Computing through Simplicity: **Towards a Theory of Dynamics**

Emanuele Natale

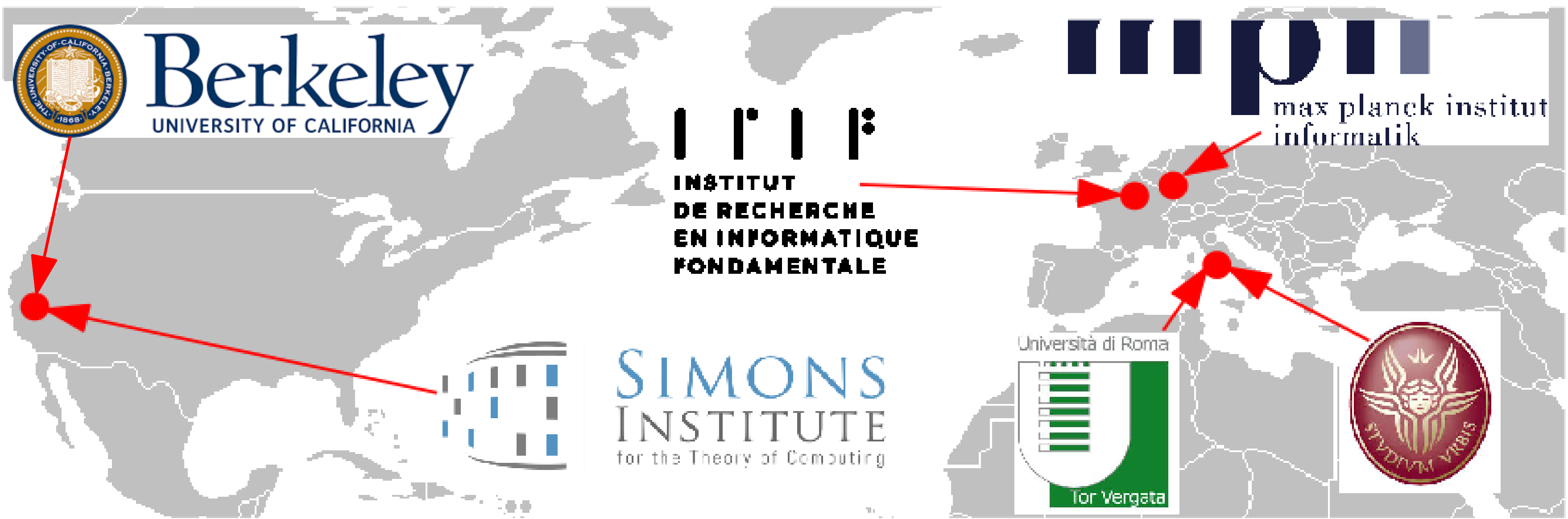
max planck institut informatik

Simons Institute for the Theory of Computing

Berkeley University of California

Academic Path

- 2017 - PhD in CS, Sapienza University
- 2014/15 (3 months)- Visiting IRIF
- 2016 & 2018-now (1 year) - **Fellow** of Simons Institute for the Theory of Computing (Brain Program)
- 2017-now - PostDoc, Max-Planck Institute for Informatics



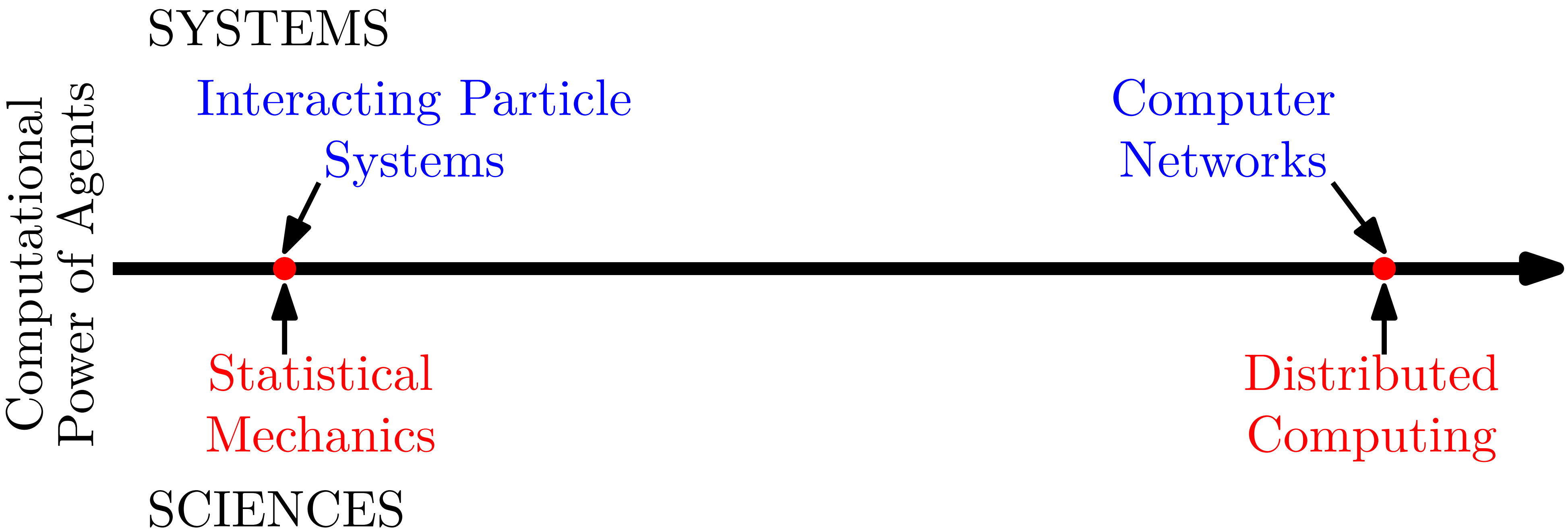
Awards

- ESA'16 **Best Student Paper**
- EATCS **Best Italian PhD Thesis** (Springer monograph in preparation)
- Sapienza 2016 **Outstanding PhD**
- Sapienza 2015 **Best CS Paper**

Invited speaker

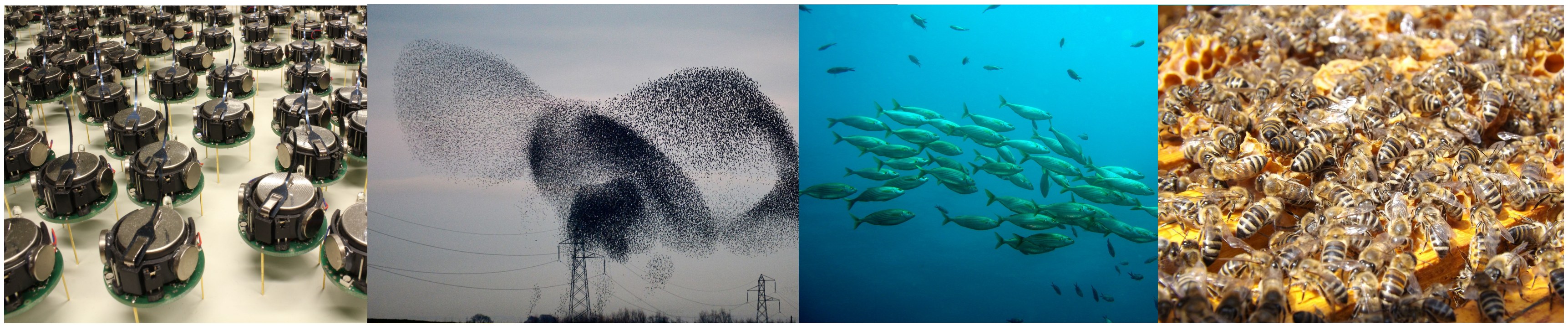
- ICTCS 2017
- Workshop on Random Processes in Discrete Structures, University of Warwick 2016

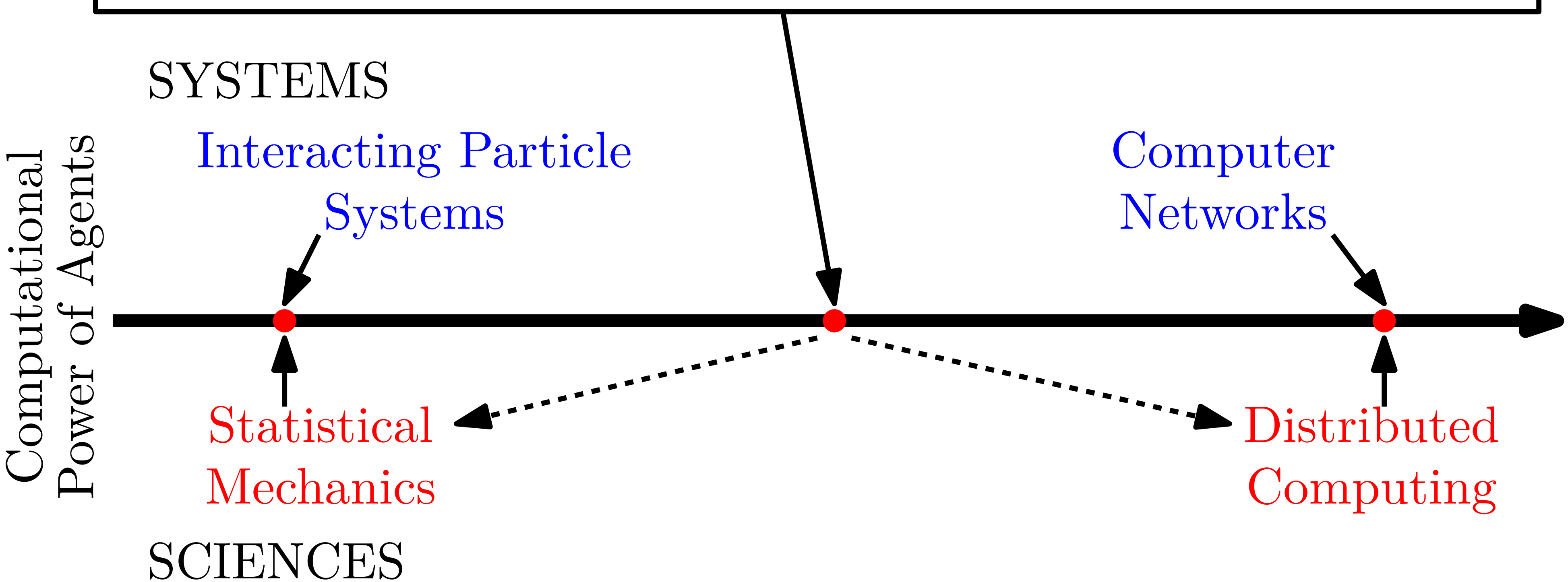
What can *Locally-Simple* Systems *Compute*?



What can *Locally-Simple* Systems *Compute*?

A **computational lens** on how
global behavior emerges from
simple local interactions among individuals





The Vision: A Theory of Dynamics

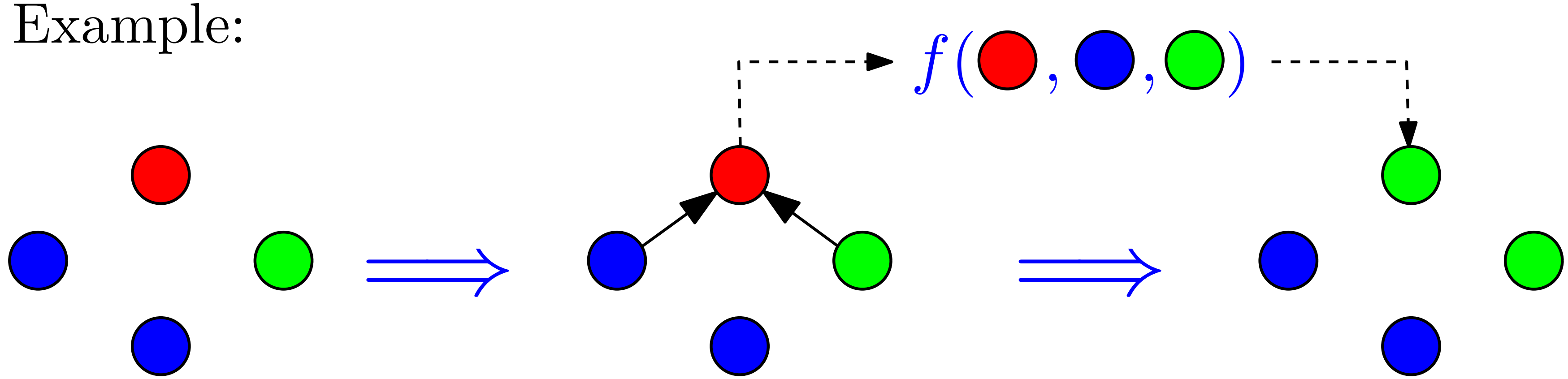
Dynamics [PhD Thesis]

Small set S and function f, agents have state $s \in S$ and

$$s^{new} = f(s^{old}, s^{\text{neighbor } u_1}, s^{\text{neighbor } u_2}, \ldots)$$

$\{u_1, u_2, \ldots\}$ typically random.

Example:



The Vision: A Theory of Dynamics

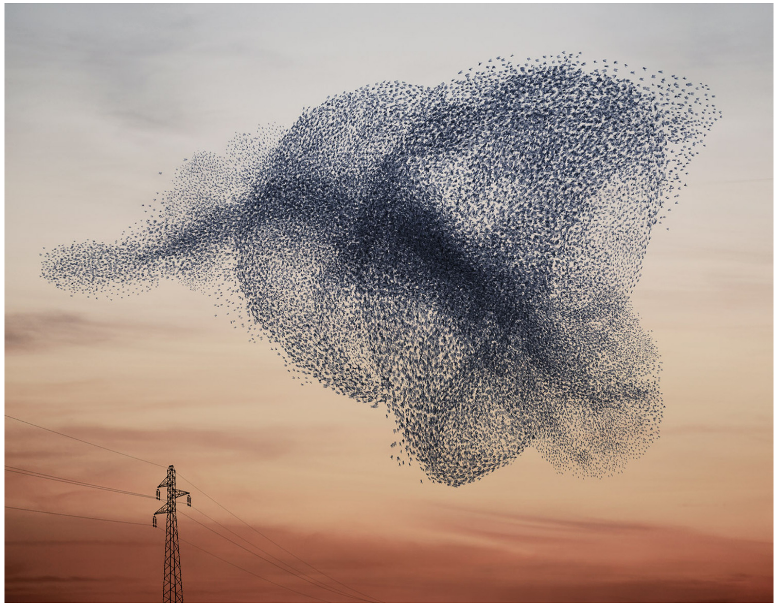
Dynamics [PhD Thesis]
Small set S and function f, agents have state $s \in S$ and

$$s^{new} = f(s^{old}, s^{\text{neighbor } u_1}, s^{\text{neighbor } u_2}, \dots)$$

$\{u_1, u_2, \dots\}$ typically random.

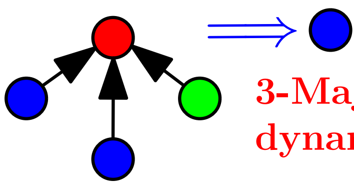
Wide applications in

- Collective behavior
- Neural networks (*Simons Institute, UC Berkeley*)
- Evolutionary dynamics (ecology)
- Many technological settings (dynamic networks)



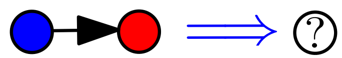
Towards the Project: Previous Work on Dynamics

Solve **majority consensus** **SPAA'14**



3-Majority dynamics

Solving majority consensus in $\mathcal{O}$(**monochromatic distance**) **SODA'15**



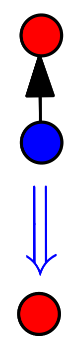
Undecided-State dynamics

General **noisy rumor-spreading** **PODC'16**

General **stabilizing consensus** **SODA'16**

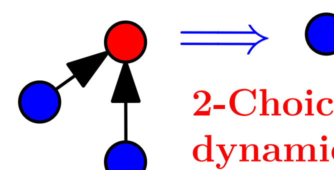
Relation between social-network structure in graph and evolution of dynamics **AAMAS'18**

Comparative techniques: use results on one dynamics to understand other dynamics **PODC'17**



Voter dynamics

Use dynamics for **clock synchronization** **SODA'17**



2-Choice dynamics

A Representative Result

Dynamics are shown
to solve difficult problems
in *distributed* settings

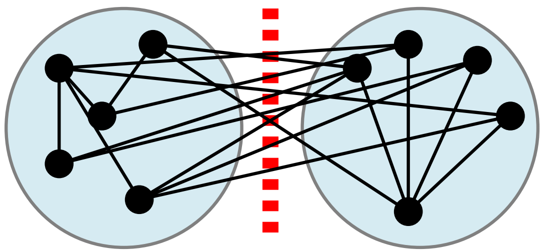
**Can we use dynamics
to solve difficult problems
even in *classical* settings?**

Start of Technical Part: Clustering

Minimum Bisection Problem

Find bipartition that minimizes cut

...NP-Complete!

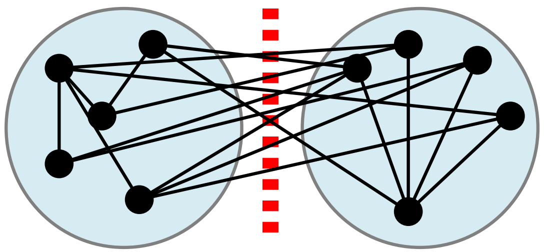


Start of Technical Part: Clustering

Minimum Bisection Problem

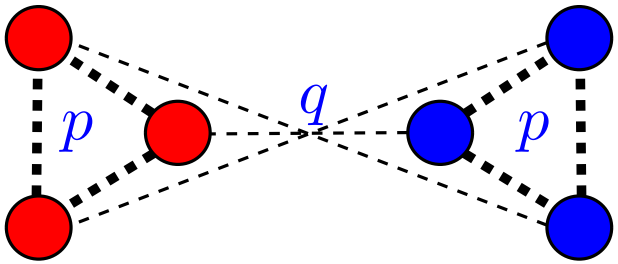
Find bipartition that minimizes cut

...NP-Complete!



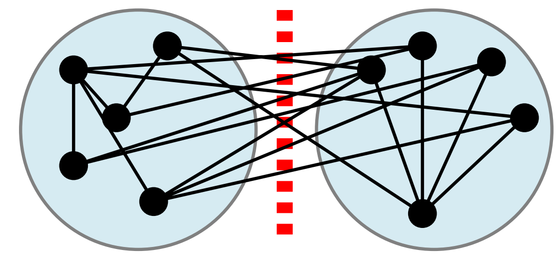
Stochastic Block Model

- "Communities" V_1, V_2, with $|V_1| = |V_2|$
- include each edge with probability
  - p if edge is inside V_1, V_2
  - q if edge is between V_1 and V_2



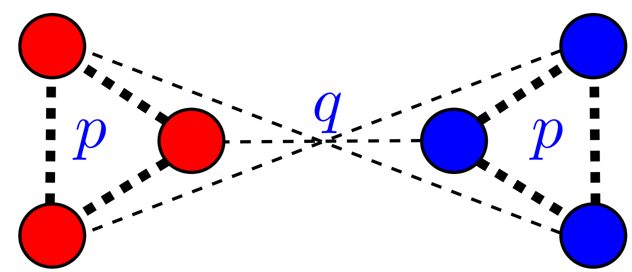
Start of Technical Part: Clustering

Minimum Bisection Problem
Find bipartition that minimizes cut
...NP-Complete!

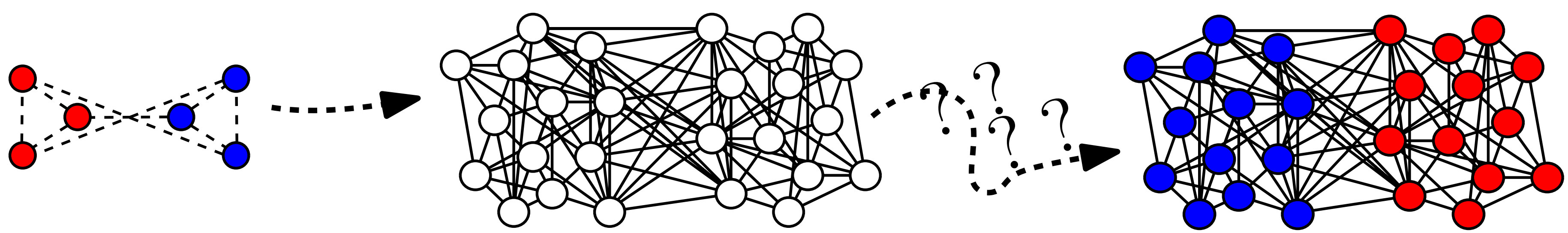


Stochastic Block Model

- "Communities" V_1, V_2, with $|V_1| = |V_2|$
- include each edge with probability
  - p if edge is inside V_1, V_2
  - q if edge is between V_1 and V_2



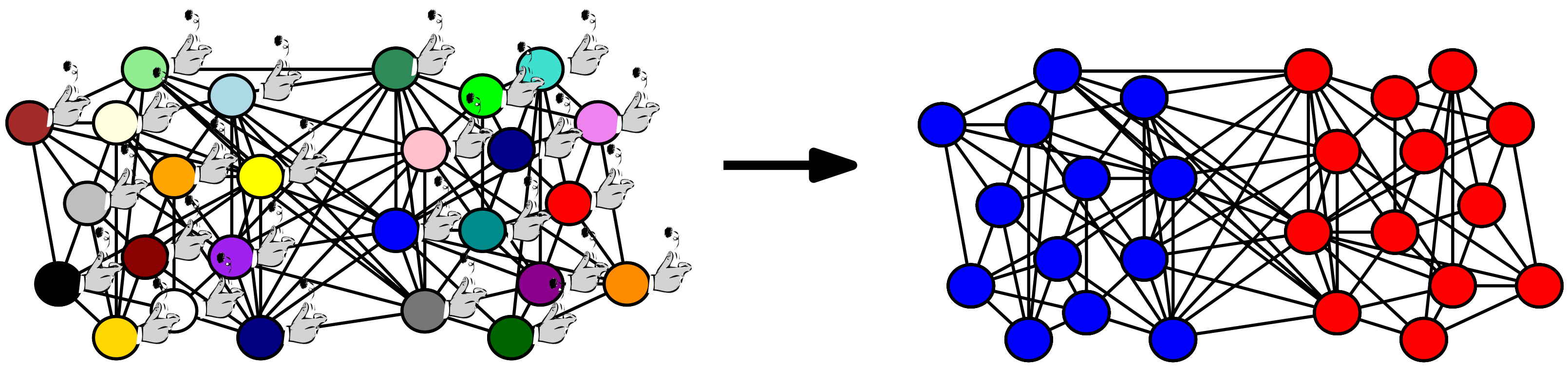
Clustering problem
Given graph generated by SBM, find original clusters



Known: clustering possible **if and only if** p and q in a precise regime

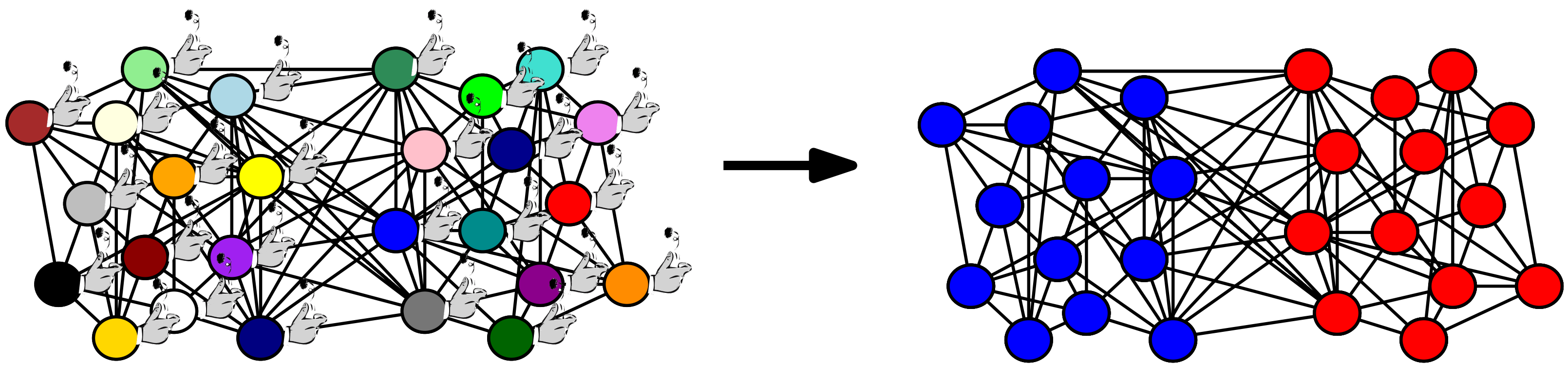
Clustering via **Averaging** [SODA'17]

Label **P**ropagation **A**lgorithms. Widely used heuristics in Data Mining: Each node initially gets random label, then updates label with simple rule of neighbors' states



Clustering via Averaging [SODA'17]

Label **P**ropagation **A**lgorithms. Widely used heuristics in Data Mining: Each node initially gets random label, then updates label with simple rule of neighbors' states



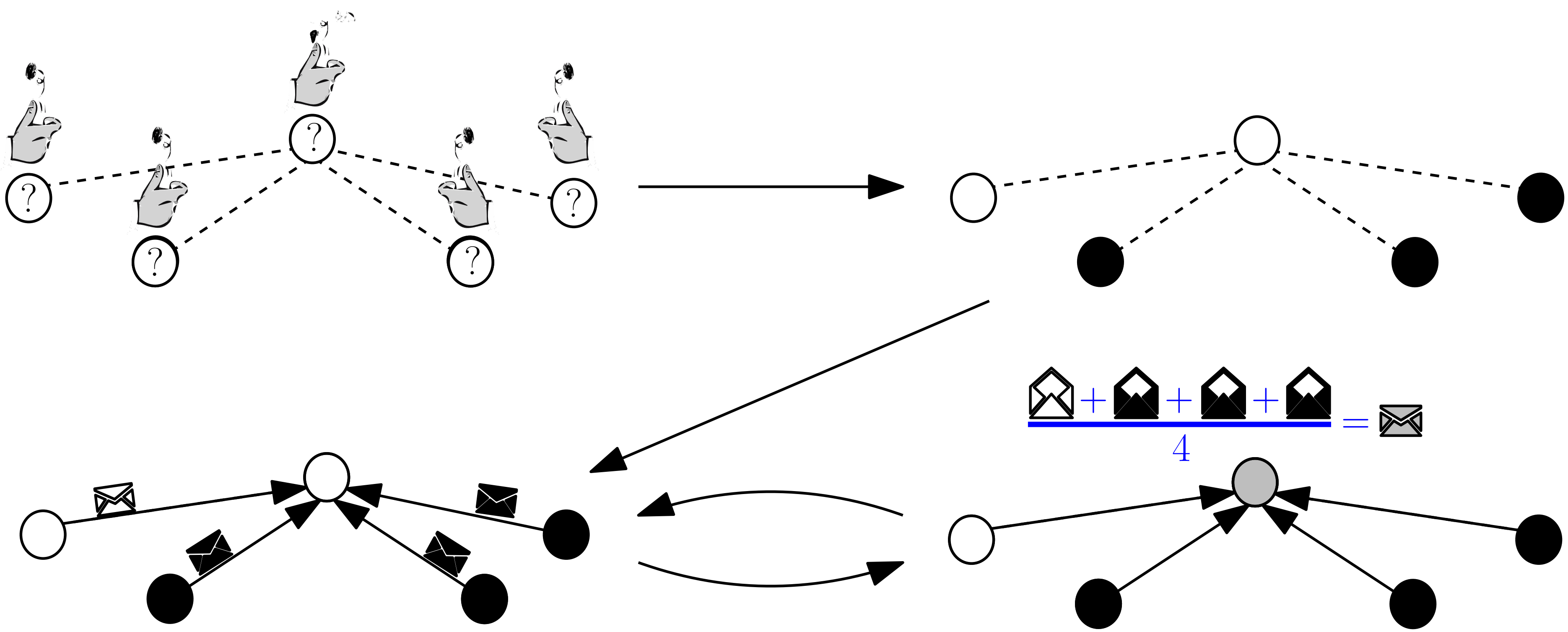
Theorem. There exists a dynamics that, in Stochastic Block Model with $p - q > \sqrt{const(p+q)/n} + \mathcal{O}(\log n/n)$, finds clusters with high probability after $\mathcal{O}(\frac{\log n}{\log \lambda_2/\lambda_3})$ steps, with $\mathcal{O}(\frac{p+q}{p-q})$ errors

- Improve state of the art in efficient distributed clustering
- Simulations show good performance in practice

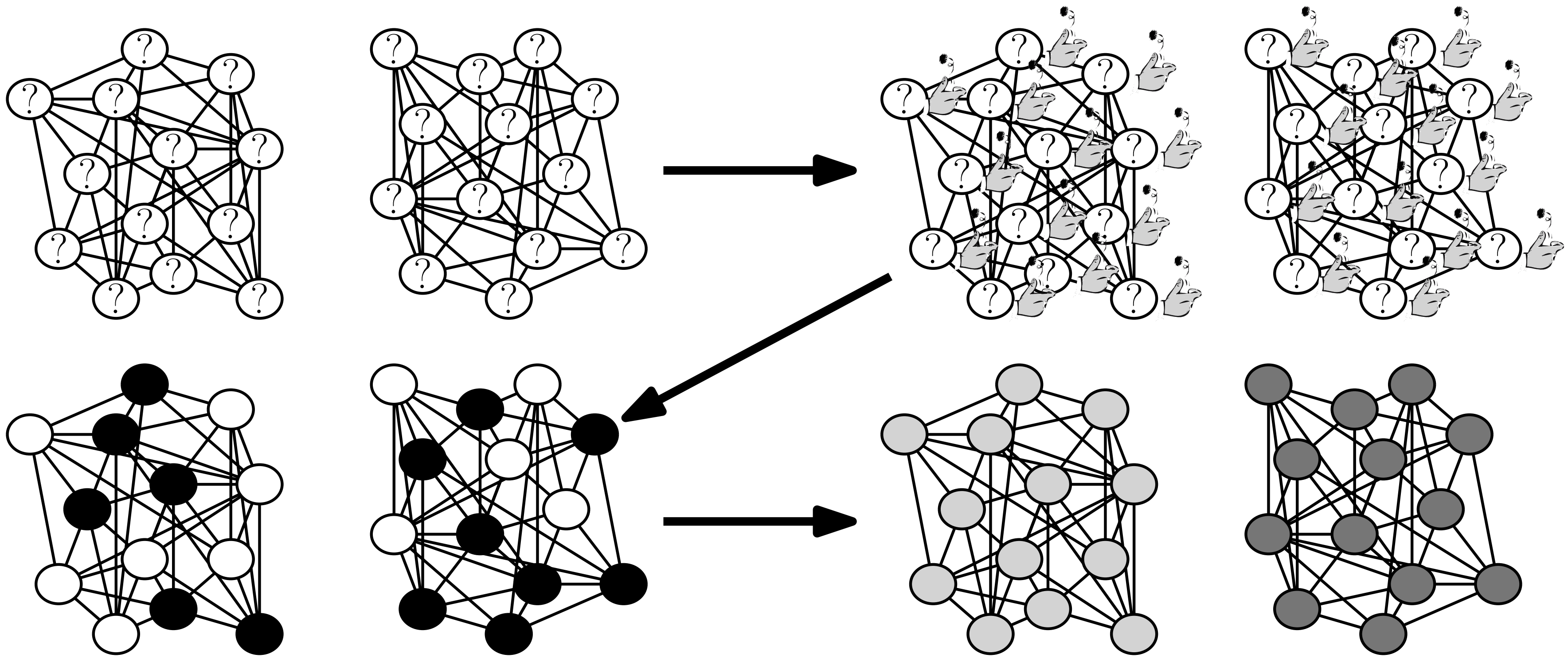
Averaging Dynamics [SODA'17]

All nodes at the same time:

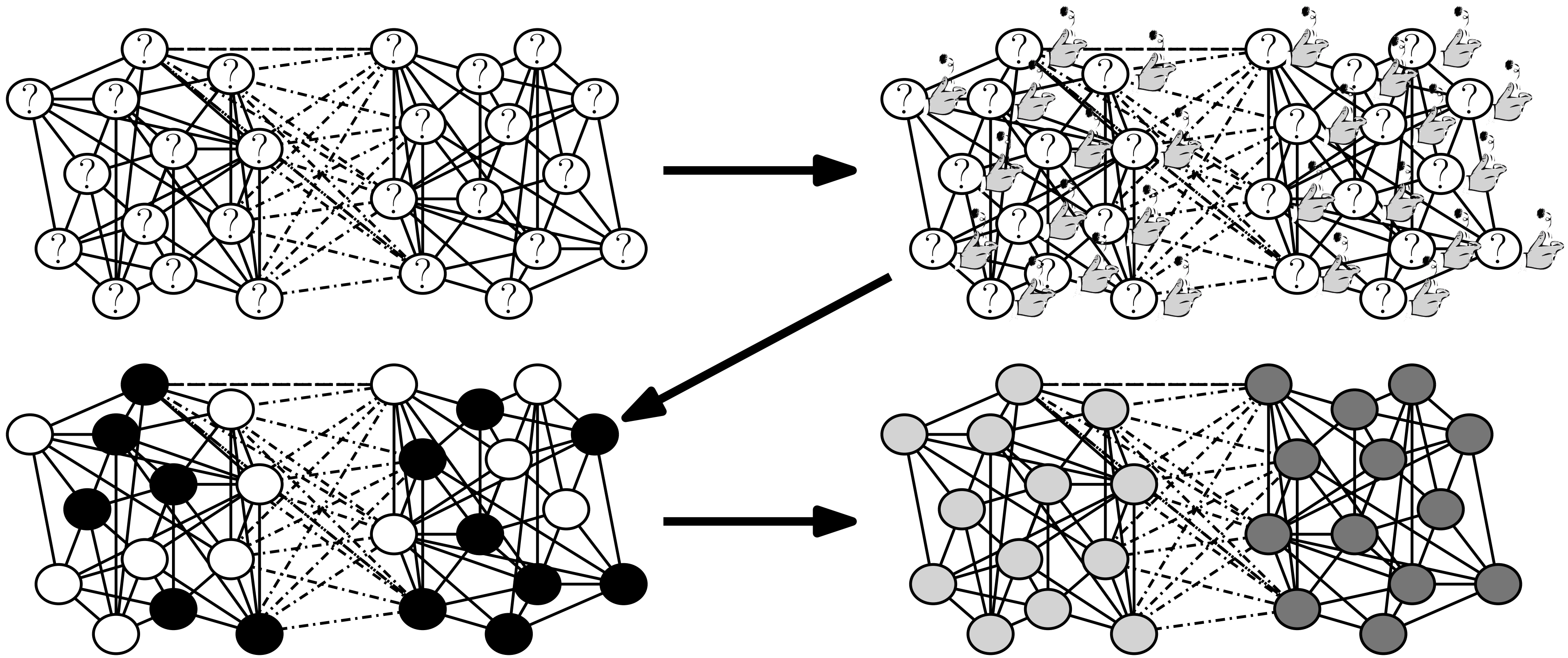
- At $t = 0$, randomly pick value $x^{(t)} \in \{\mathbf{+1}, \mathbf{-1}\}$
- Then, at each round
  set value $x^{(t)}$ to average of neighbors



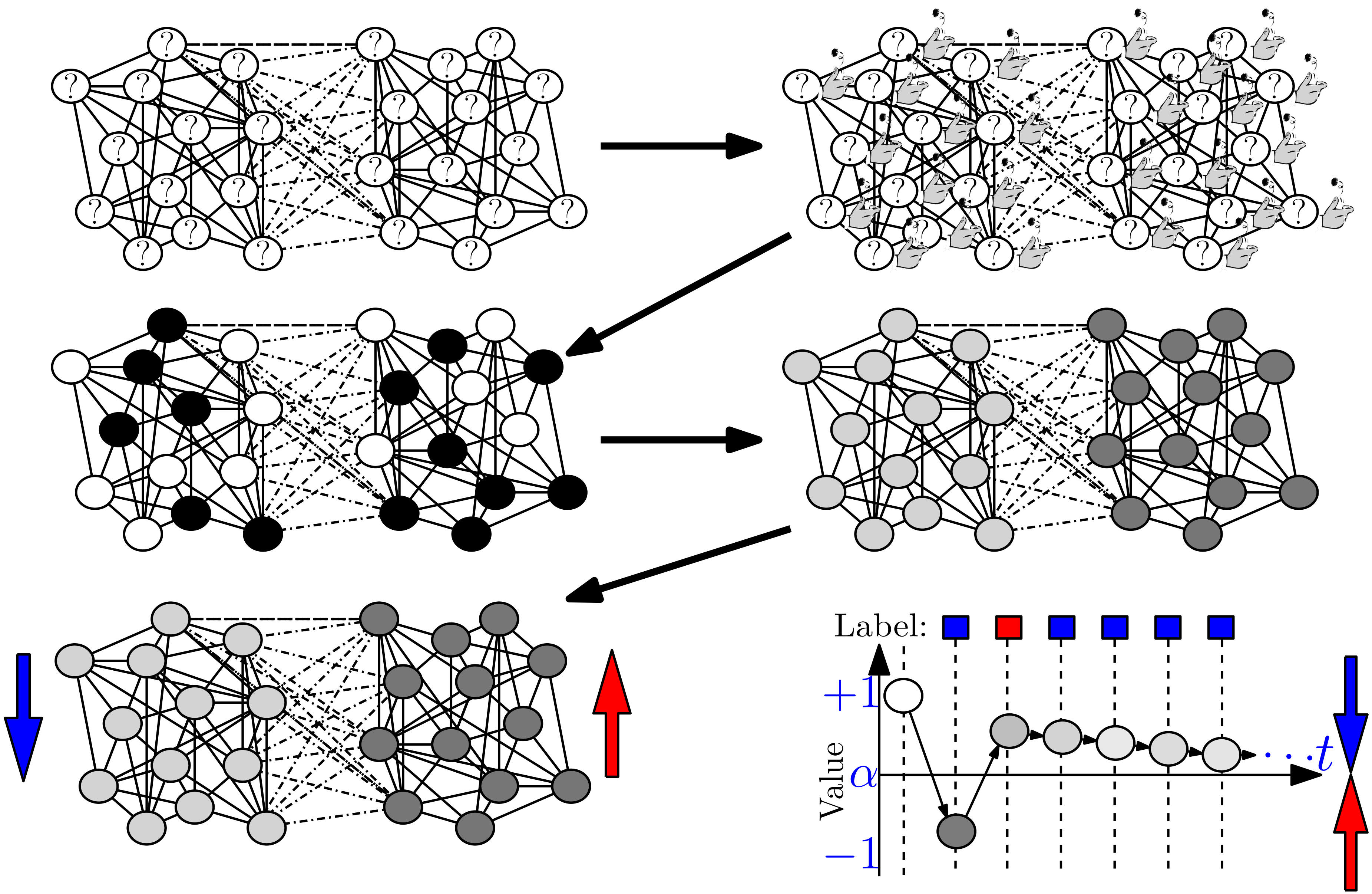
Why it Works: Intuition



Why it Works: Intuition



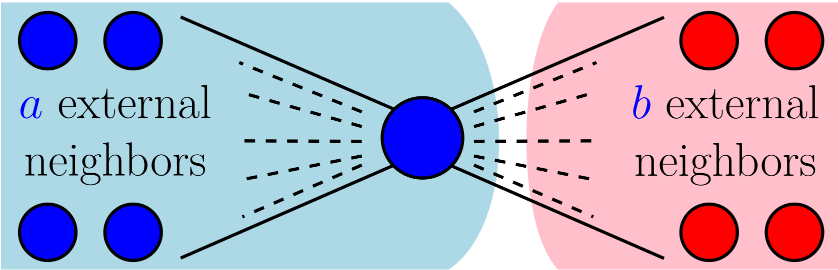
Why it Works: Intuition



- Set label to **blue** if $x^{(t)} < x^{(t-1)}$, **red** otherwise

End of Technical Part: Idea of Proof [SODA'17]

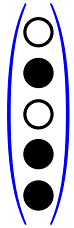
G instance of Regular Stochastic Block Model



Theorem. In Regular Stochastic Block Model with $a - b > \sqrt{2(a+b)}$, Averaging Dynamics finds clusters after $\frac{\log n}{\log \lambda_2/\lambda_3}$ steps with high probability

Averaging is a **linear** dynamics:

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

$\mathbf{x}^{(t)} =$ 

P transition matrix of random walk on G

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

End of Technical Part: Idea of Proof [SODA'17]

2nd eigenvalue $\gg \lambda_3(1+\delta)$ w.h.p.

negligible after

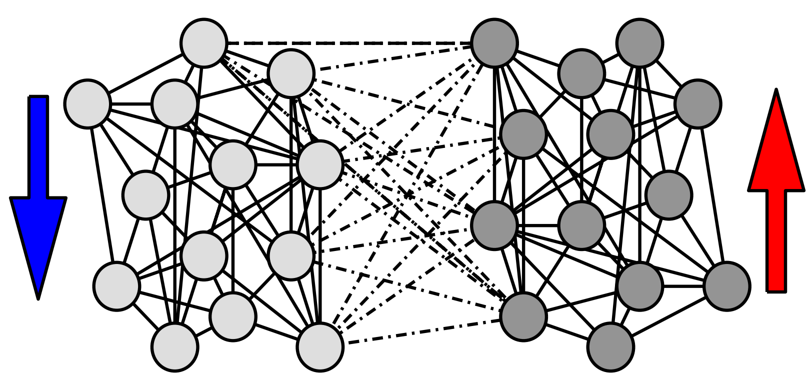
$t \gg \frac{\log n}{\log \lambda_2/\lambda_3}$

$$\mathbf{x}^{(t)} = \frac{1}{n}\mathbf{1}^\intercal \mathbf{x}^{(0)}\mathbf{1} + \left(\frac{a-b}{a+b}\right)^t \frac{1}{n}\chi^\intercal\mathbf{x}^{(0)}\chi + \mathbf{e}^{(t)}$$

1st eigenvector: all ones

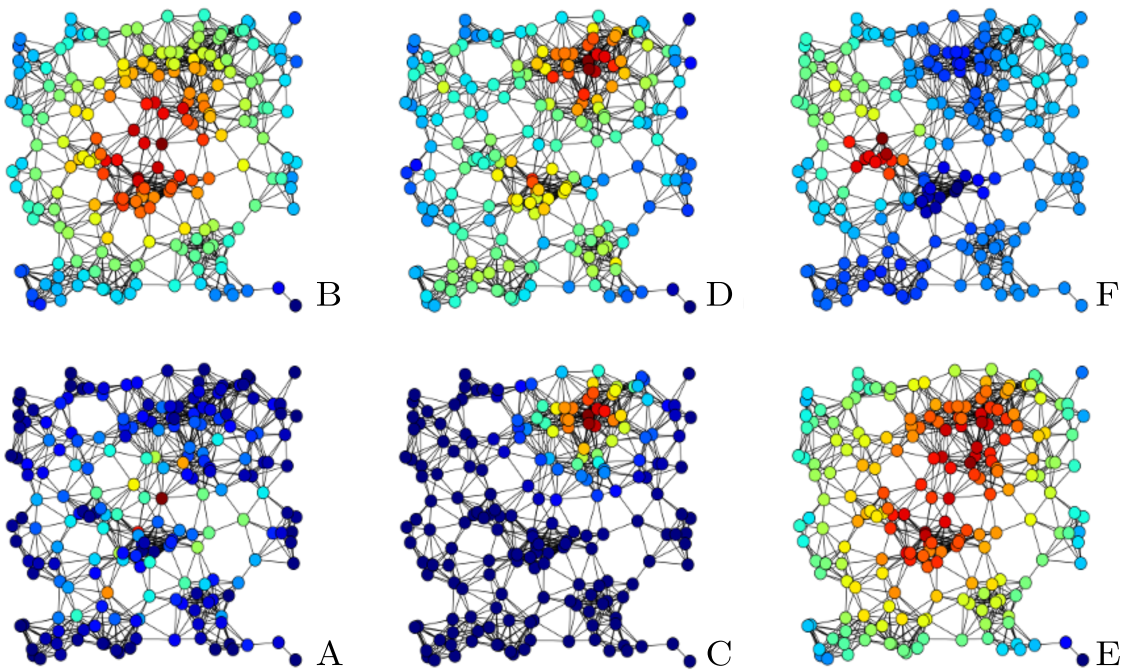
$(1, \dots, 1, -1, \dots, -1)$

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



QED

Theory-Driven Algorithm Engineering



Examples of A) Betweenness centrality, B) Closeness centrality, C) Eigenvector centrality, D) Degree centrality, E) Harmonic Centrality and F) Katz centrality of the same graph.

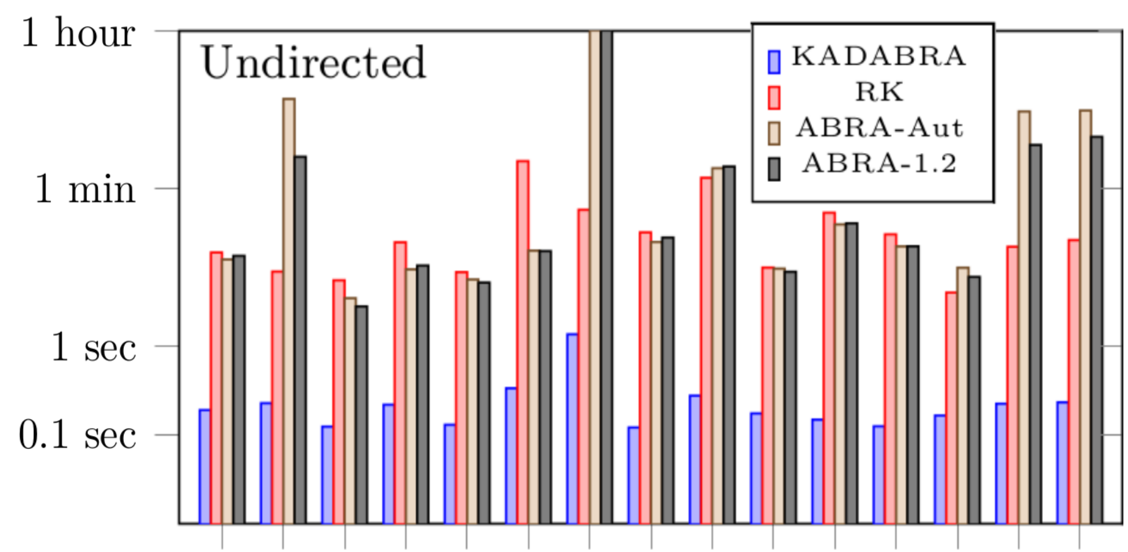
Betweenness centrality of u: Probability to transit through u when
- sample nodes s, t and
- sample s-t-shortest path.

!!! Betweenness is fundamental but (assuming SETH), hard to compute: must use randomized algorithms

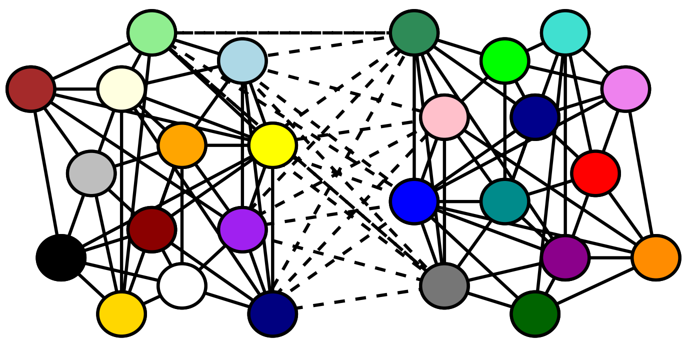
Previous state-of-art: ABRA [KDD'16].

We *proved* new results on
- Bidirectional Balanced BFS and
- Adaptive Sampling,

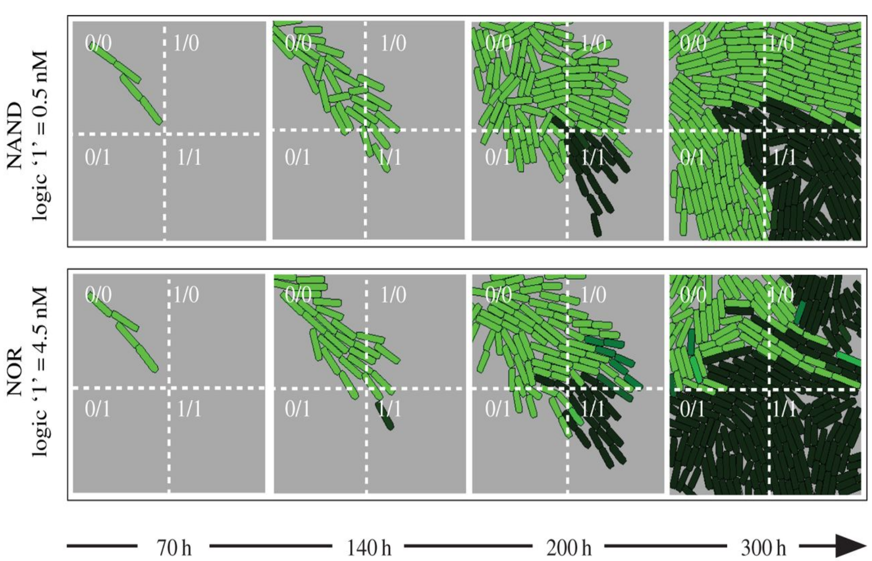
and developed KADABRA:



The Future with GANG



- Computing in **Dynamic Networks**
  - **L**abel **P**ropagation **A**lgorithms made rigorous
  - Computation through dynamic processes (A. Kosowski)



- Biological experiment *from* TCS predictions
  - algorithmic analysis of biological systems: ants, bacteria, molecules... (A. Korman, A. Kosowski)

The Far Future

- From theoretical neuroscience to neuromorphic machine learning
  - Simons Institute, Brain Program
  - Comparative compuational neuroscience (A. Korman)
  - New approaches to neural networks (S. Denéve (ENS))

- Engineering a CYBENT
  - Reverse-engineering neural networks: *understanding* neural code of insects
  - Hacking the neural code: *developing interfaces* to control insects (neural recording and stimulation)

Merci pour votre attention

Conference (16)

⋆16 Cruciani†, E.; N., E.; Nusser, A.; Scornavacca†, G. **AAMAS'18**

⋆15 Becchetti L., Bonifaci V., N. E. **AAMAS'18** † = supervised interns

⋆14 Boczkowski, L.; Korman, A.; N. E. **ITCS'18**

⋆13 Becchetti, L.; Clementi, A.; N., E.; Pasquale, F.; Trevisan, L. **SODA'17**

⋆12 Boczkowski, L.; Korman, A.; N., E. **SODA'17**

⋆11 Berenbrink, P.; Clementi, A.; Elsässer, R.; Kling, P.; Mallmann-Trenn, F.; N., E. **PODC'17**.

⋆10 Becchetti, L.; Clementi, A.; N., E.; Pasquale, F.; Trevisan, L.. **SODA'16**

9 Gualà, L.; Leucci, S.; N. E.; Tauraso, R. **FUN'16**

8 Borassi, M.; N. E. **ESA'16 (Best Student Paper Award)**

⋆7 Kaaser, D.; Mallmann-Trenn, F.; N., E. **MFCS'16**

⋆6 Fraigniaud, P.; N., E. **PODC'16**

⋆5 Becchetti, L.; Clementi, A.; N., E.; Pasquale, F.; Silvestri, R. **SODA'15** **(Best 2015 PhD Paper in CS at Sapienza Award)**

4 Becchetti, L.; Clementi, A.; N. E.; Pasquale, F.; Posta, G. **SPAA'15**

3 Gualà, L.; Leucci, S.; N. E. Bejeweled, **CIG'14**

⋆2 Becchetti, L.; Clementi, A.; N., E.; Pasquale, F.; Silvestri, R.; Trevisan, L. **SPAA'14**

⋆1 Clementi, A.; Di Ianni, M.; Gambosi, G.; N. E.; Silvestri, R. **SIROCCO'13**

Red text= Important/A^* conference

⋆ item = Related to Dynamics

Blue box = New

authors in alphabetical order

Journal (5)

⋆5 Fraigniaud P., N. E. **Distributed Computing**

⋆4 Boczkowski L., Korman A., N. E. **Distributed Computing 2018**

3 Becchetti, L.; Clementi, A.; N. E.; Pasquale, F.; Posta, G. **Distributed Computing 2017**

⋆2 Becchetti, L.; Clementi, A.; N. E.; Pasquale, F.; Silvestri, R.; Trevisan, L. **Distributed Computing 2017**

1 Clementi, A.; Di Ianni, M.; Gambosi, G.; N. E.; Silvestri, R. **Theor. Comp. Scie. 2015 (Special Issue)**

Submitted (2 conference, 2 journal)

⋆4 Boczkowski L., Feinerman O., Korman A., N. E. Submitted revision **PLOS Computational Biology**

⋆3 Becchetti, L.; Clementi, A.; Manurangsi, P.; N. E.; Pasquale, F.; Raghavendra, P.; Trevisan, L. **Sub. ESA**

⋆2 Clementi, A.; Gualà, L.; Pasquale, F.; Scornavacca, G.; N. E.; Ghaffari, M. **Submitted to MFCS**

⋆1 Borassi, M.; N. E.. **Submitted to J. of Experimental Algorithms (Special Issue)**