\mathbf{e}^{(t)}\| \leq (\max\{\lambda_3, |\lambda_n|\})^t \sqrt{n}$

Analysis on Regular SBM

$\frac{1}{n}\sum_{u \in V} \mathbf{x}^{(0)}(u)$



$\left(\frac{1}{n}\sum_{u \in V_1} \mathbf{x}^{(0)}(u) - \frac{1}{n}\sum_{u \in V_2} \mathbf{x}^{(0)}(u)\right)$



W.h.p. $\max\{\lambda_3, |\lambda_n|\}(1+\delta) < \frac{a-b}{a+b} = \lambda_2$, then

$$\mathbf{x}^{(t+1)} = \frac{1}{n}(\mathbf{1}^\intercal \mathbf{x}^{(0)})\mathbf{1} + \lambda_2^t \frac{1}{n}(\chi^\intercal \mathbf{x}^{(0)})\chi + \mathbf{e}^{(t)}$$

with $\|\mathbf{e}^{(t)}\| \leq (\max\{\lambda_3, |\lambda_n|\})^t \sqrt{n}$

Analysis on Regular SBM

$$\mathbf{x}^{(t)} = \frac{1}{n}(\mathbf{1}^\intercal \mathbf{x}^{(0)})\mathbf{1} + \lambda_2^t \frac{1}{n}(\chi^\intercal \mathbf{x}^{(0)})\chi + \mathbf{e}^{(t)}$$

Analysis on Regular SBM

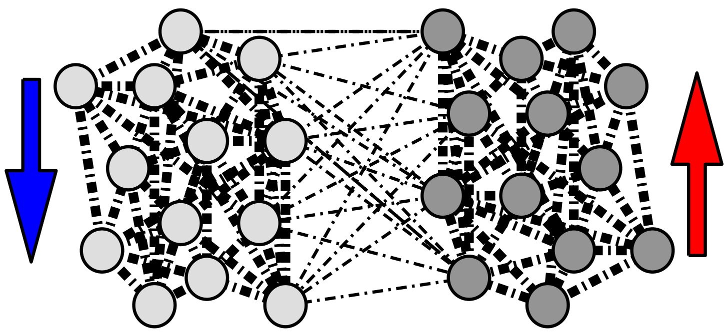
$$\mathbf{x}^{(t)} = \frac{1}{n}(\mathbf{1}^\intercal \mathbf{x}^{(0)})\mathbf{1} + \lambda_2^t \frac{1}{n}(\chi^\intercal \mathbf{x}^{(0)})\chi + \mathbf{e}^{(t)}$$

$$\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)} = (\chi^\intercal \mathbf{x}^{(0)})\lambda_2^{t-1}(\lambda_2 - 1)\chi + \underbrace{\mathbf{e}^{(t)} - \mathbf{e}^{(t-1)}}_{\ll \lambda_2^{t-1} \text{ if } t = \Omega(\log n)}$$

Analysis on Regular SBM

$$\mathbf{x}^{(t)} = \frac{1}{n}(\mathbf{1}^\intercal \mathbf{x}^{(0)})\mathbf{1} + \lambda_2^t \frac{1}{n}(\chi^\intercal \mathbf{x}^{(0)})\chi + \mathbf{e}^{(t)}$$

$$\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)} = (\chi^\intercal \mathbf{x}^{(0)})\lambda_2^{t-1}(\lambda_2 - 1)\chi + \underbrace{\mathbf{e}^{(t)} - \mathbf{e}^{(t-1)}}_{\ll \lambda_2^{t-1} \text{ if } t=\Omega(\log n)}$$



$$\boxed{\mathrm{sign}(\mathbf{x}^{(t)}(u) - \mathbf{x}^{(t-1)}(u)) \propto \mathrm{sign}(\chi(u))}$$