An introduction to network analysis and modeling with applications to social contagion processes

Part I

Juan G. Restrepo
University of Colorado at Boulder

June 28, 2022
ICERM

Talk outline

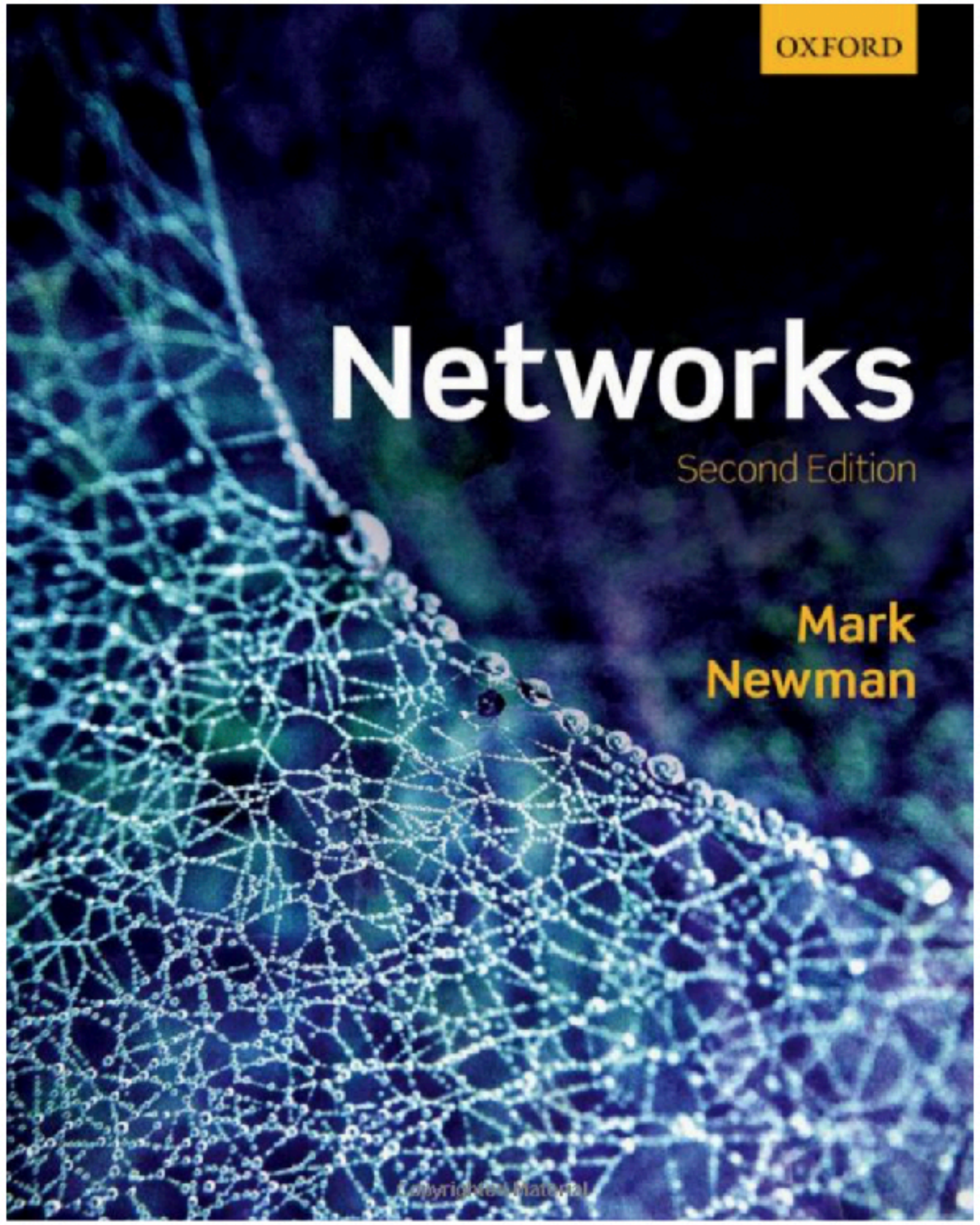
Part 1

- Motivation: how are **networks useful** for **social justice** analysis?
- Basic network **properties**
- Network **models**
- Beyond networks

Part 2

- Examples

This tutorial will review concepts covered in depth, for example, in Newman's Networks book:



Talk outline

Part 1

- Motivation: how are networks useful for social justice analysis?
- Basic network **properties**
- Network **models**
- Beyond networks

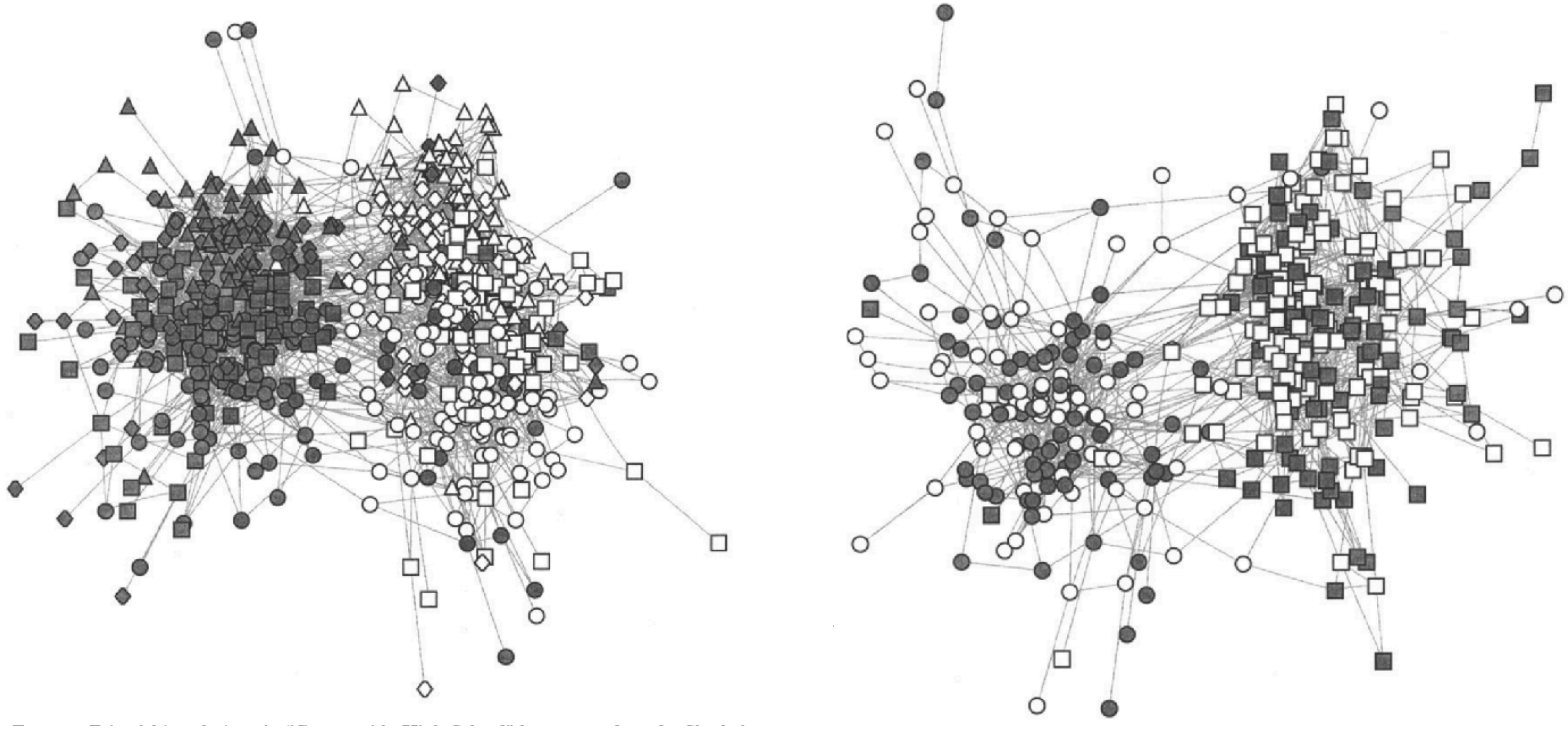
Part 2

- Examples

Networks: motivation for social justice analysis

Example: friendship network in middle and high schools: mixing by **race** and **grade**

At first sight there seems to be segregation on the left school and not on the right. Is this the case, or **does it depend on how the network is drawn**?



Symbol colors represent **race**. Symbol shapes represent **grade**.

James Moody, American Journal of Sociology, v 107, Number 3, pp. 679-716 (2001).

Networks: motivation for social justice analysis

Quantifying relationships can make biases and prejudices evident

Example: prestige in faculty hiring networks

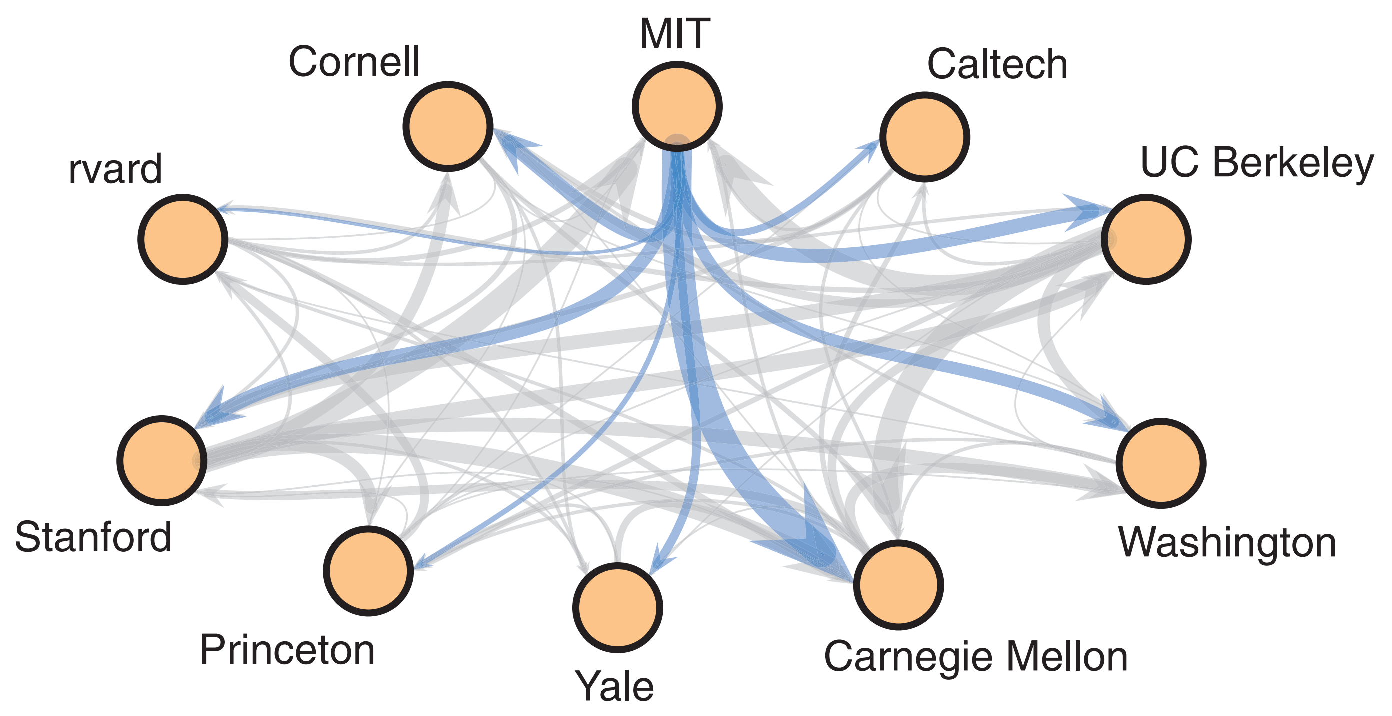
RESEARCH ARTICLE

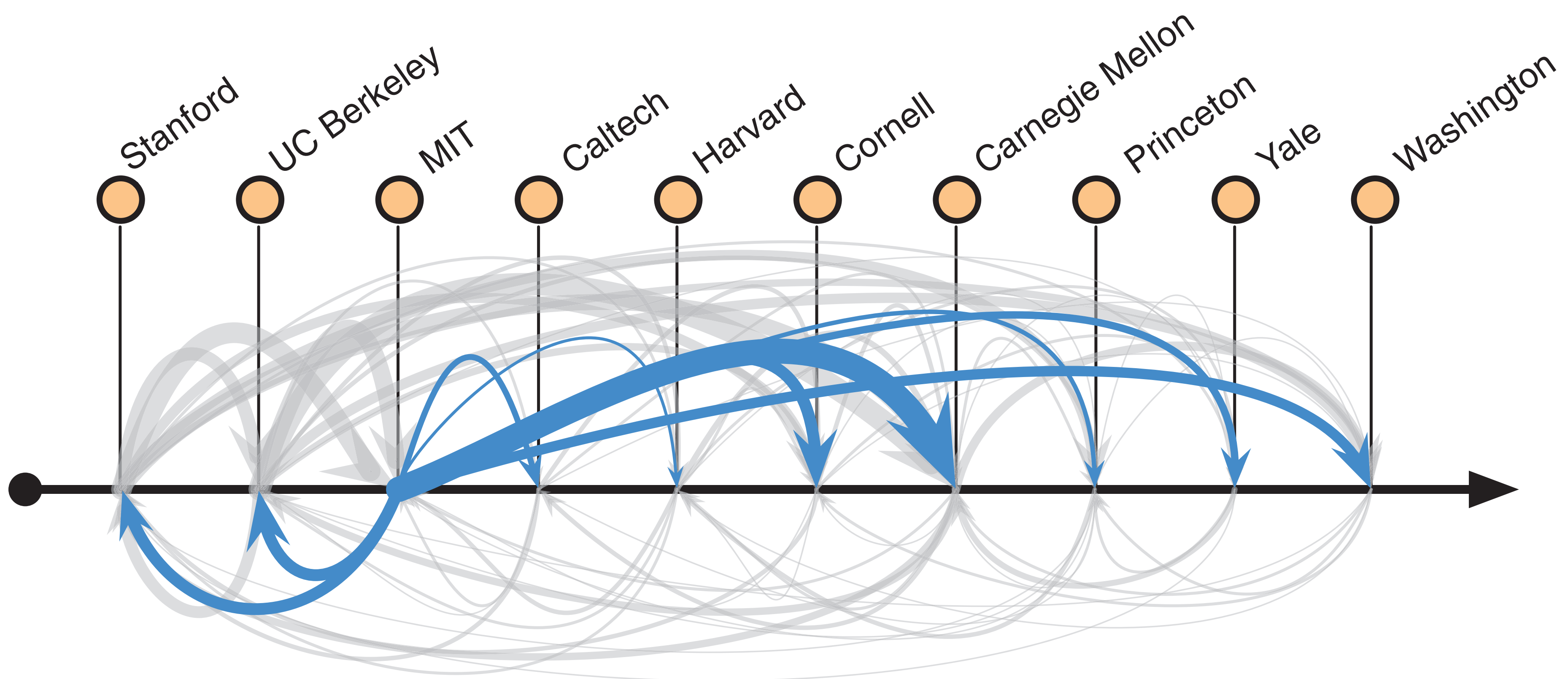
NETWORK SCIENCES

Systematic inequality and hierarchy in faculty hiring networks

Aaron Clauset,[1,2,3]* Samuel Arbesman,[4] Daniel B. Larremore[5,6]

The faculty job market plays a fundamental role in shaping research priorities, educational outcomes, and career trajectories among scientists and institutions. However, a quantitative understanding of faculty hiring as a system is lacking. Using a simple technique to extract the institutional prestige ranking that best explains an observed faculty hiring network—who hires whose graduates as faculty—we present and analyze comprehensive placement data on nearly 19,000 regular faculty in three disparate disciplines. Across disciplines, we find that faculty hiring follows a common and steeply hierarchical structure that reflects profound social inequality. Furthermore, doctoral prestige alone better predicts ultimate placement than a *U.S. News & World Report* rank, women generally place worse than men, and increased institutional prestige leads to increased faculty production, better faculty placement, and a more influential position within the discipline. These results advance our ability to quantify the influence of prestige in academia and shed new light on the academic system.





Talk outline

Part 1

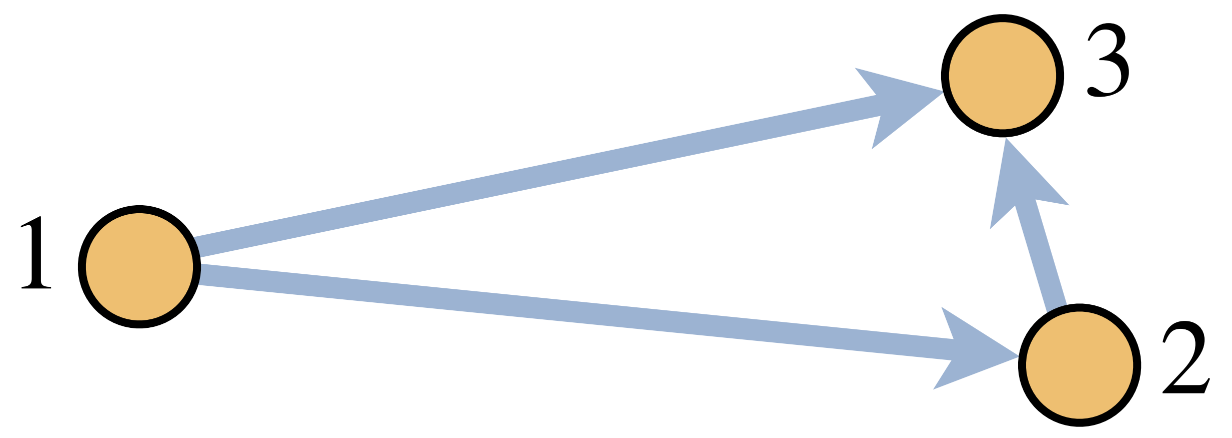
- ✓ Motivation: how are networks useful for social justice analysis?
- Basic network **properties**
- Network **models**
- Beyond networks

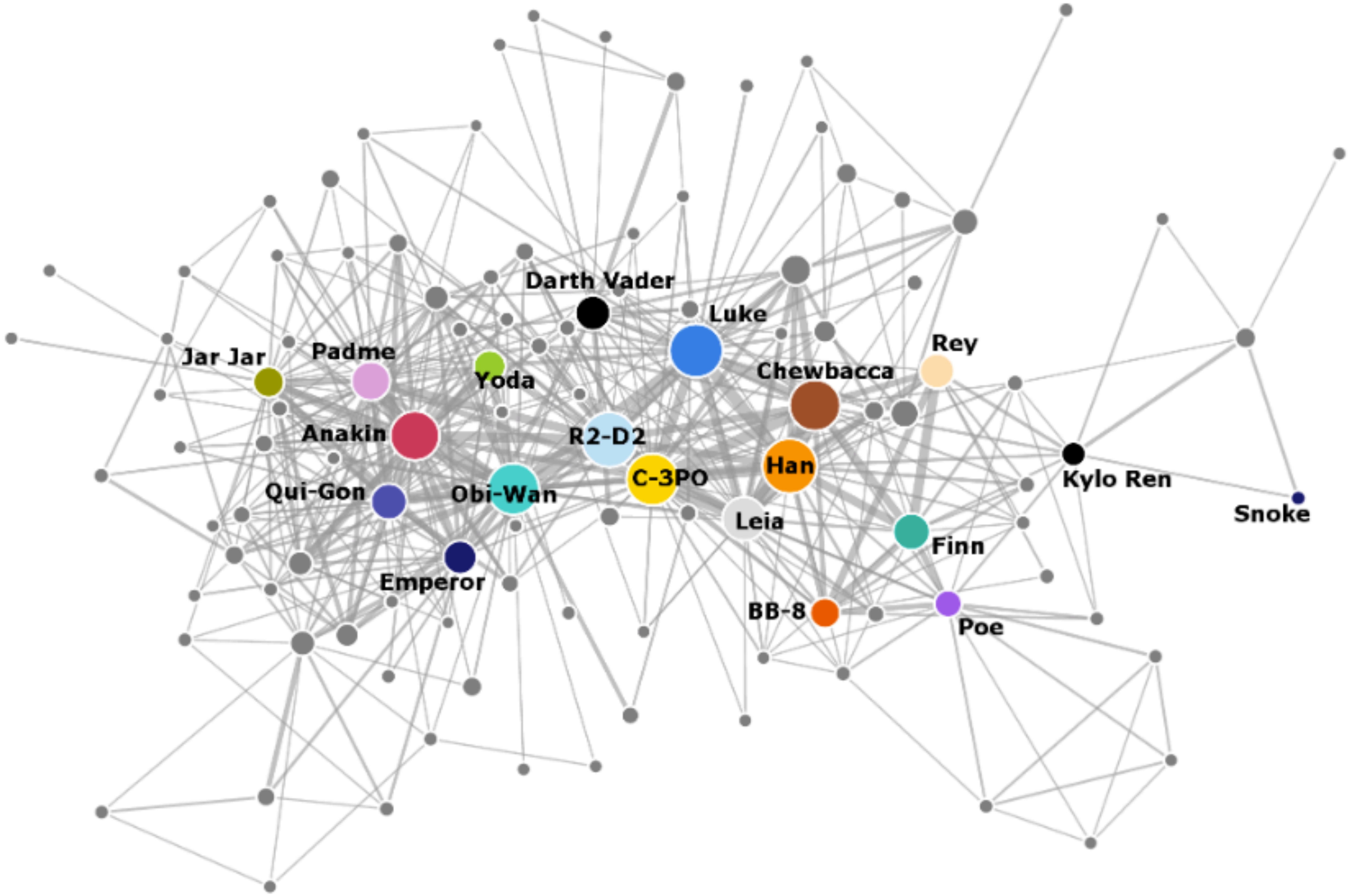
Part 2

- Examples

Networks: the basics

A network is a collection of **nodes** and **links**



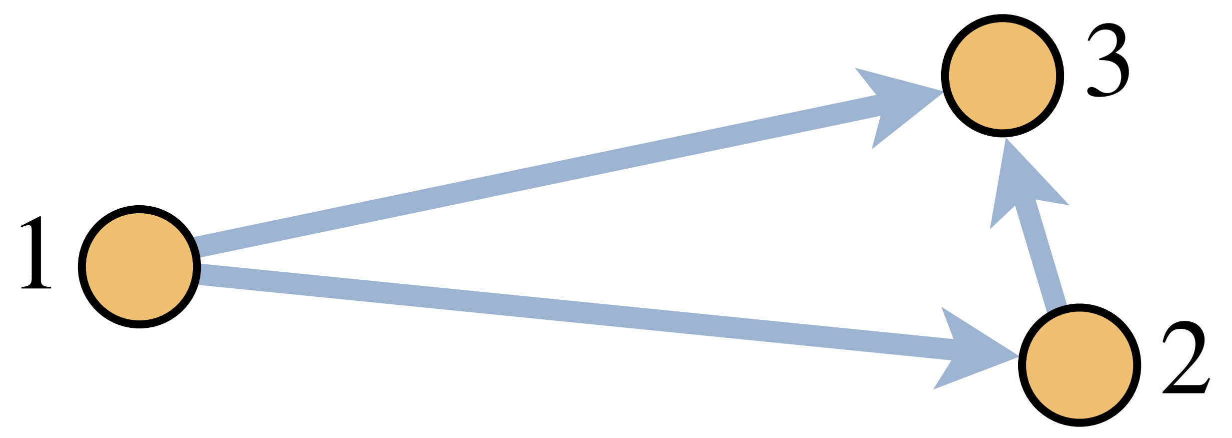


Nodes = Star Wars characters
Links = co-appearances in Star Wars movies

http://evelinag.com/blog/2015/12-15-star-wars-social-network/

Networks: representations

A network is a collection of **nodes** and **links**



Definition

Precise but cumbersome

Nodes $= \{1,2,3\}$

Links $= \{(1,2), (1,3), (2,3)\}$

Edge list

Computationally efficient

1,2
1,3
2,3

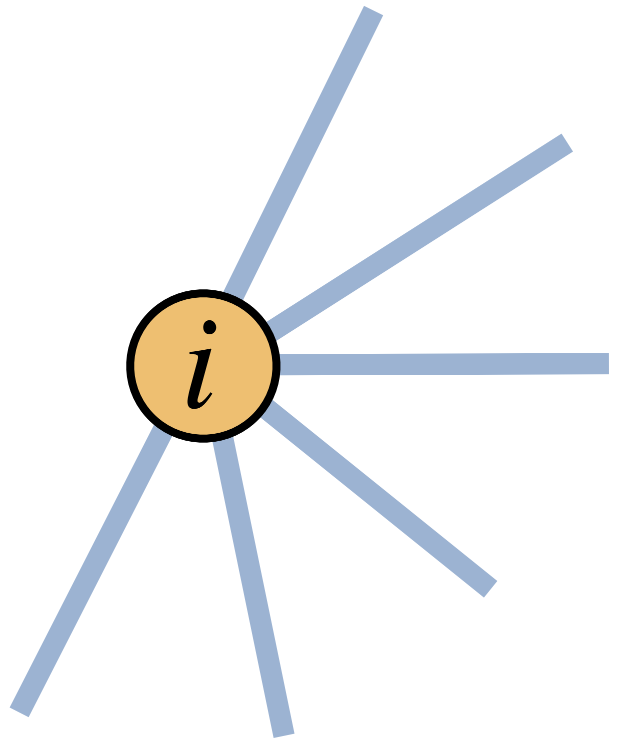
Adjacency matrix = A

Useful for theory, inefficient when storing many zeros

$$\begin{array}{c c} & \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \end{array}$$

“2 points to 3”

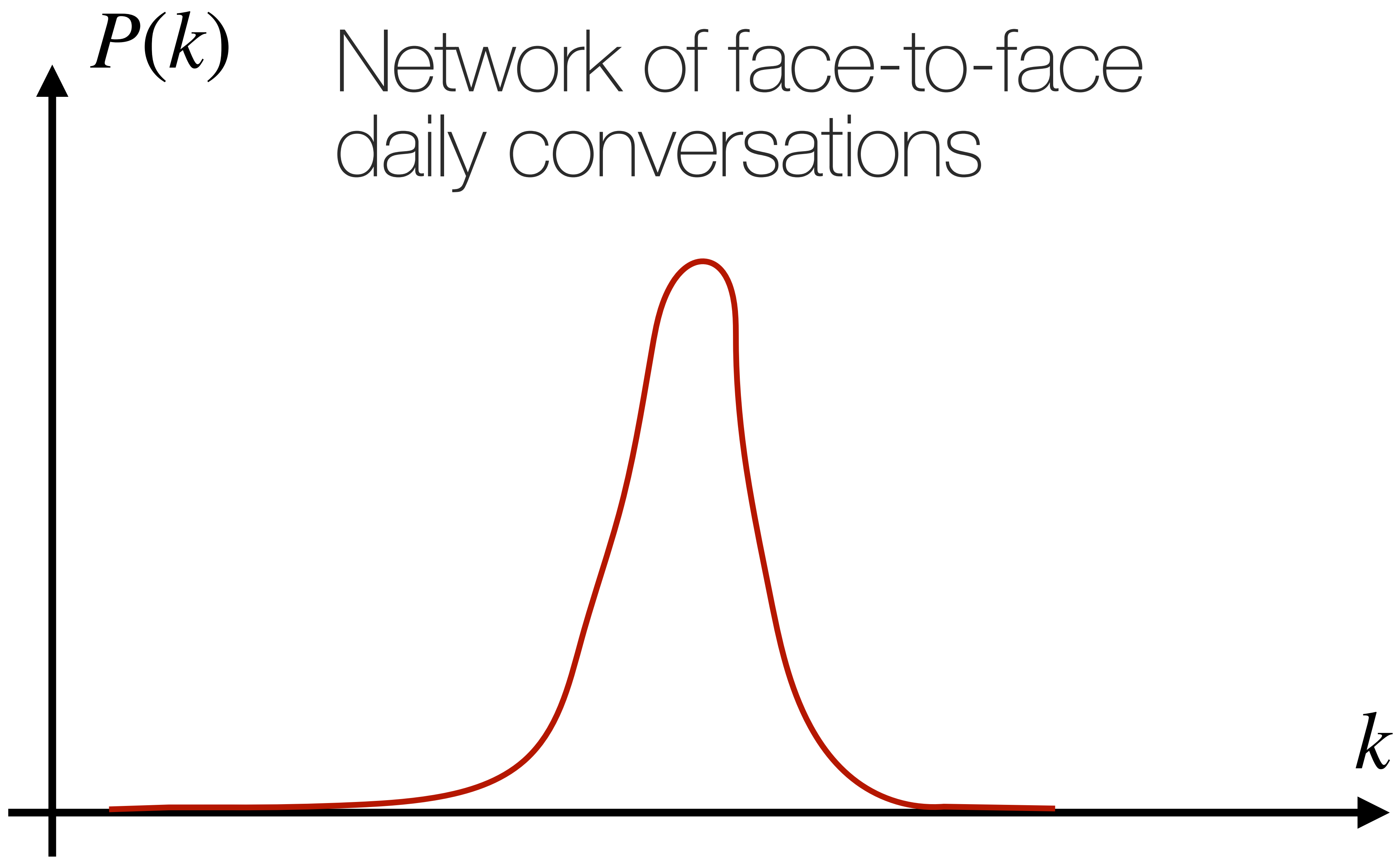
Basic properties: degree and degree distribution

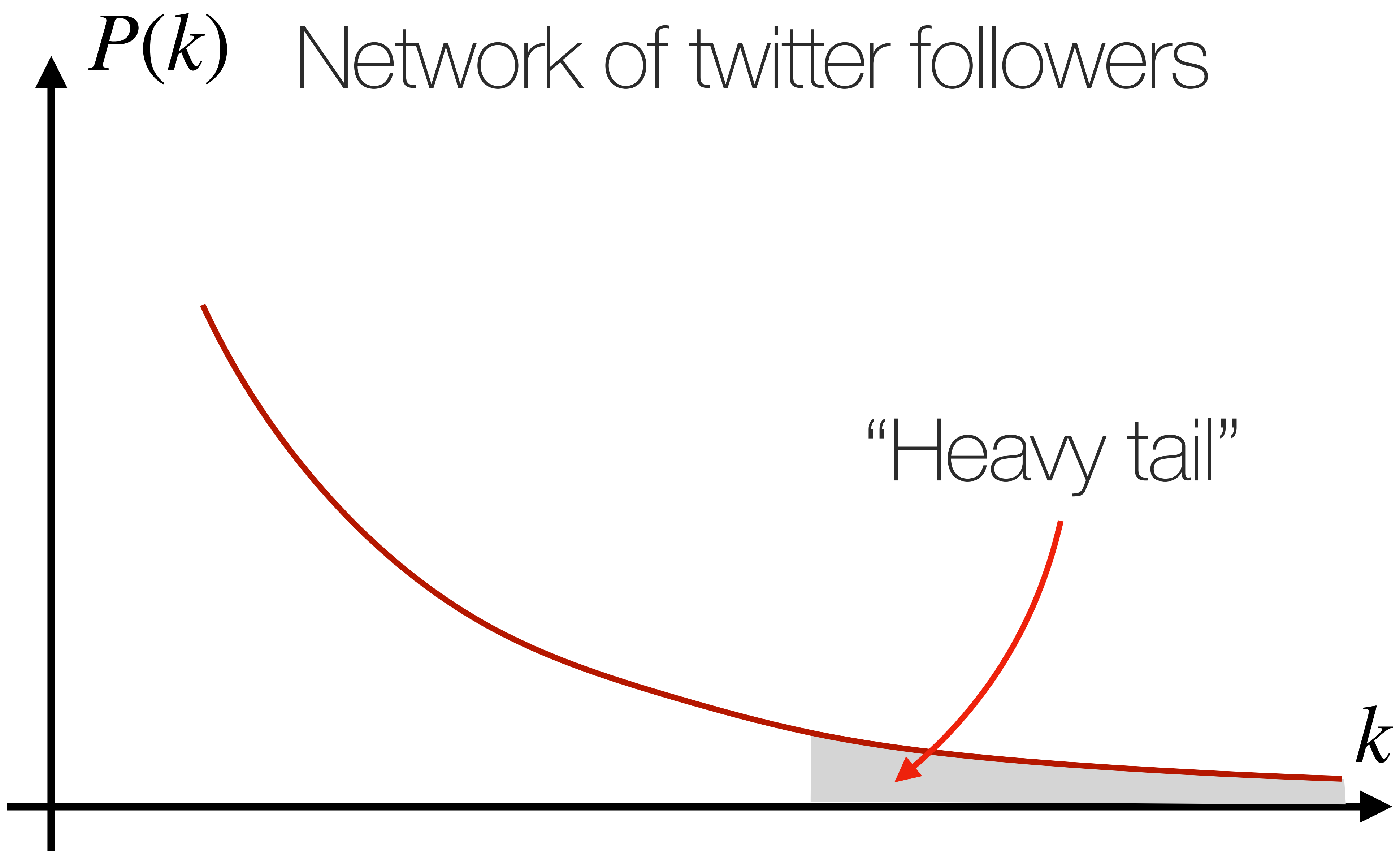


The **degree** of a node is the number of links it has to other nodes

Degree of node $i = k_i$

The **degree distribution** is the fraction of nodes with degree k. Consider these two examples:

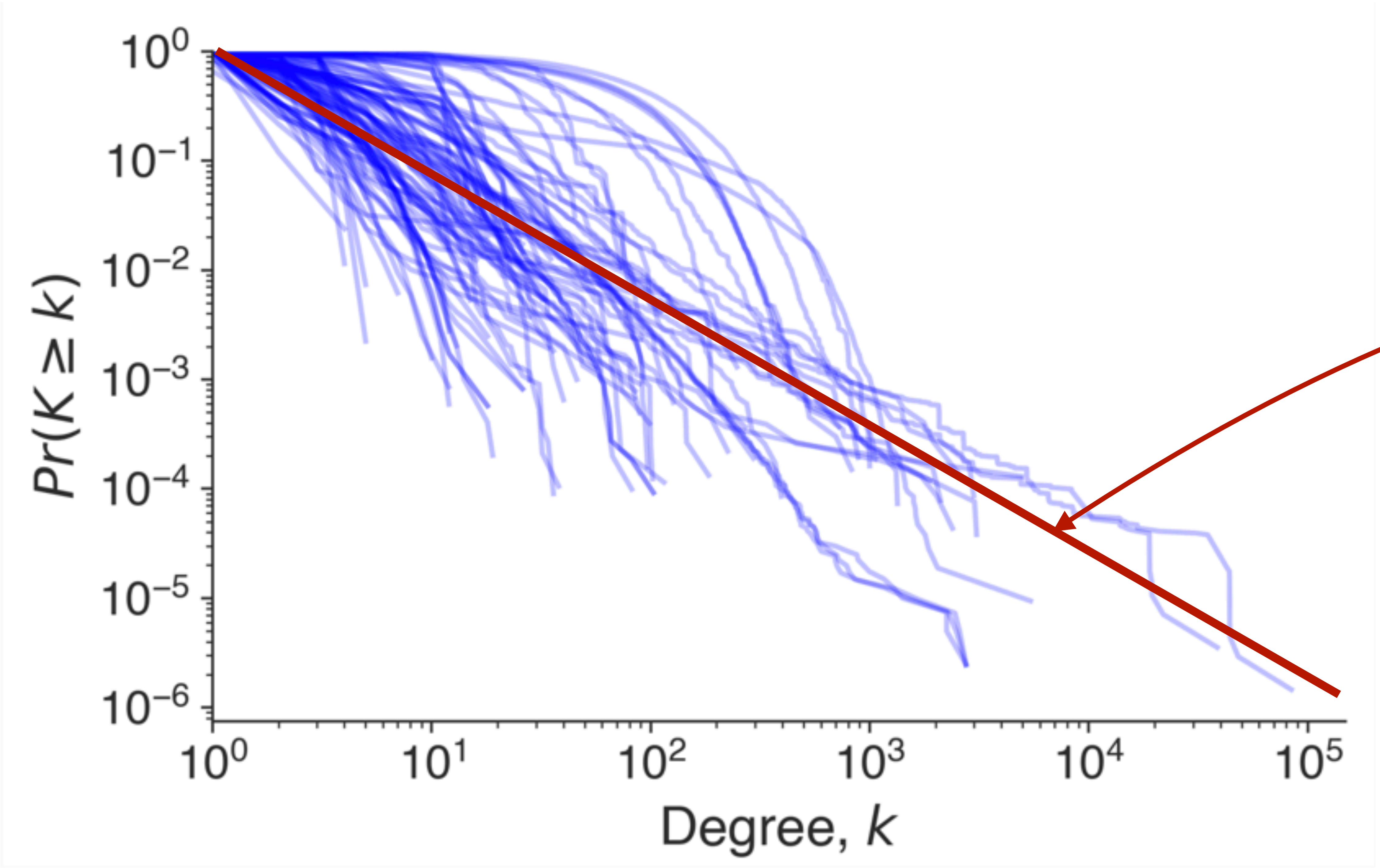




Basic properties: degree distribution

The **degree distribution** of real-world networks is often **heavy-tailed.**

In practice this means there is a significant number of nodes with very high degrees (think popular twitter/instagram users or web pages)

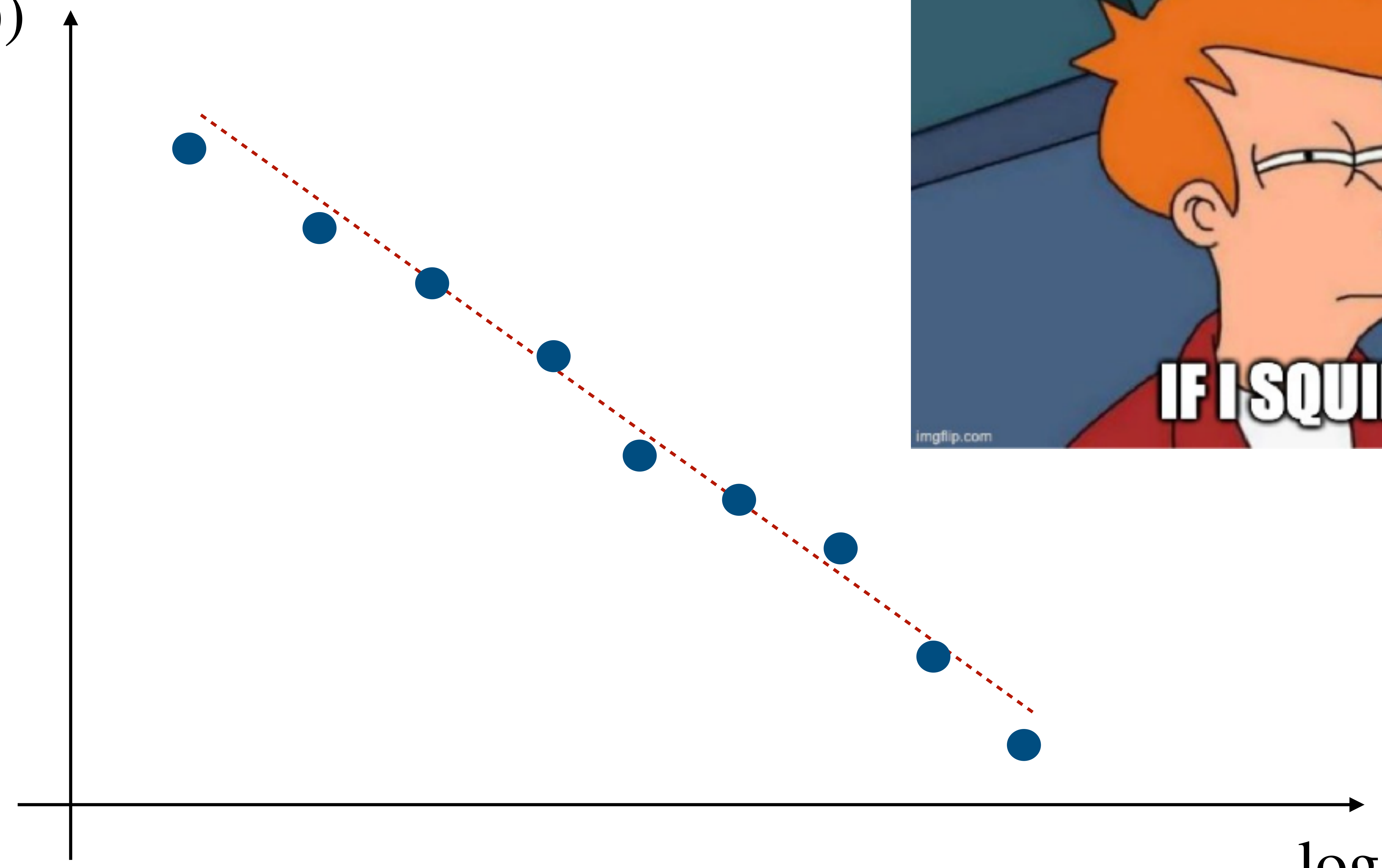


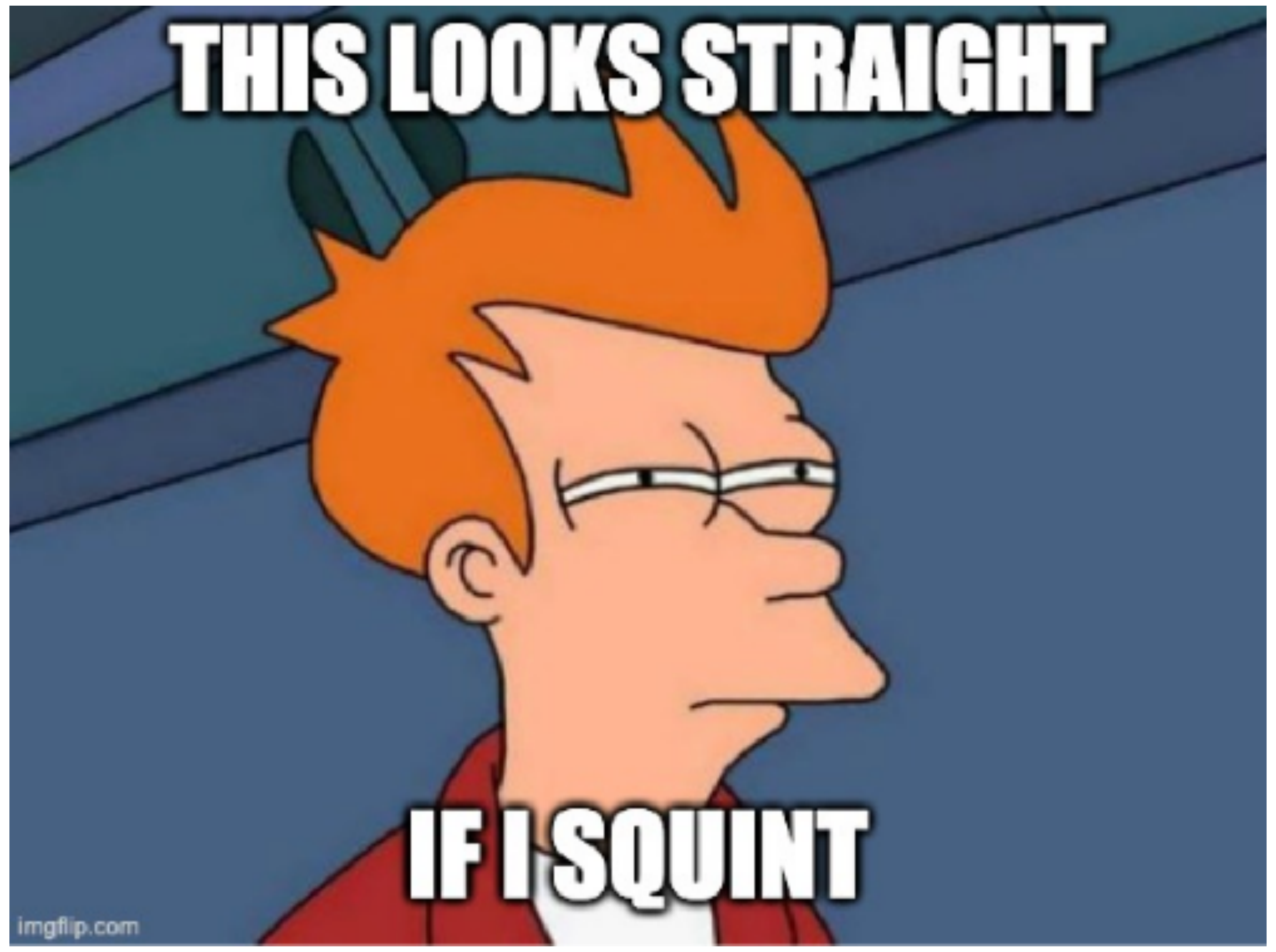
For some networks
$P(k) \sim k^{-\gamma}$
"Scale-free networks"

Credit: Aaron Clauset

Scale-free networks

$$P(k) = Ck^{-\gamma} \quad \longrightarrow \quad \log(P(k)) = \log(C) - \gamma \log(k)$$

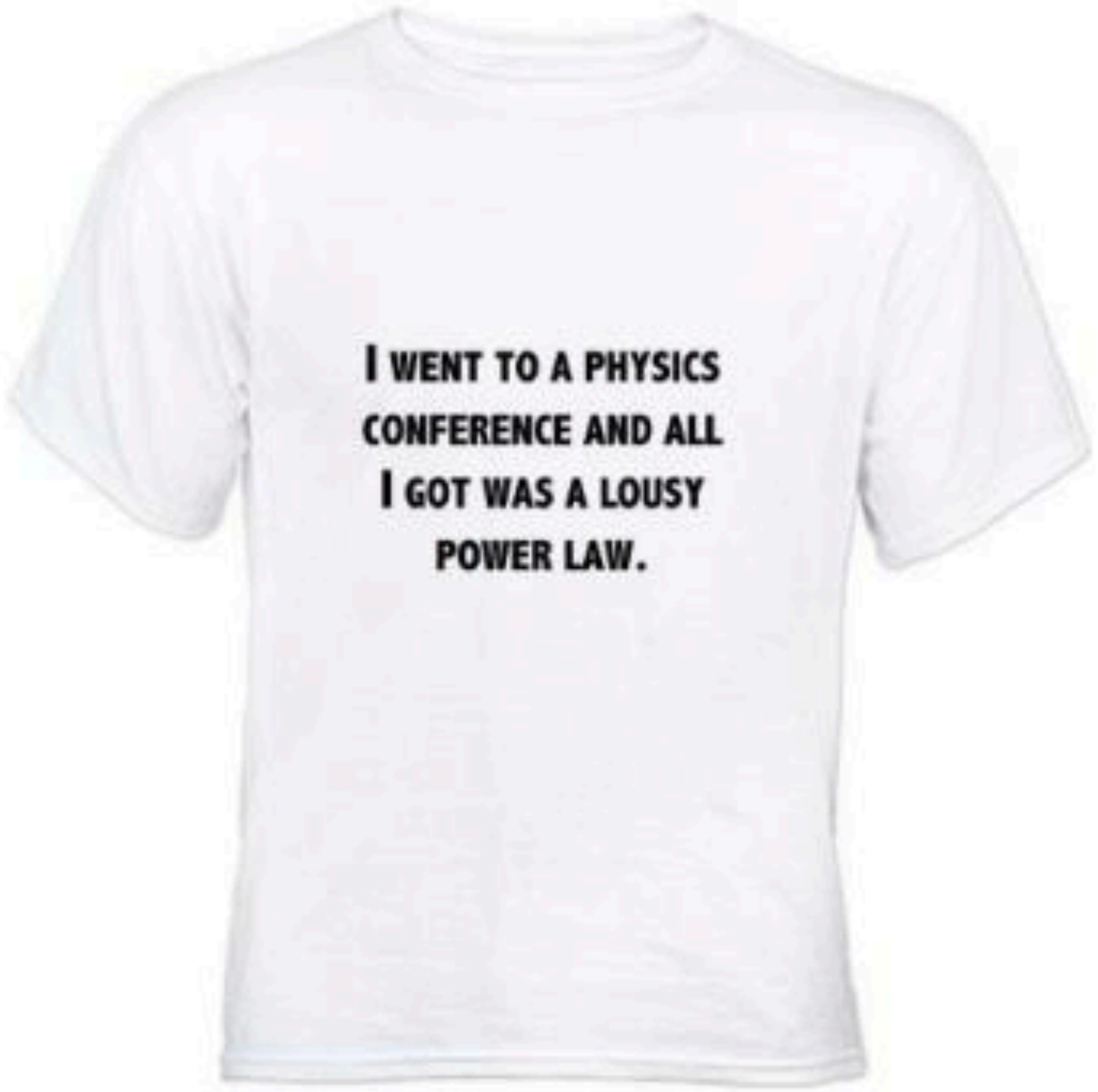




Mason Porter’s Power Law Shop

(now sadly inactive as far as I know)

http://www.cafepress.com/thepowerlawshop



Scale-free networks - why all the fuss?

Pushing Networks to the Limit

PERSPECTIVE

Scale-Free Networks: A Decade and Beyond

Albert-László Barabási

Article | Open Access | Published: 04 March 2019

Scale-free networks are rare

Anna D. Broido & Aaron Clauset

Nature Communications **10**, Article number: 1017 (2019) | Cite this article

59k Accesses | **314** Citations | **604** Altmetric | Metrics

Power laws suggest the possibility of common organizing principles

Physicists attracted to the similarities with statistical physics

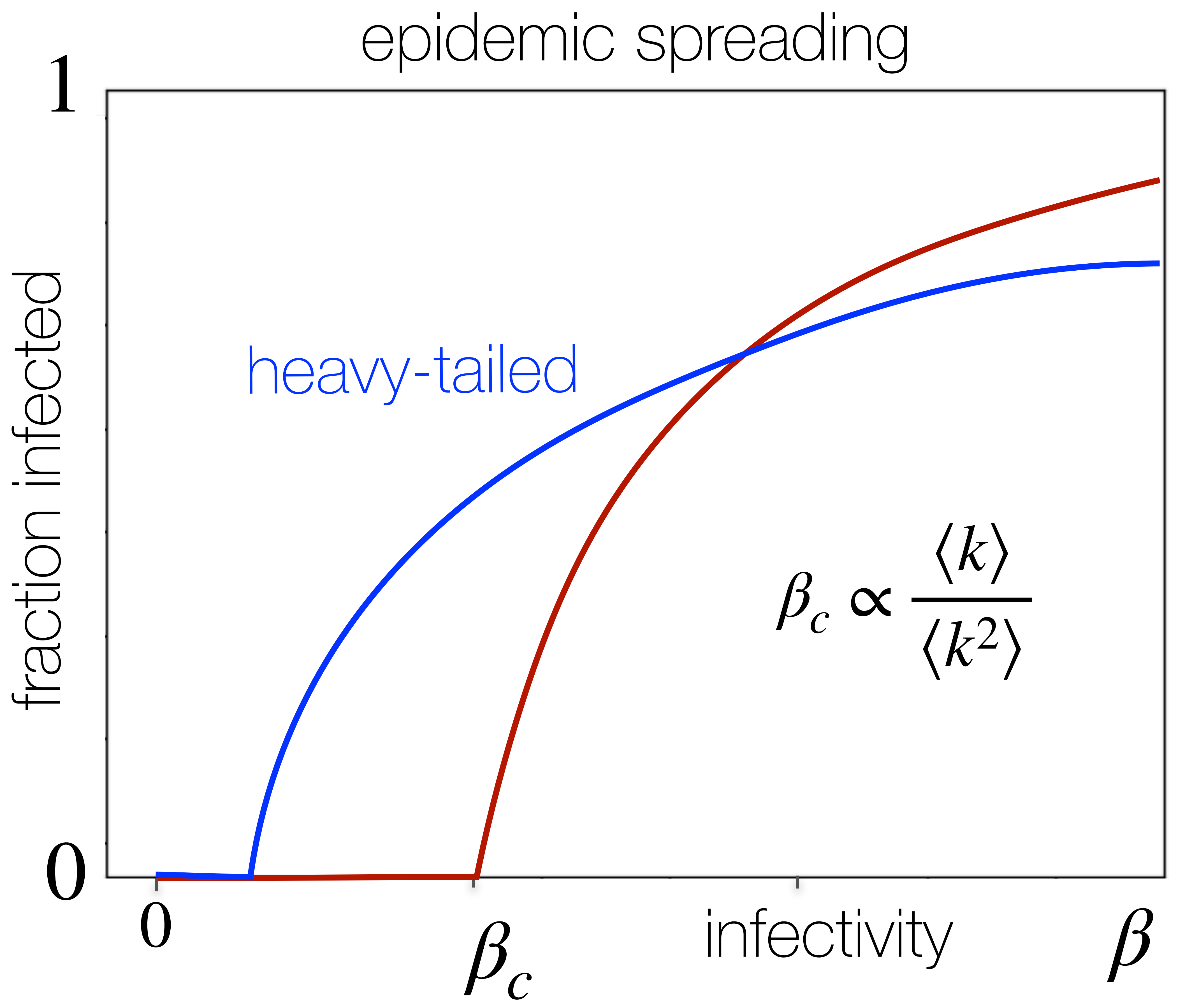
The fit to power-law scaling is not statistically justified in many situations

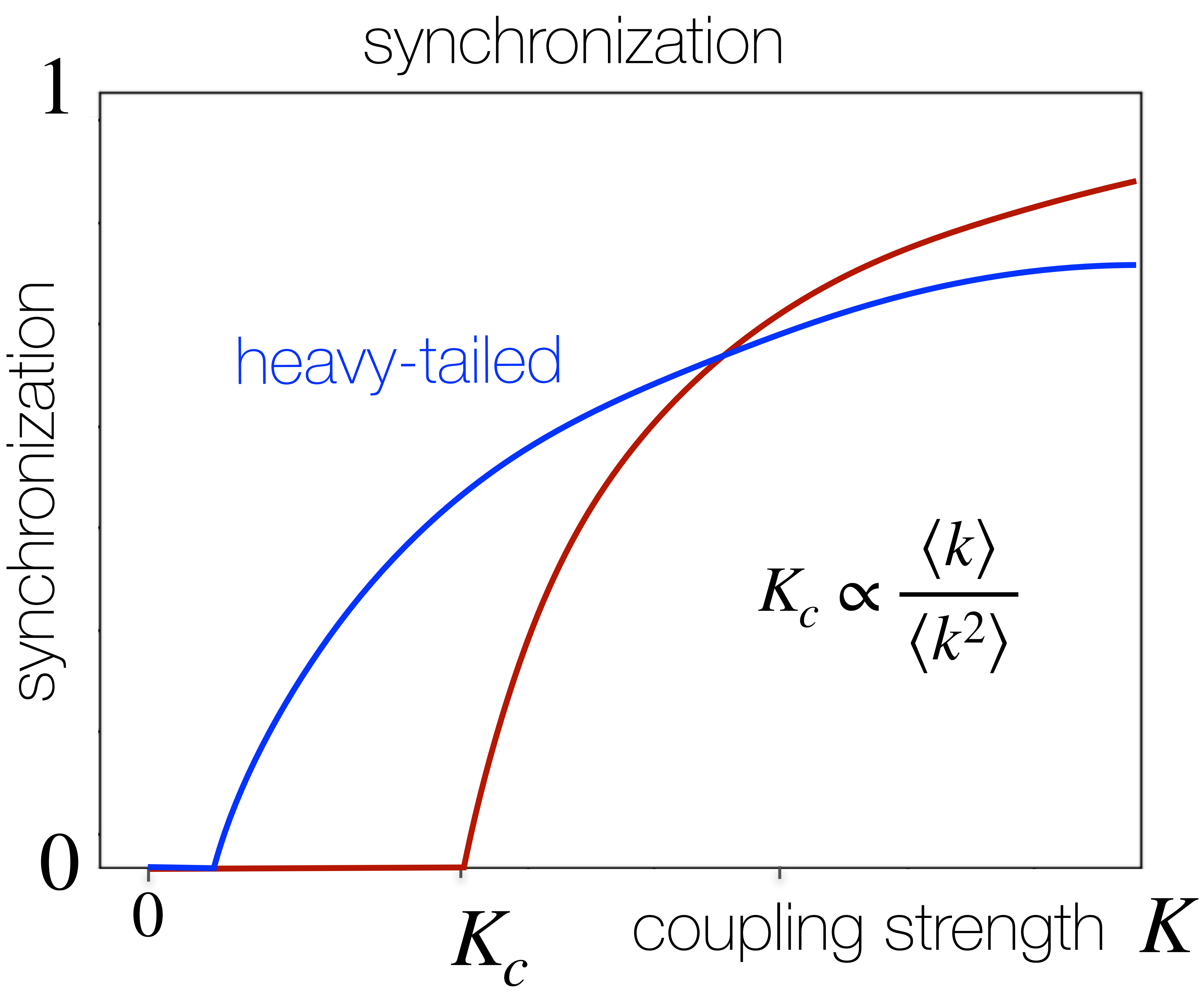
The consensus (if there's one) is that many networks have a heavy-tailed degree distribution, and that's what matters

Heavy-tailed degree distribution - moments

The qualitative behavior of processes on networks often depend strongly on the value of $\langle k^2 \rangle / \langle k \rangle$

where $\langle k^n \rangle$ is the nth moment of the degree distribution, $\langle k^n \rangle = \frac{1}{N} \sum_{i=1}^{N} k_i^n$





Assortative mixing

Sometimes nodes are **more likely to connect if they have similar characteristics**

Perhaps the single most relevant concept in network science for social justice analysis!

Facebook users more likely to connect if they went to the same high school 😐

Employers might want to hire people who look like them 🤨

Assortative mixing

Sometimes nodes are **more likely to connect if they have similar characteristics**

Example: friendship network in middle and high schools: mixing by **race** and **grade**

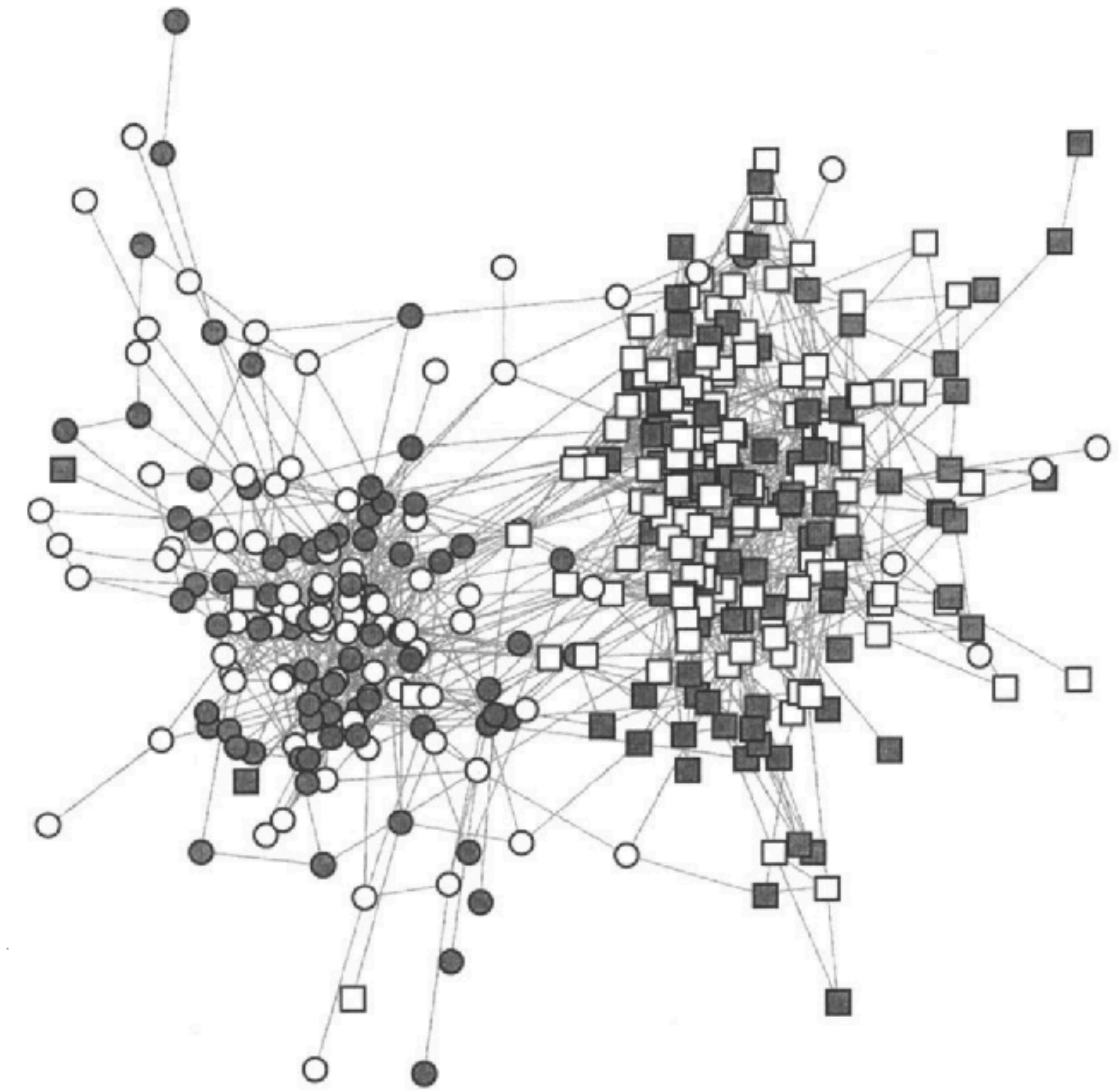


FIG. 2.—Friendship relations in "Mountain Middle School" by race and grade. Shaded figures represent nonwhite students. Circles = seventh graders and squares = eighth graders.

FIG. 1.—Friendship relations in "Countryside High School" by race and grade. Shaded figures represent nonwhite students. Circles = ninth graders, squares = tenth graders, hexagons = eleventh graders, and triangles = twelfth graders.

James Moody, "Race, School Integration, and Friendship Segregation in America", American Journal of Sociology, v 107, Number 3, pp. 679-716 (2001).

Assortativity coefficient

How assortative is the network? How can we tell, **without visualizing and eyeballing** the patterns, if there is assortative mixing?

Suppose that $\boldsymbol{e_{ij}}$ is the fraction of edges connecting a node of type $\boldsymbol{i}$ to a node of type $\boldsymbol{j}$. An assortativity coefficient can be defined as

From $\boldsymbol{i}$ to $\boldsymbol{i}$ From $\boldsymbol{i}$ To $\boldsymbol{i}$

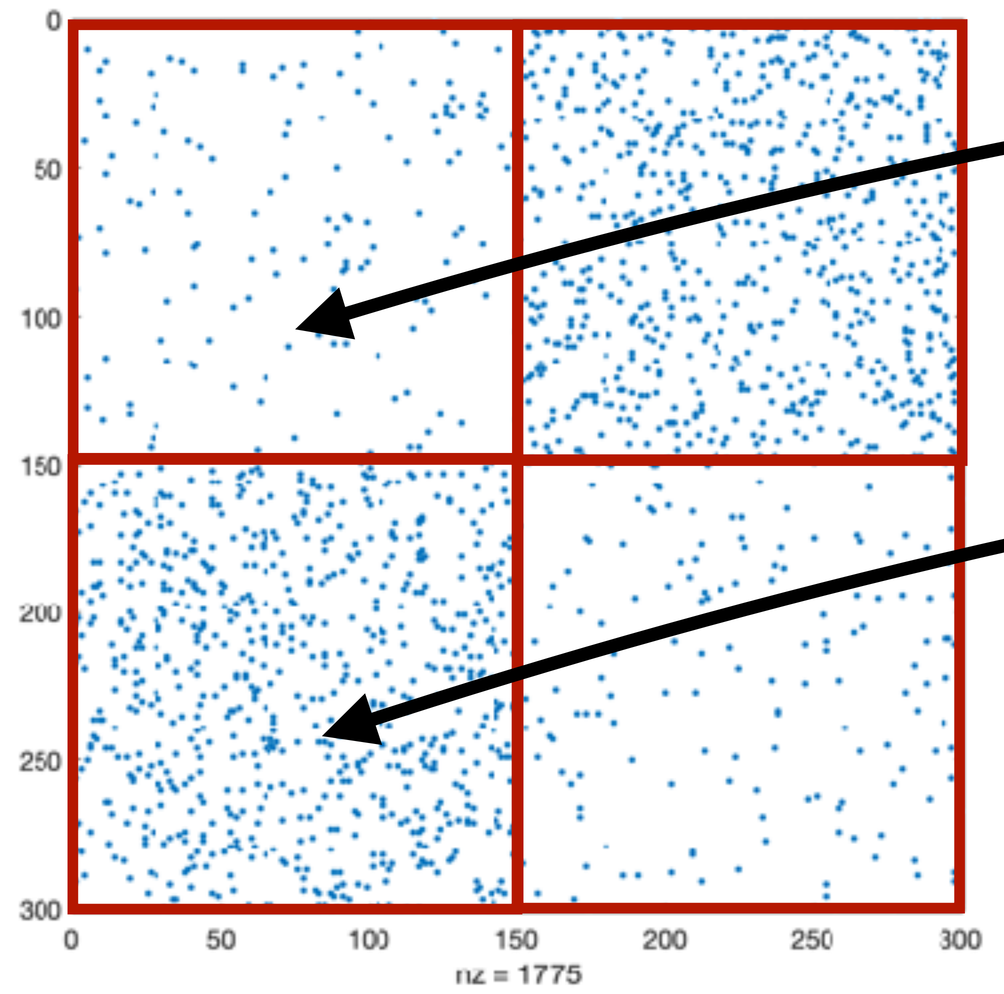
$$r = \frac{\sum_i e_{ii} - \sum_i a_i b_i}{1 - \sum_i a_i b_i}$$

where $a_i = \sum_j e_{ij}$ $b_j = \sum_i e_{ij}$

$r = 0$: randomly connected; $r = 1$: perfectly segregated

Newman, Mark EJ. "Mixing patterns in networks." Physical review E 67.2 (2003): 026126.

Assortativity coefficient



e_{11} = number of 1's here divided by total number of 1's

e_{21} = number of 1's here divided by total number of 1's

$r = -0.68$ $r = 0.05$ $r = 0.59$ $r = 0.99$

Assortativity coefficient

In this example: **only two categories**

We know exactly which nodes belong to which category.

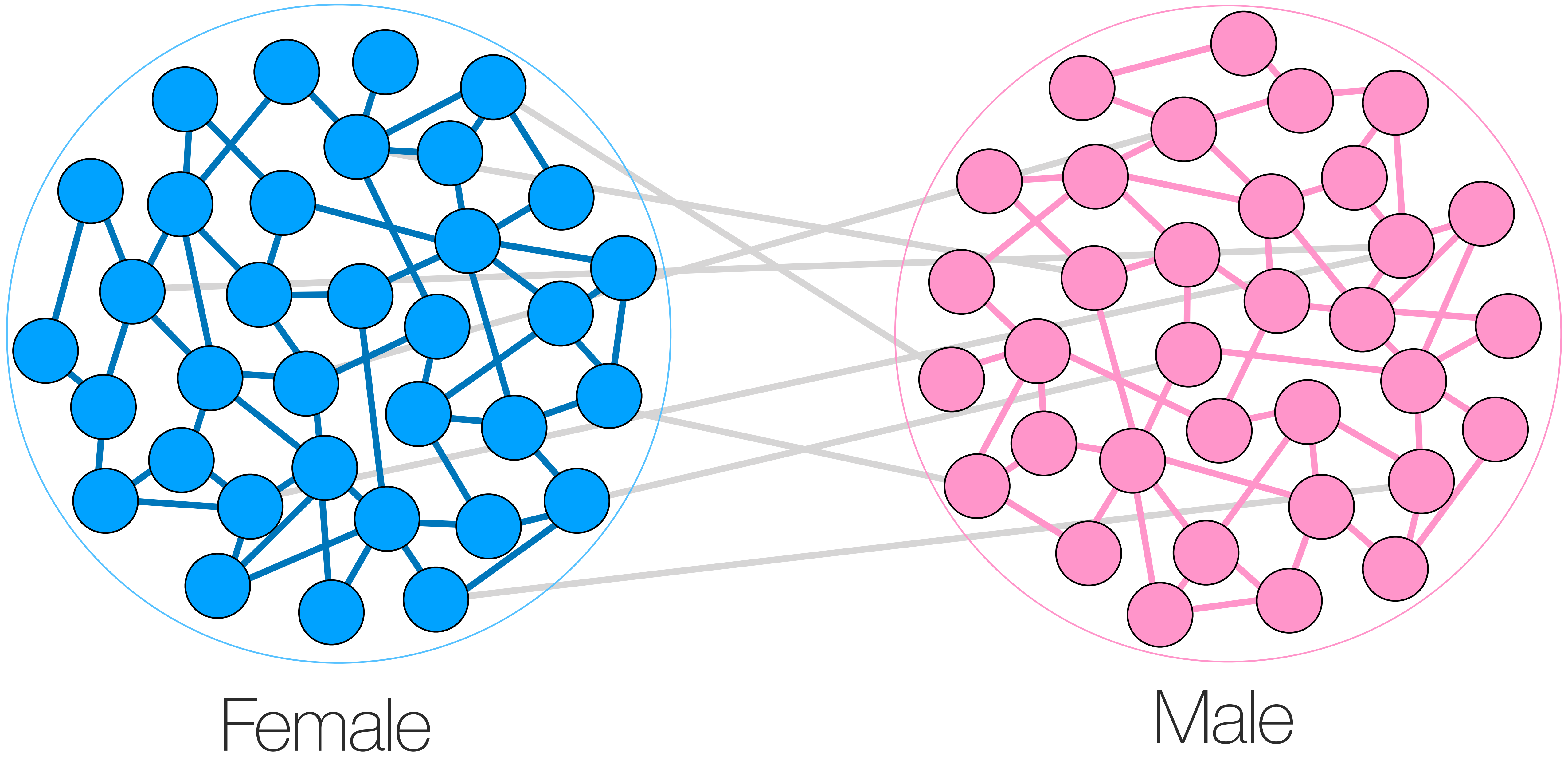
In general: we might want to include more categories, and data on categories might not be clear.

Which category describes you? Please select all that apply: *

- [ ] American Indian or Alaska Native
- [ ] Asian
- [ ] Black or African American
- [ ] Hispanic, Latino or Spanish Origin
- [ ] Middle Eastern or North African
- [ ] Native Hawaiian or Other Pacific Islander
- [ ] White
- [ ] Multiethnic
- [ ] Prefer not to disclose
- [ ] Other: ____

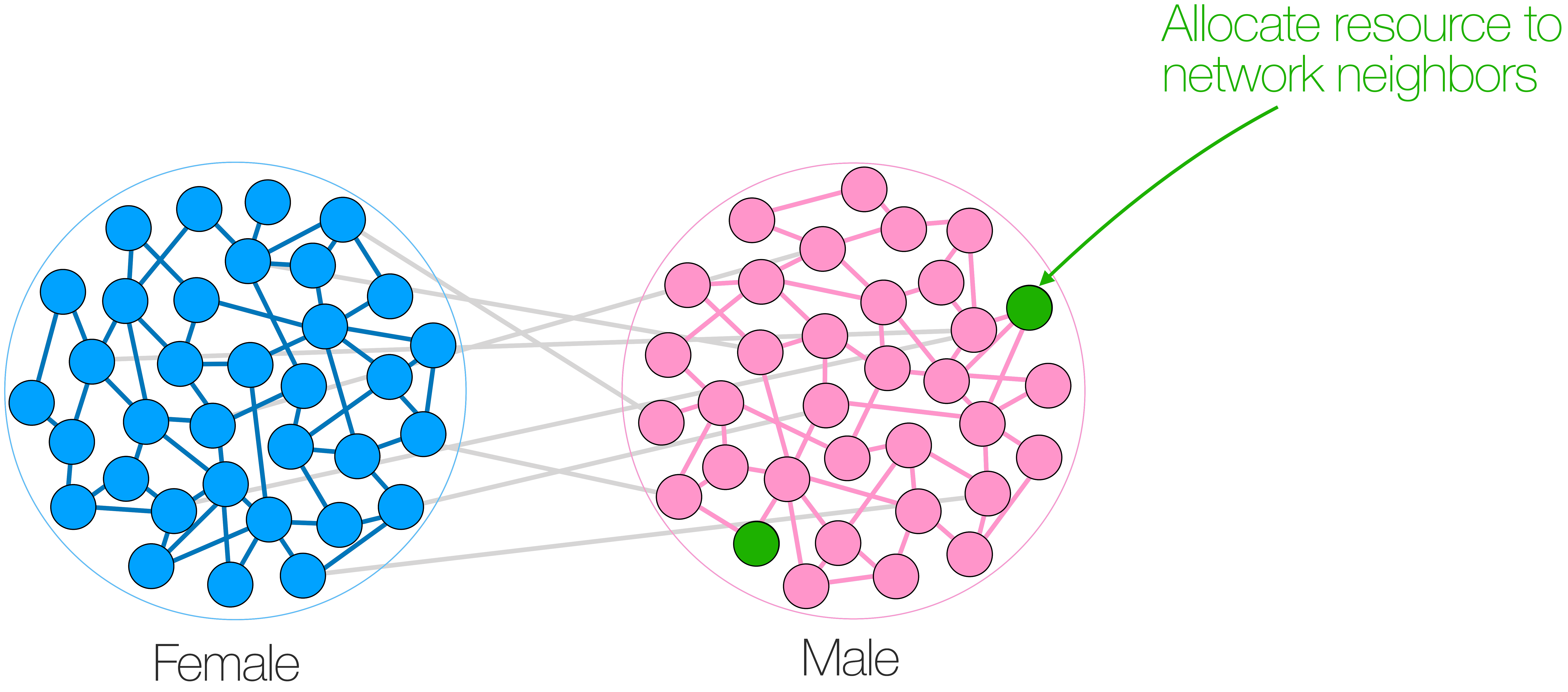
Assortative mixing

Example: organization with overall gender balance (50%-50%), but with assortative mixing by gender and with resources controlled by males



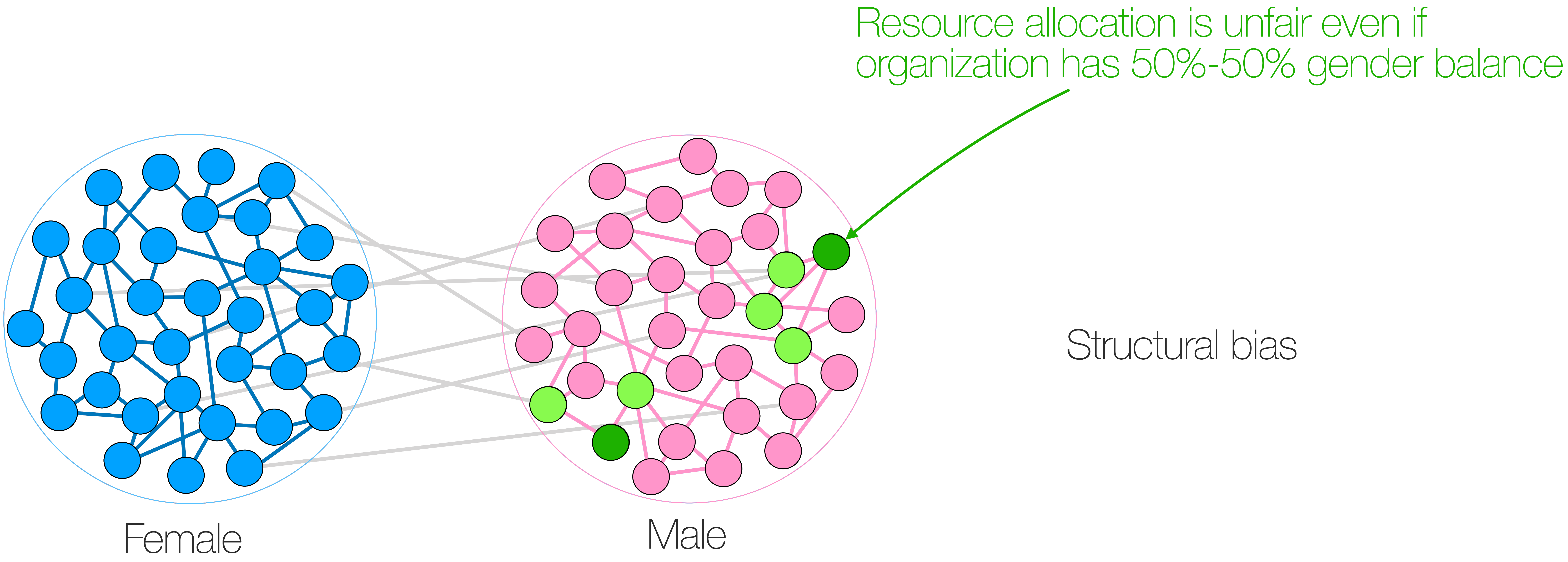
Assortative mixing

Example: organization with overall gender balance (50%-50%), but with assortative mixing by gender and with resources controlled by males



Assortative mixing

Example: organization with overall gender balance (50%-50%), but with assortative mixing by gender and with resource distribution controlled by males



Structural bias

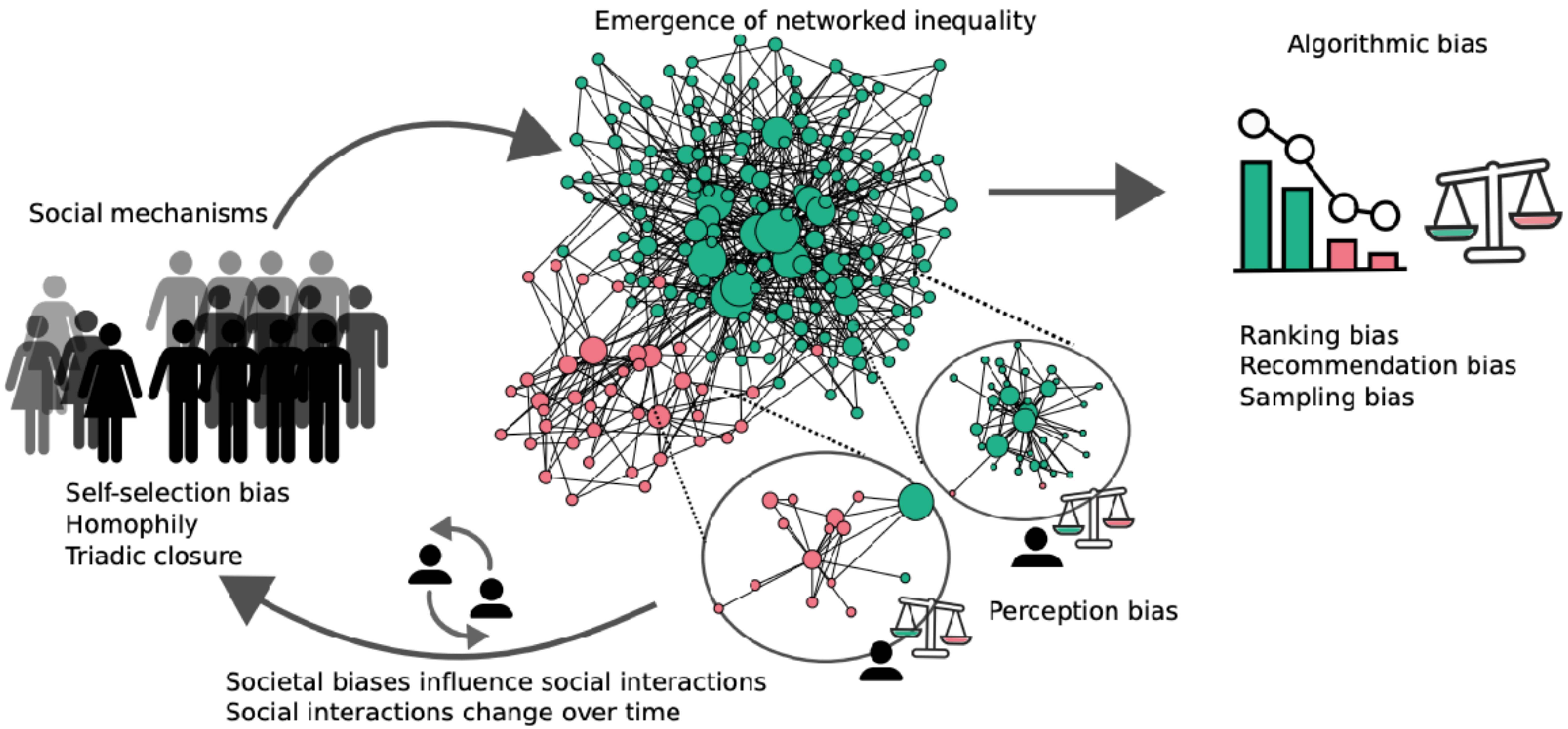
Relevant read

Minorities in networks and algorithms

Fariba Karimi[1*], Marcos Oliveira[2], Markus Strohmaier[4,3,1]

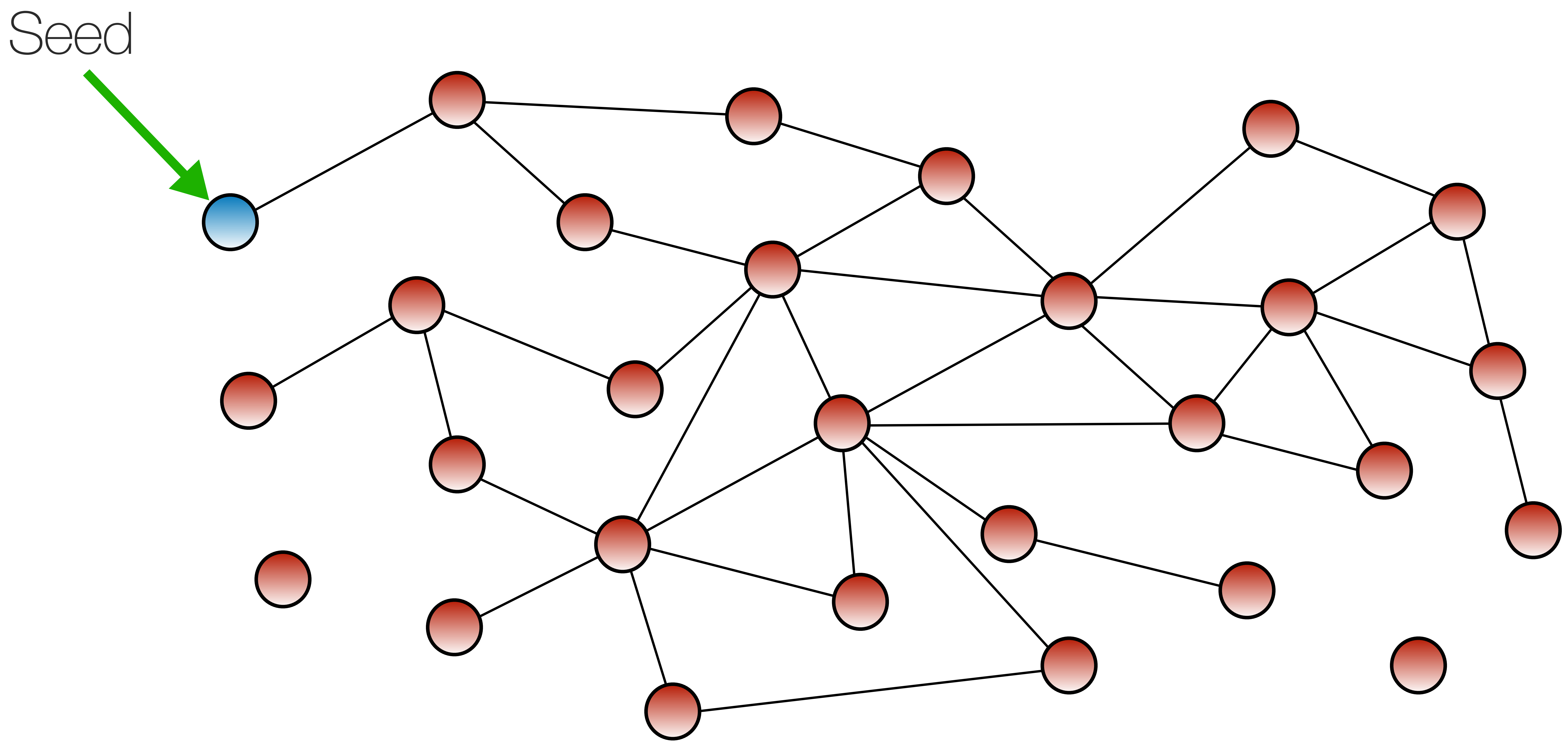
https://arxiv.org/abs/2206.07113

(submitted to arXiv June 14)



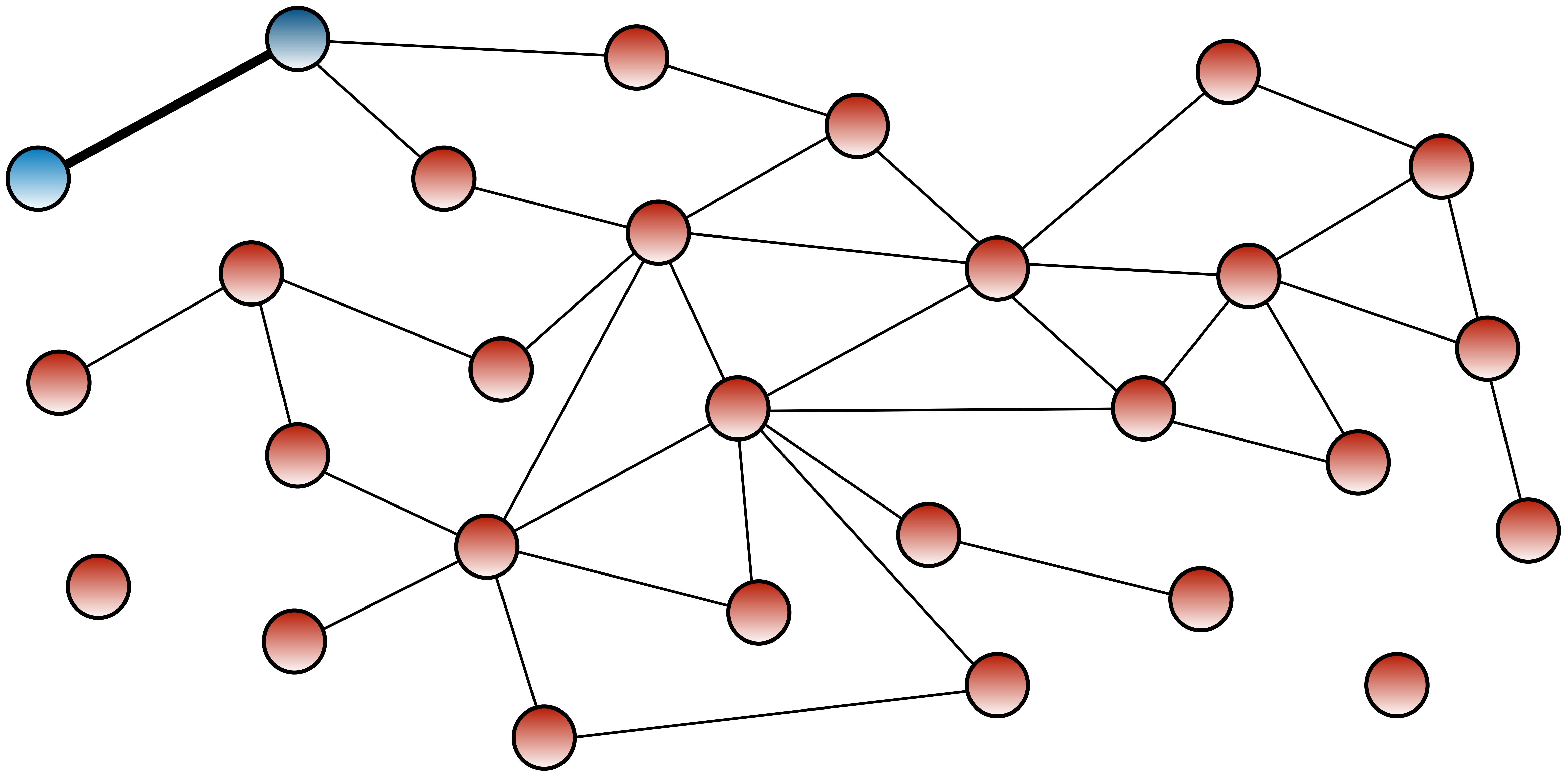
Centrality

Motivation: spreading processes on networks



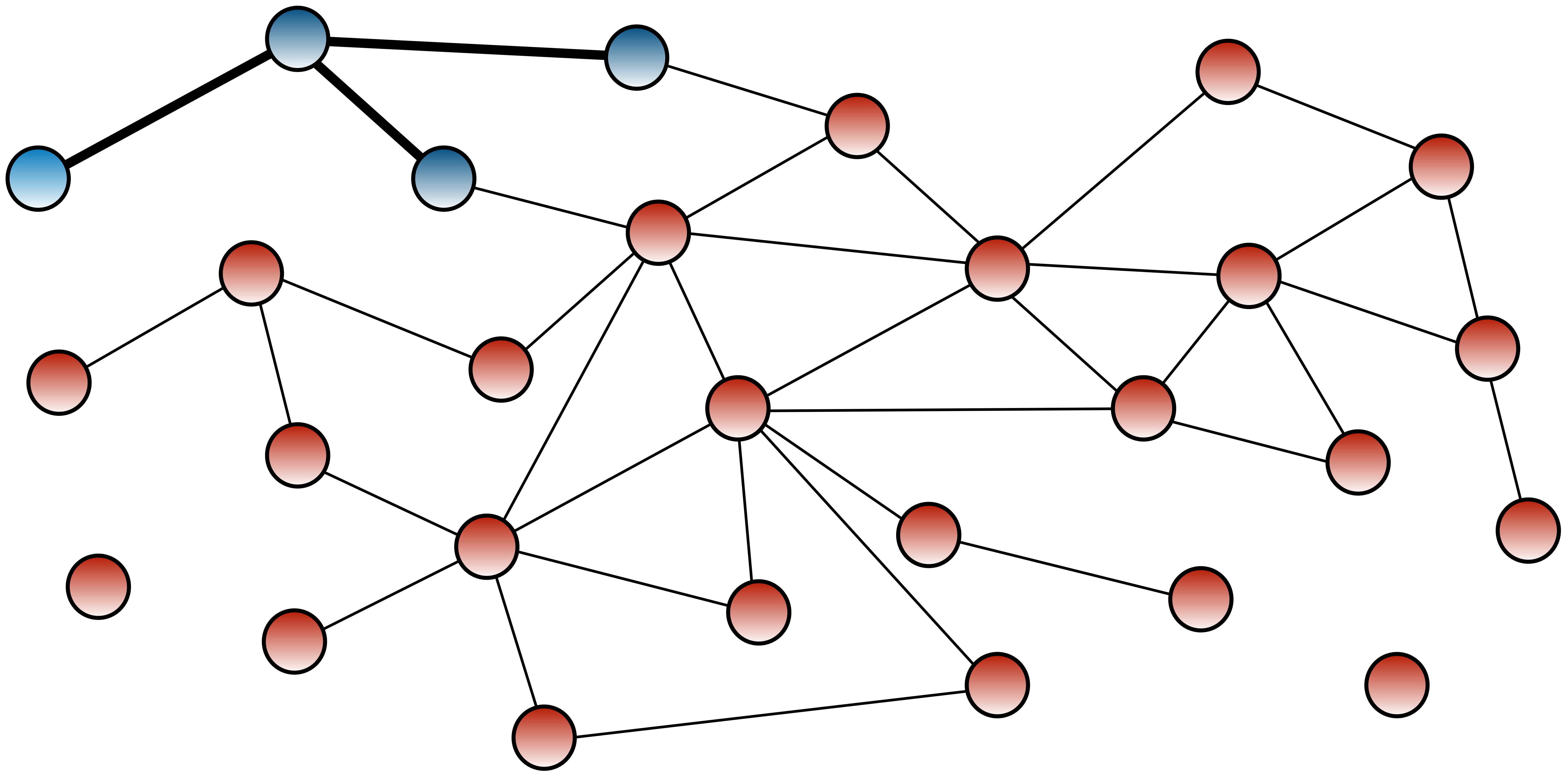
Centrality

Motivation: spreading processes on networks



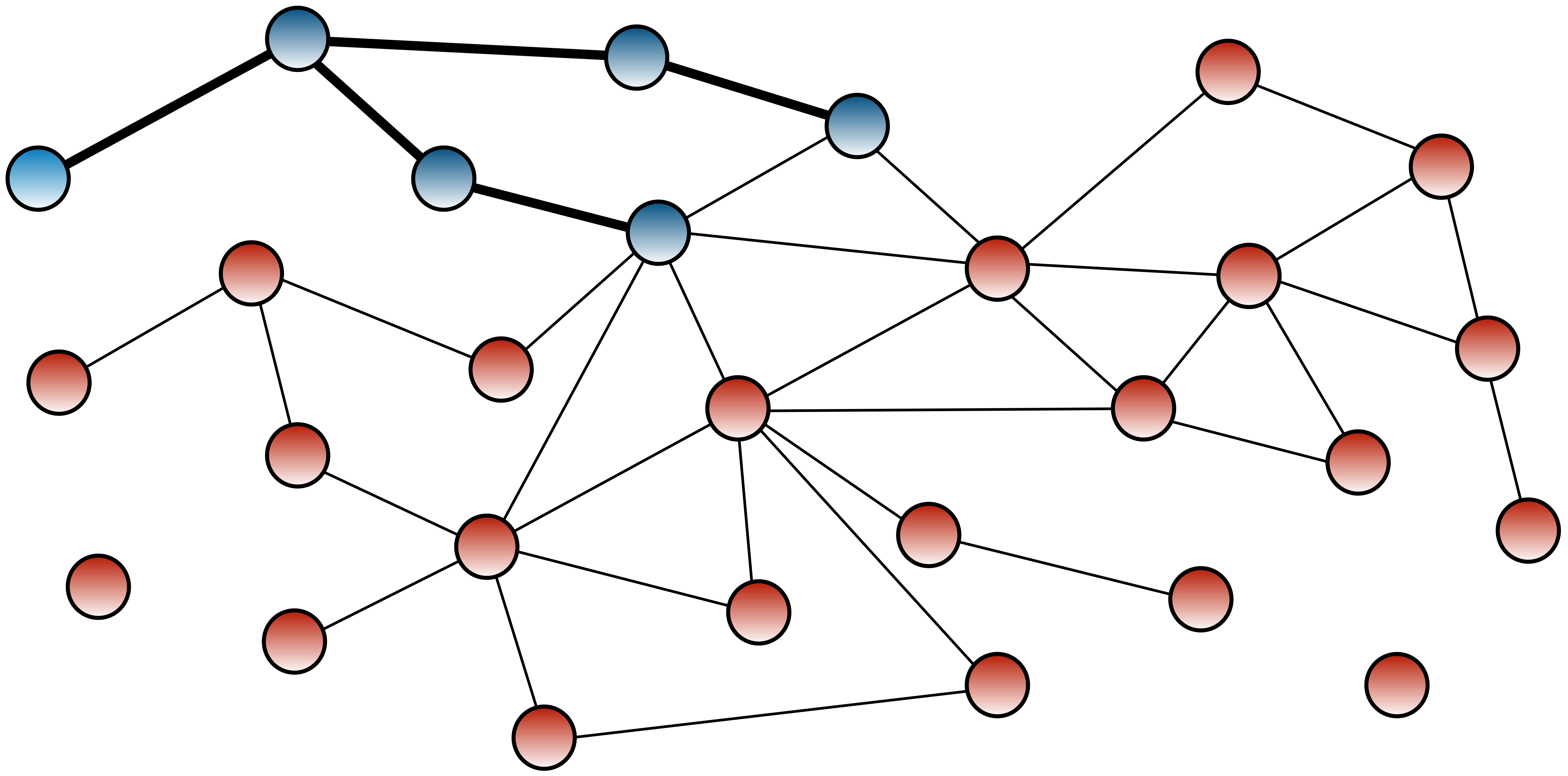
Centrality

Motivation: spreading processes on networks



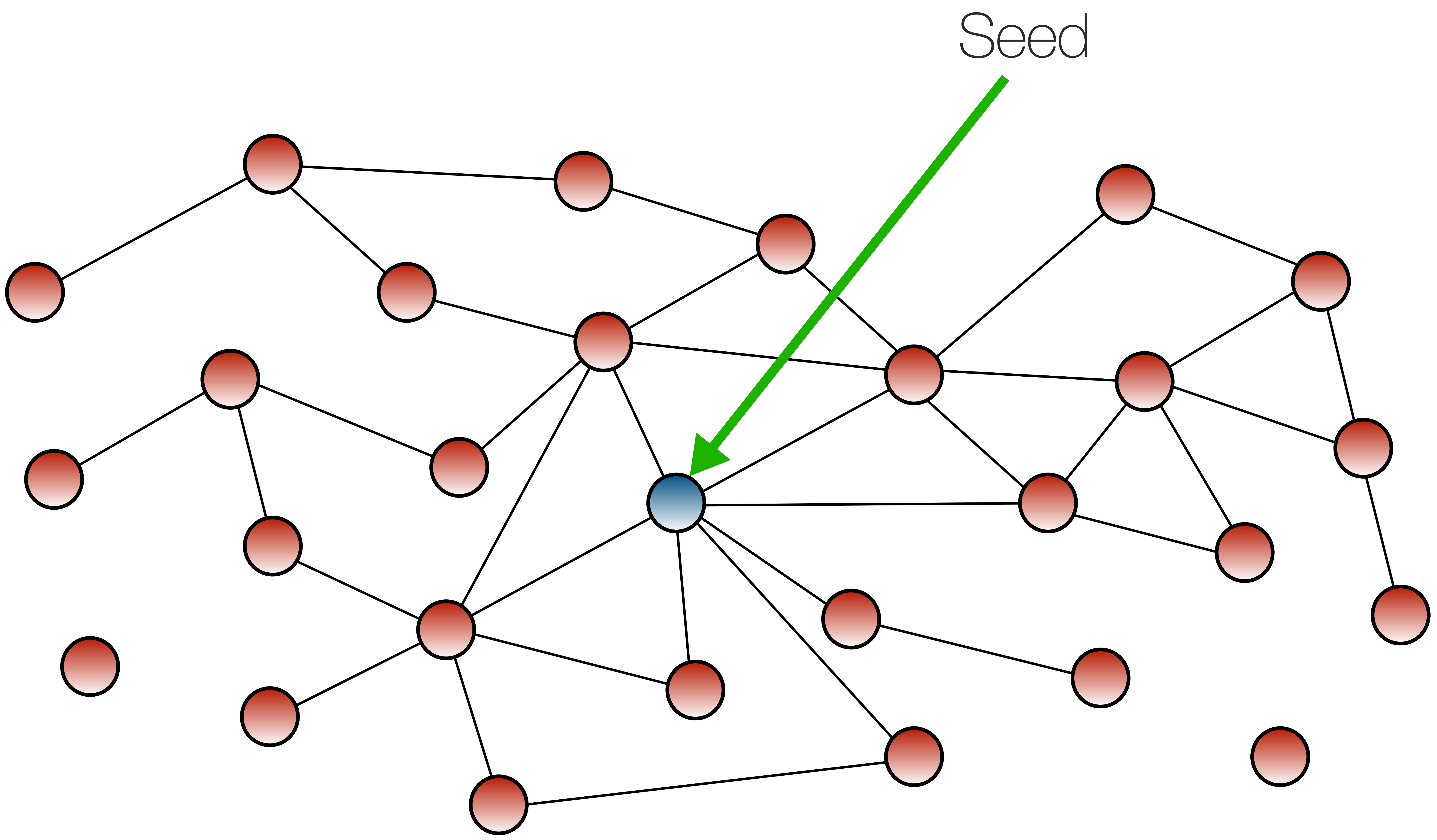
Centrality

Motivation: spreading processes on networks



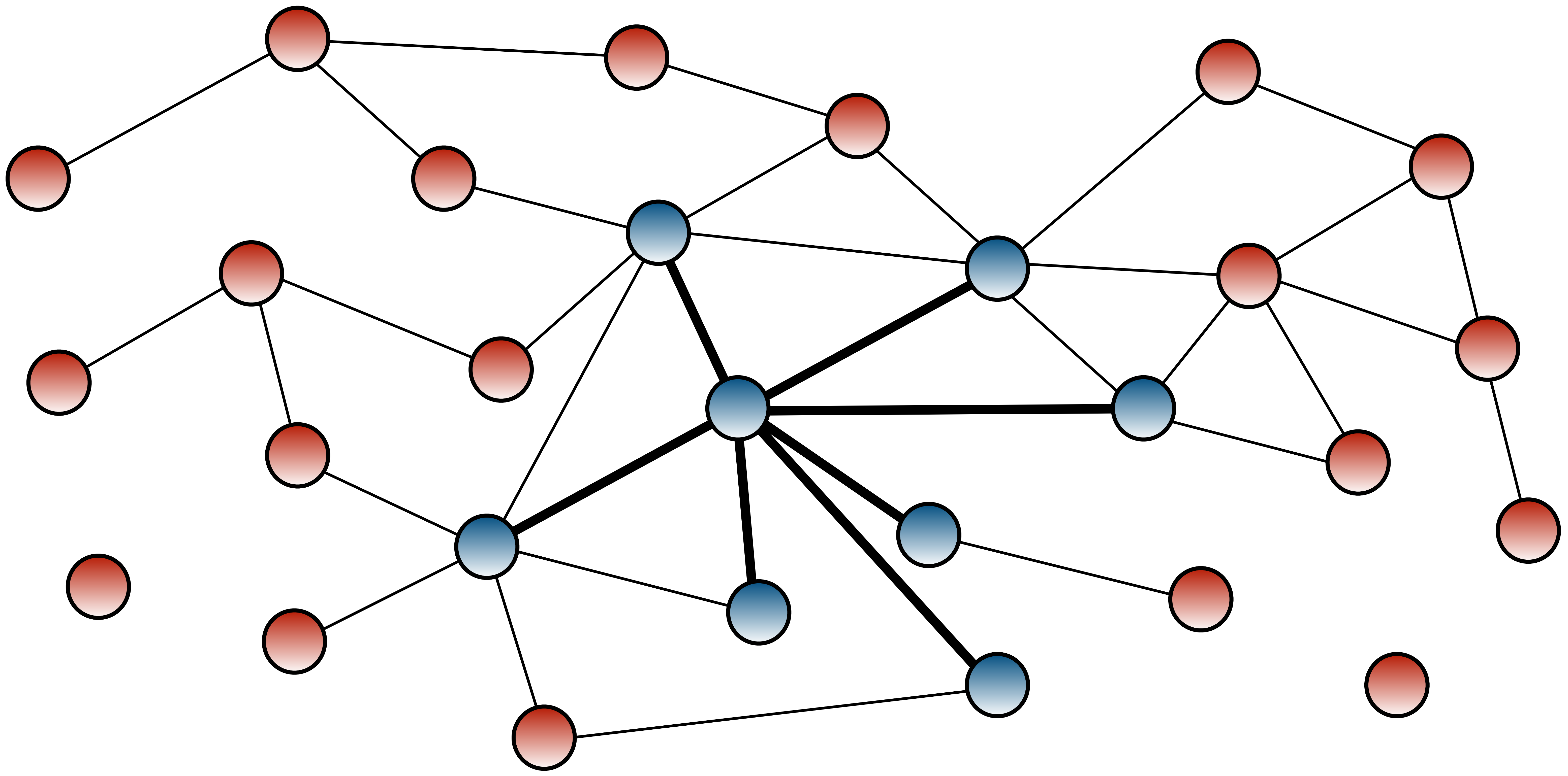
Centrality

Motivation: spreading processes on networks



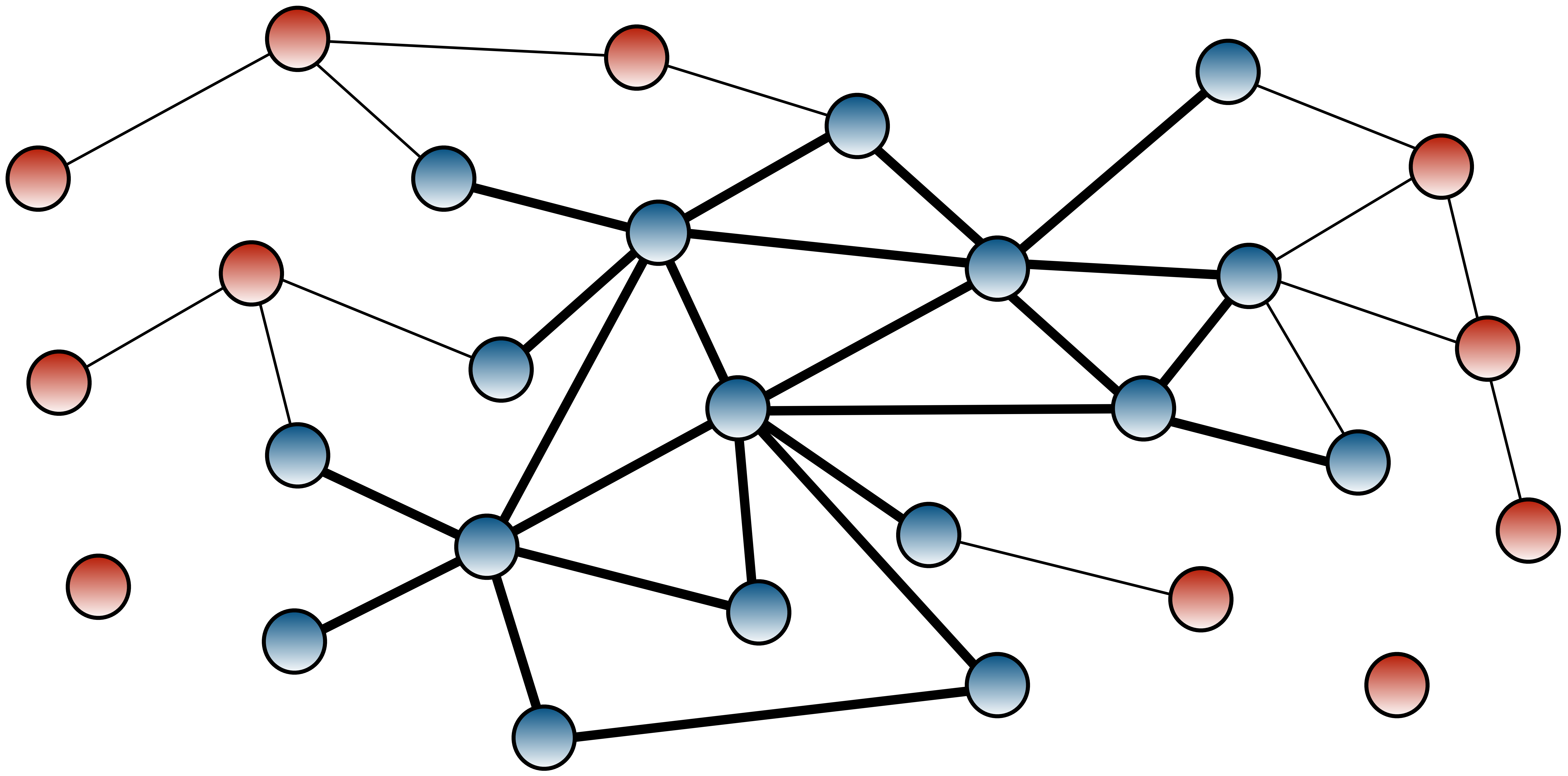
Centrality

Motivation: spreading processes on networks



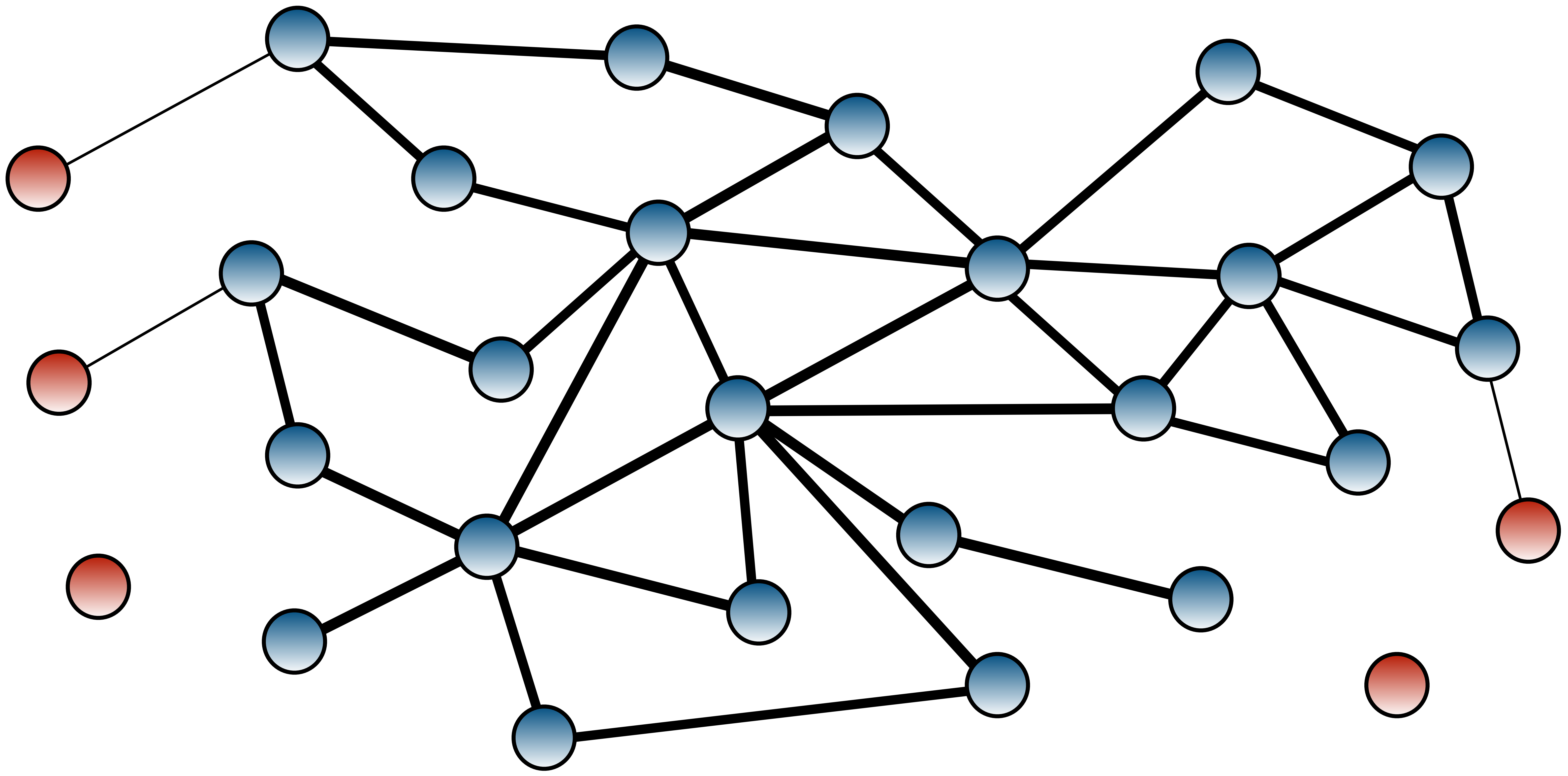
Centrality

Motivation: spreading processes on networks



Centrality

Motivation: spreading processes on networks



Many ways to define centrality

Degree centrality: the "degree centrality" of a node is its degree

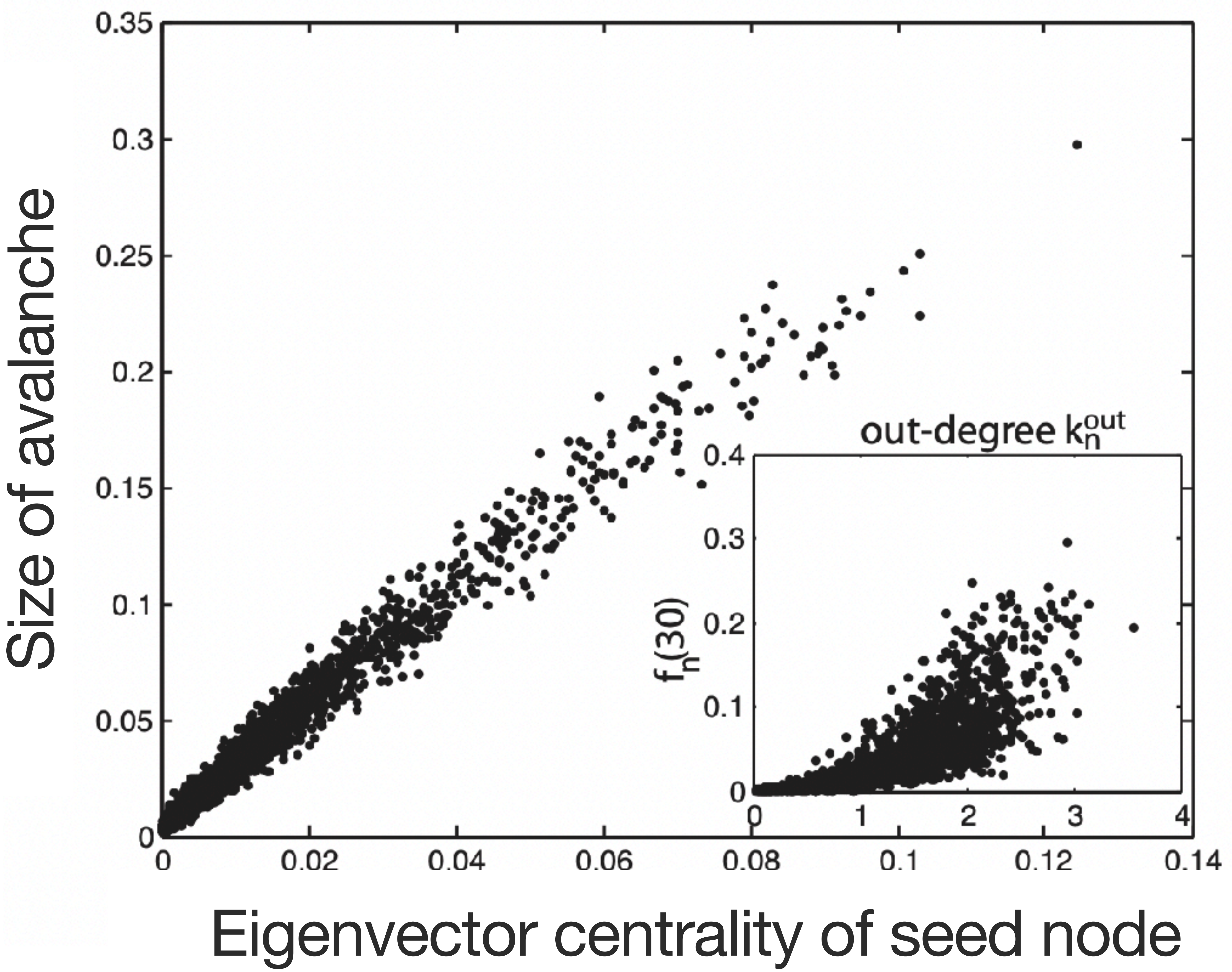
Eigenvector centrality: the centrality u_i of node i is proportional to the sum of its neighbors' centralities

$$u_i = C \sum_{j=1}^{N} A_{ij} u_j$$

→ The eigenvector centrality u_i of node i is proportional to the ith entry of the network's adjacency matrix **eigenvector** corresponding to its largest eigenvalue.

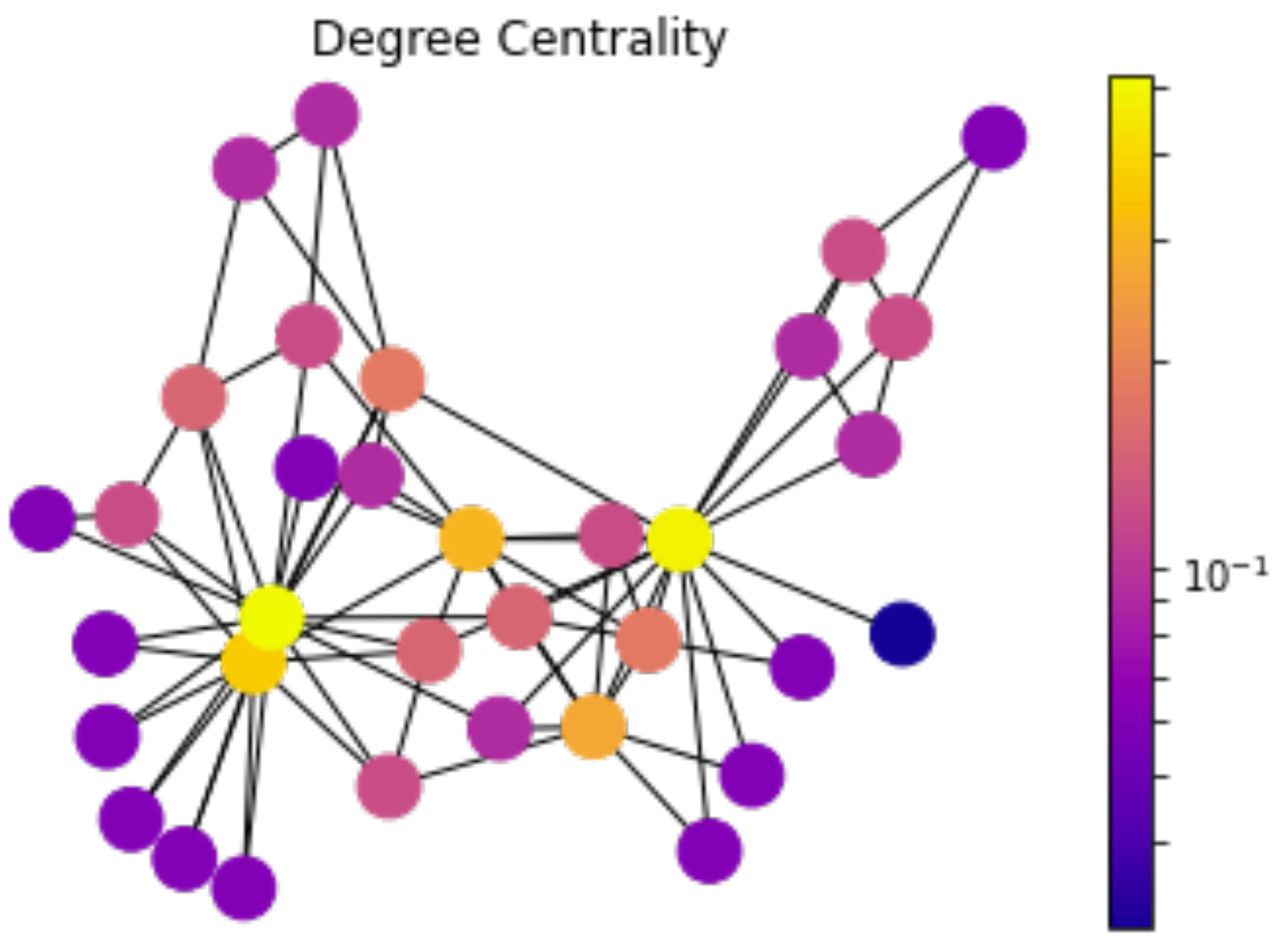
Betweenness centrality: how many shortest paths between pairs of nodes have to go through node i

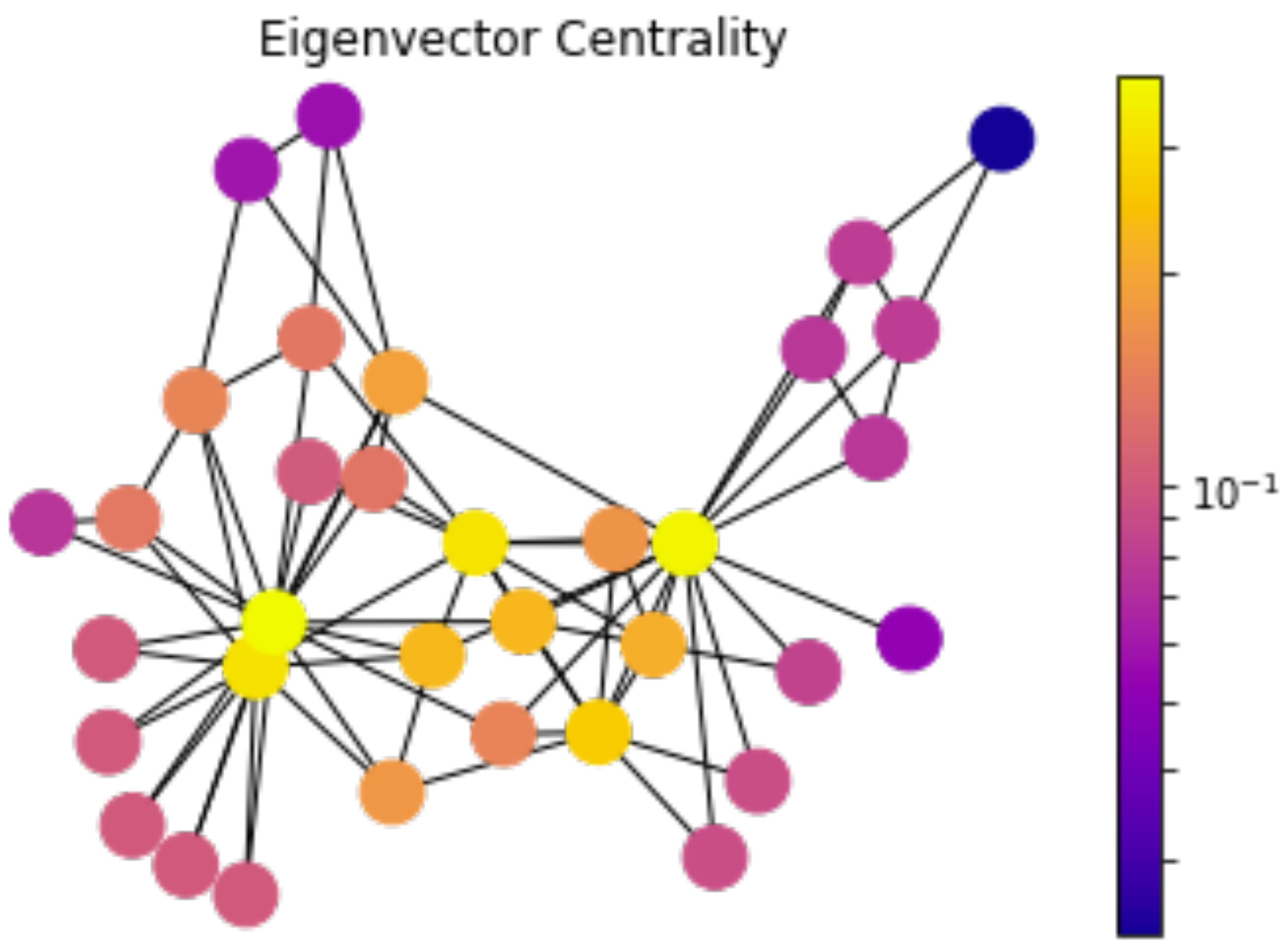
Eigenvector centrality

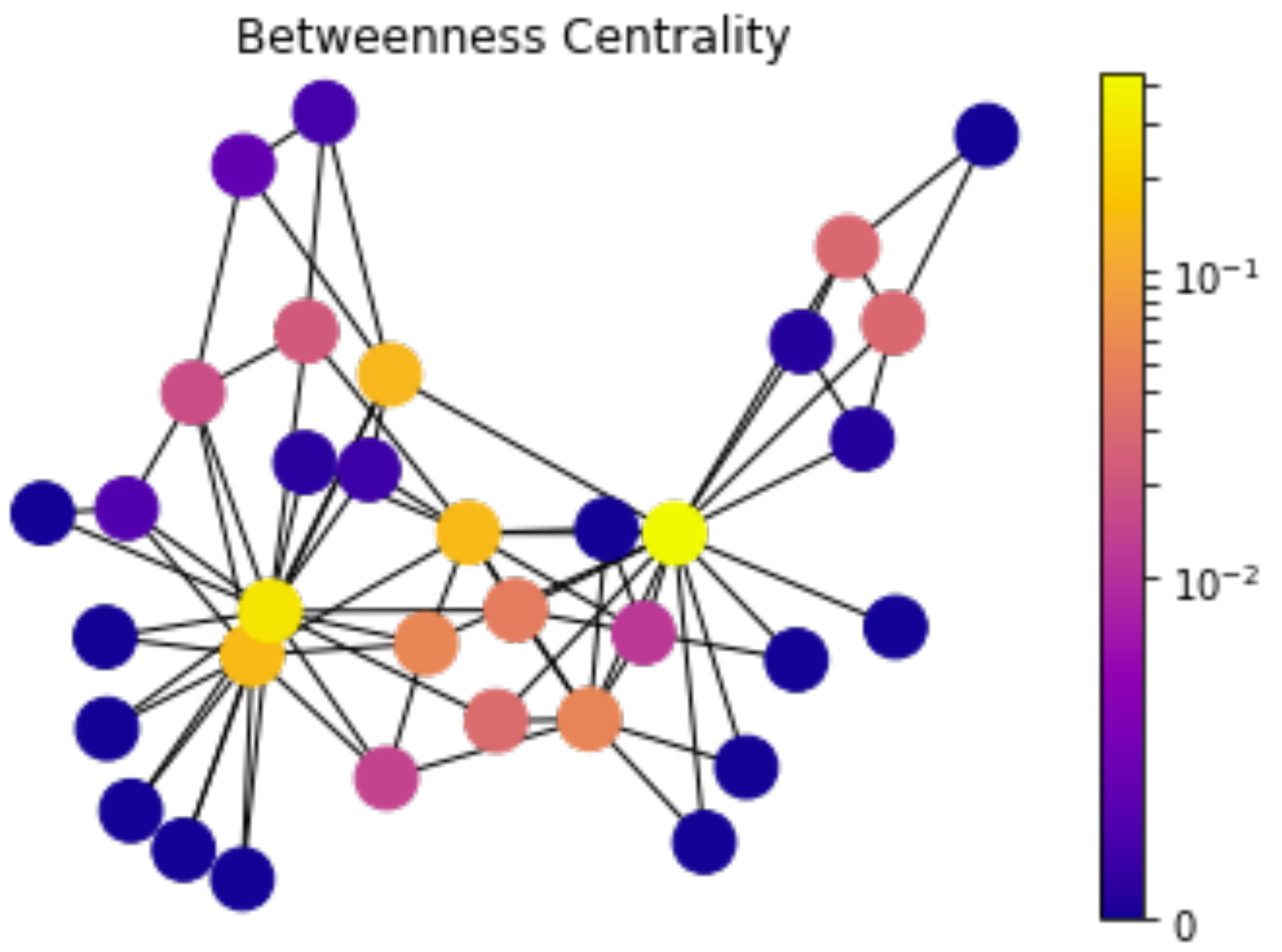


Larremore, Daniel B., et al. "Statistical properties of avalanches in networks." *Physical Review E* 85.6 (2012): 066131.

Many ways to define centrality







https://aksakalli.github.io/2017/07/17/network-centrality-measures-and-their-visualization.html

Talk outline

Part 1

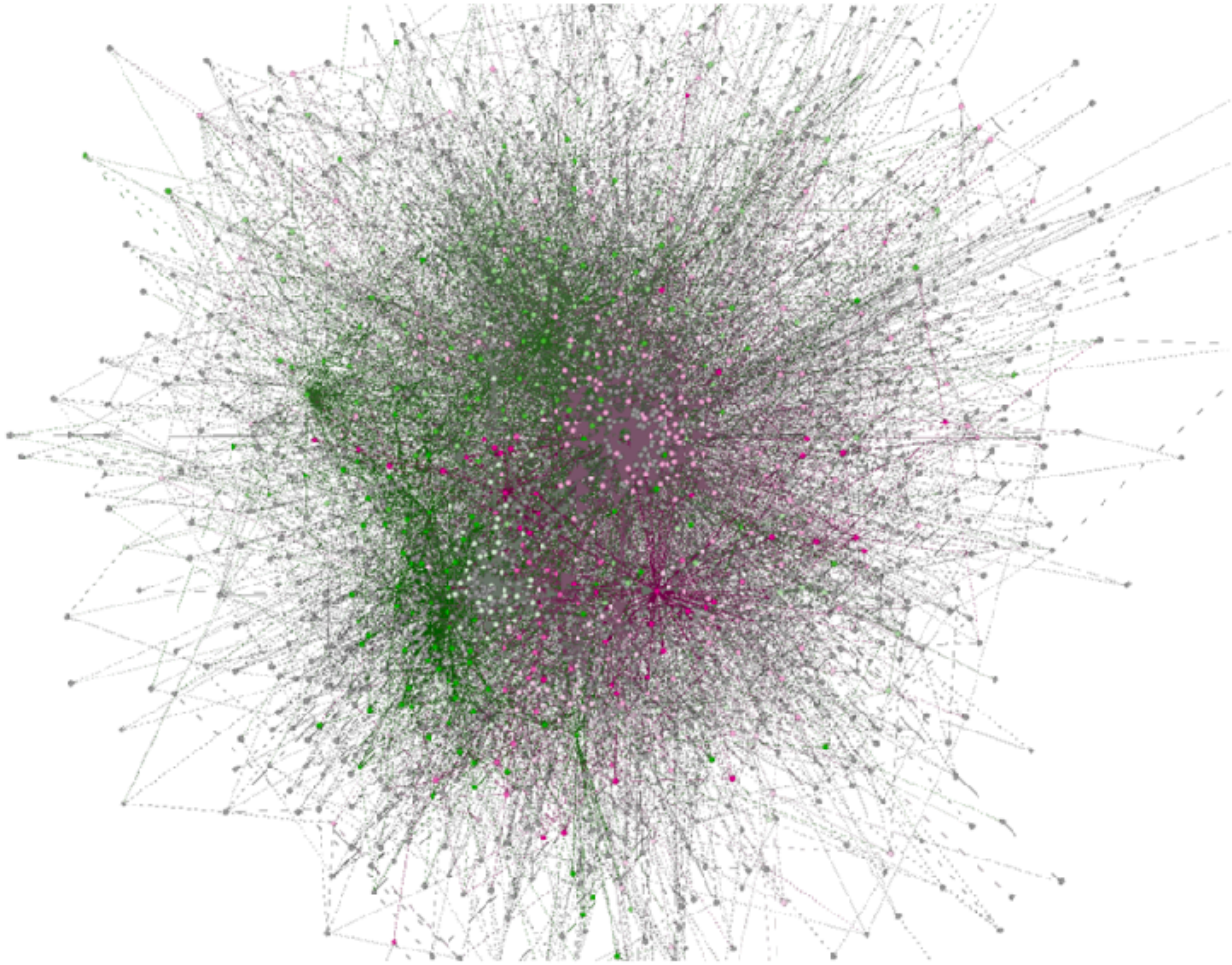
- ✓ Motivation: how are networks useful for social justice analysis?
- ✓ Basic network **properties**
- Network **models**
- Beyond networks

Part 2

- Examples

Social networks are big and messy

It is useful to have **null models** where we can remove complications and study the effect of separate features individually



Network of social interactions on a facebook page
(nodes with degree less than 6 are not shown)
Nia et al., 2012 International Conference on Social Informatics, 2012

Network models

Network models allow us to add and change properties in a **controlled** way

Some uses:

As null models: Construct a network which a desired property but otherwise random.

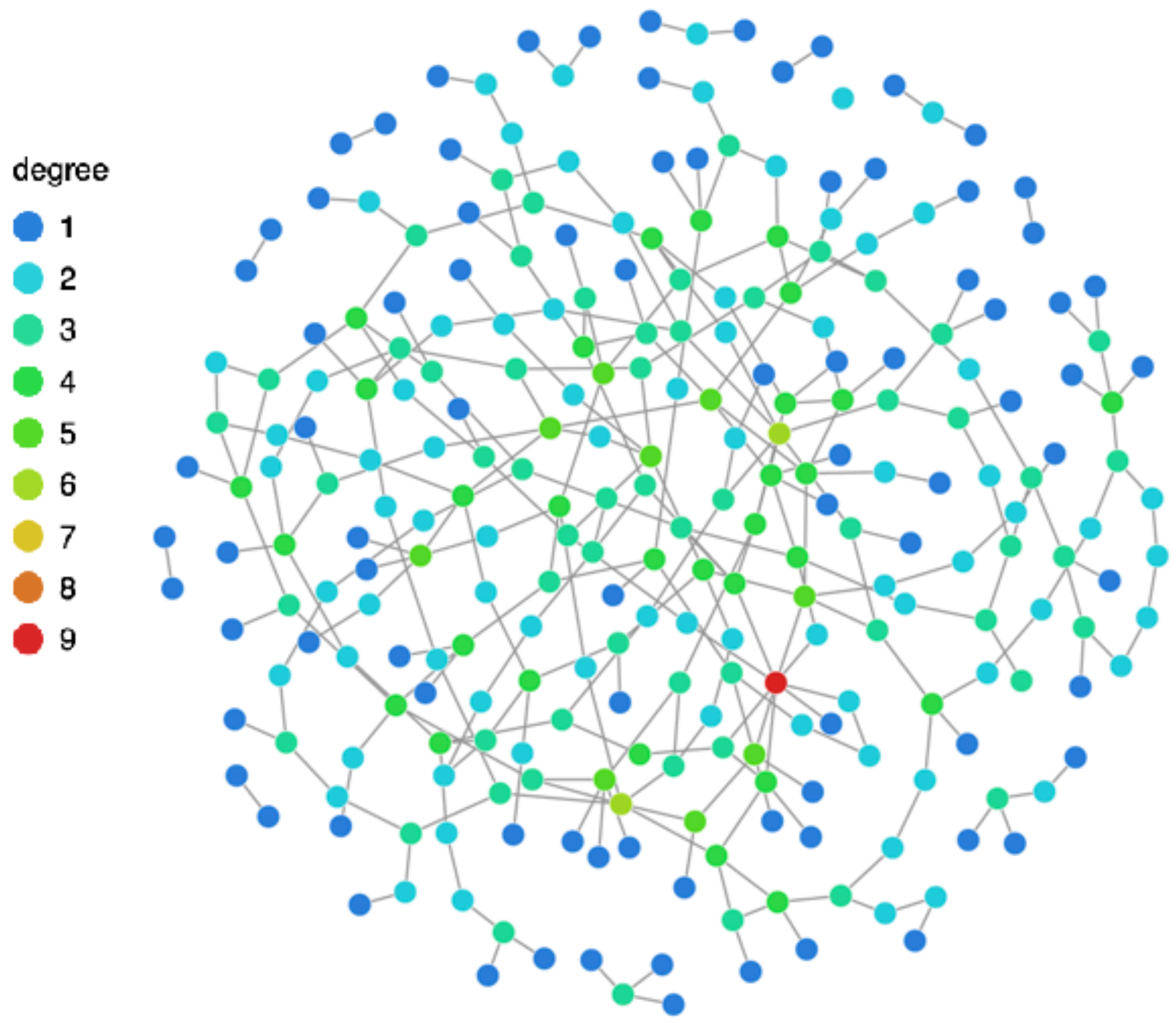
To explore the effect of network properties: Construct a network where a property can be adjusted.

As testbeds: Test your model of social dynamics in a simple network first.

The simplest: Erdös-Rényi network

$G(N, p)$: Random network with N nodes and probability of connection p

Start with N nodes, create a link between every pair with probability p



No heavy-tail

No community structure

Easy to work with

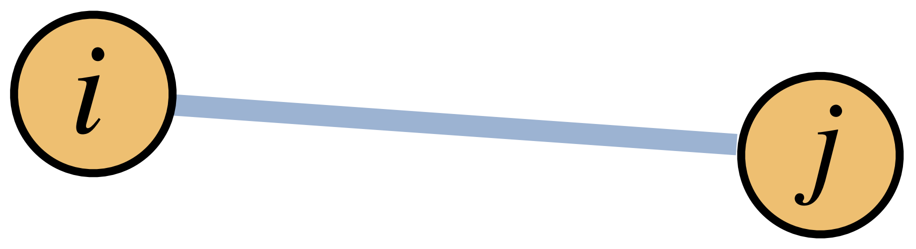
Too simple

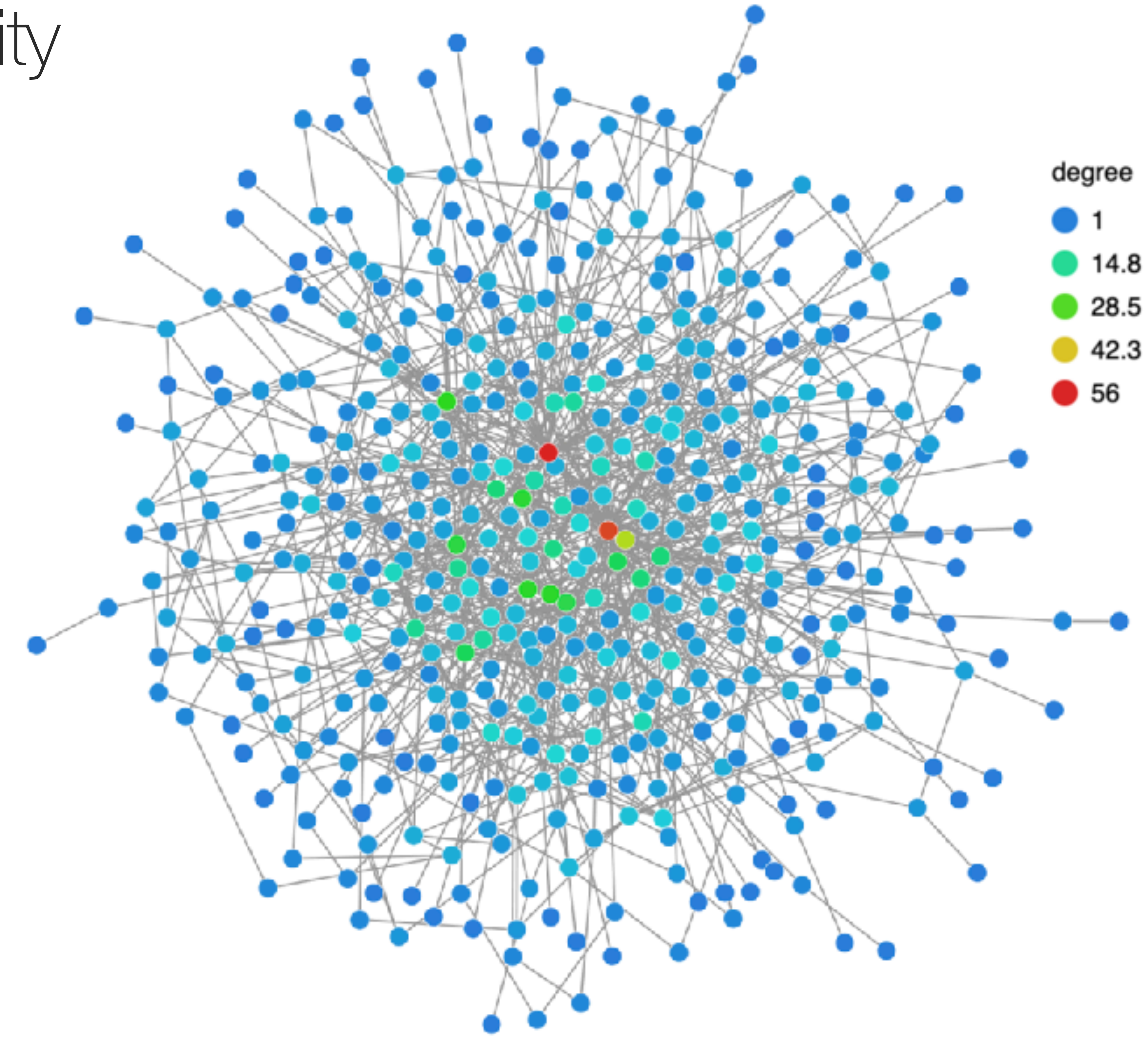
Degree-based models - Chung-Lu model

Random network with **prescribed** degree sequence $\{k_1, k_2, \ldots, k_N\}$

Start with N nodes, create a link between nodes i, j with probability

$$\frac{k_i k_j}{\sum_n k_n}$$





Degree distribution can be prescribed

No community structure

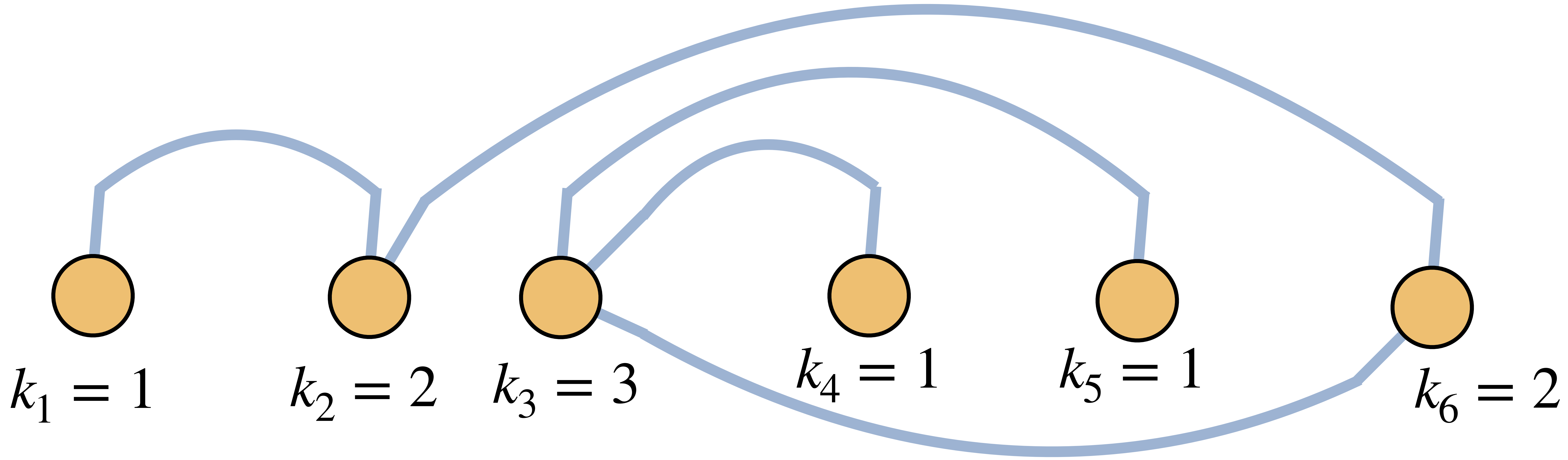
Chung, Fan, Linyuan Lu, and Van Vu. "The spectra of random graphs with given expected degrees." Internet Mathematics 1.3 (2004): 257-275.

Degree-based models - configuration model

Random network with **prescribed** degree sequence $\{k_1, k_2, \ldots, k_N\}$

Assign stubs to each node according to their degree

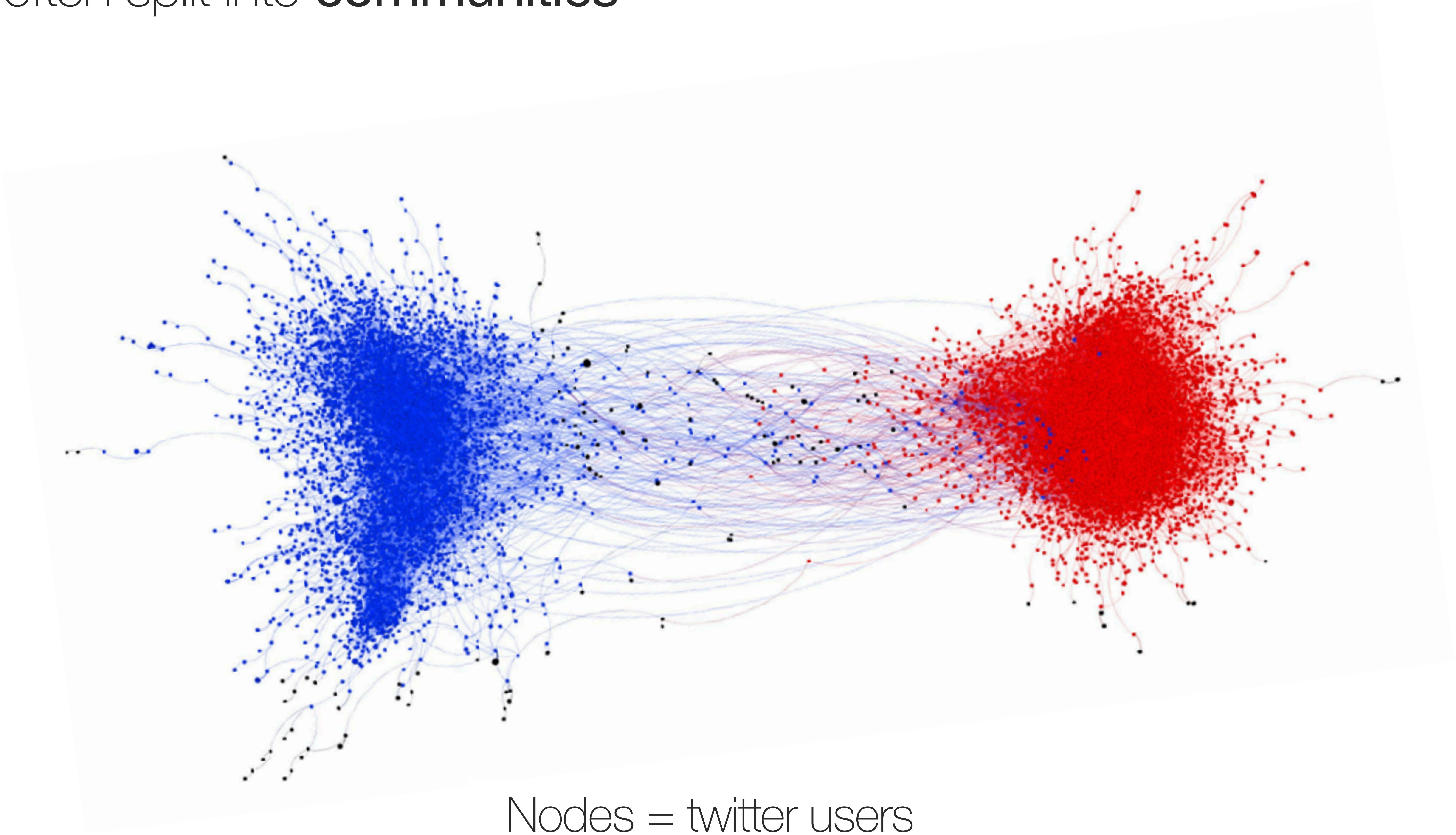
Match pairs of stubs randomly



Fosdick, Bailey K., et al. "Configuring random graph models with fixed degree sequences." Siam Review 60.2 (2018): 315-355.

Community structure

Nodes in networks often split into **communities**



Nodes = twitter users
Links = retweets
Color = ideology

Brady et al. PNAS 2016

Stochastic block model

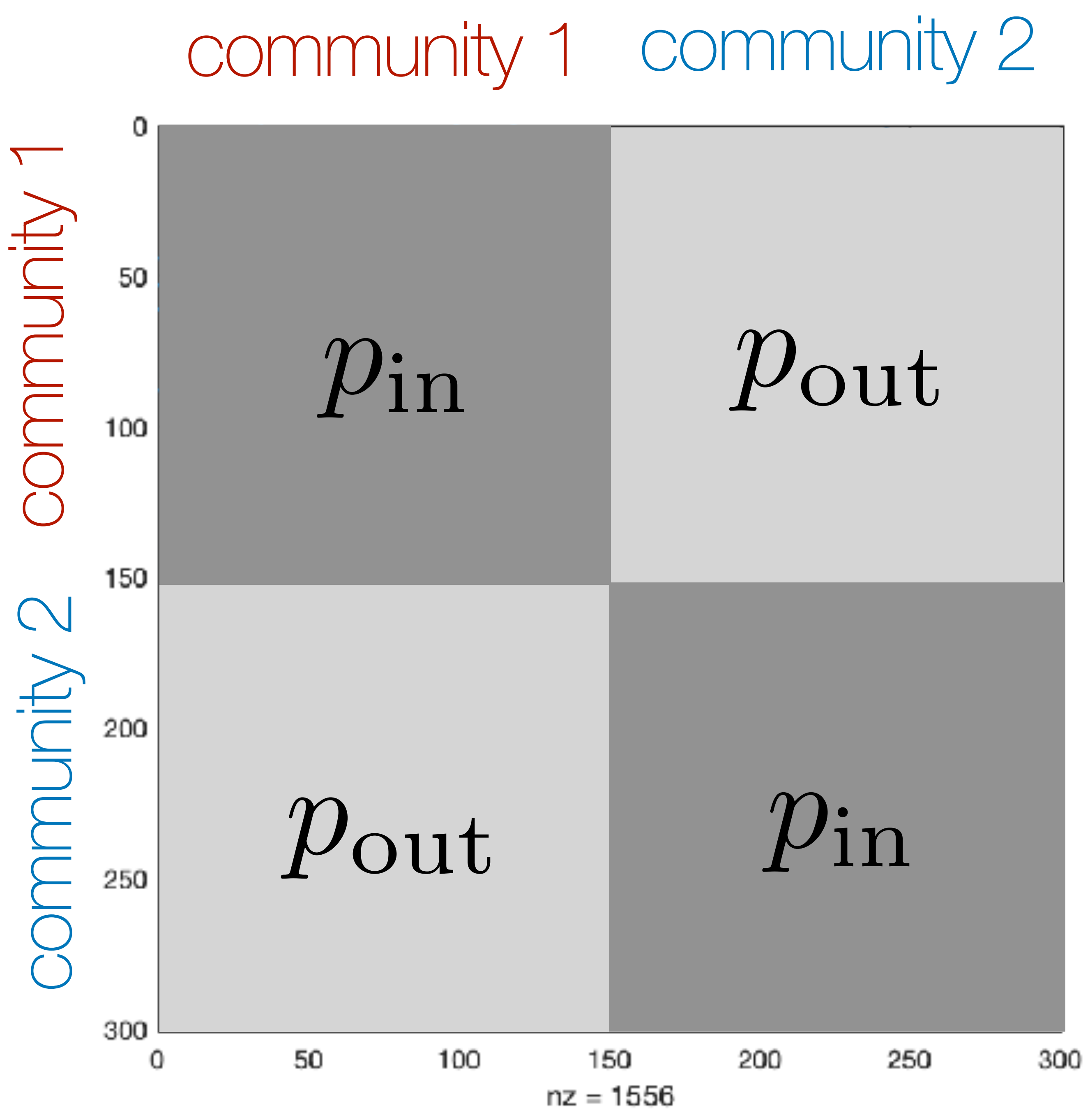
Specify communities in advance, and the create links with higher probability for nodes within the same community

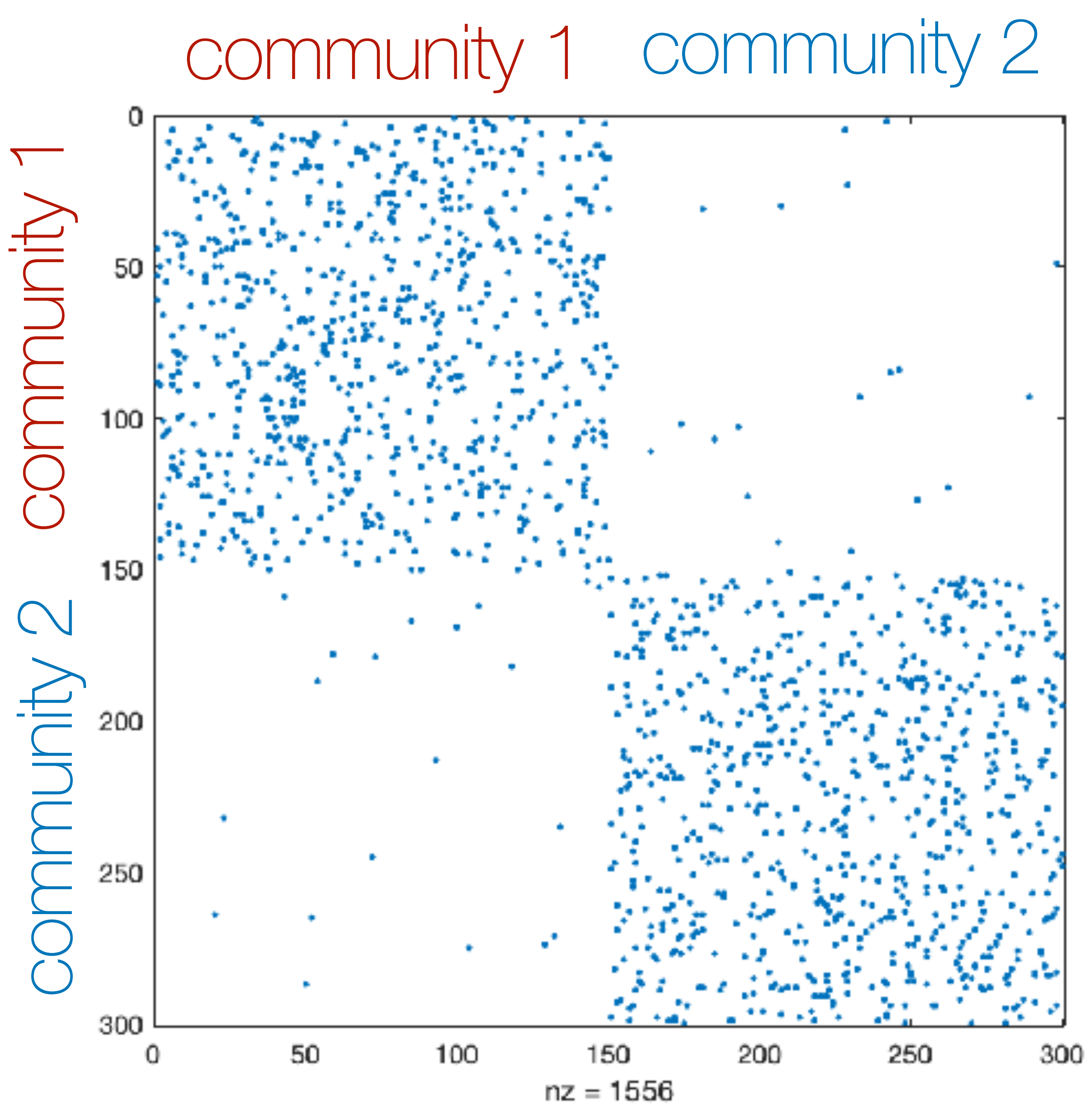
Start with N nodes, create a link between every pair of nodes i, j with probability

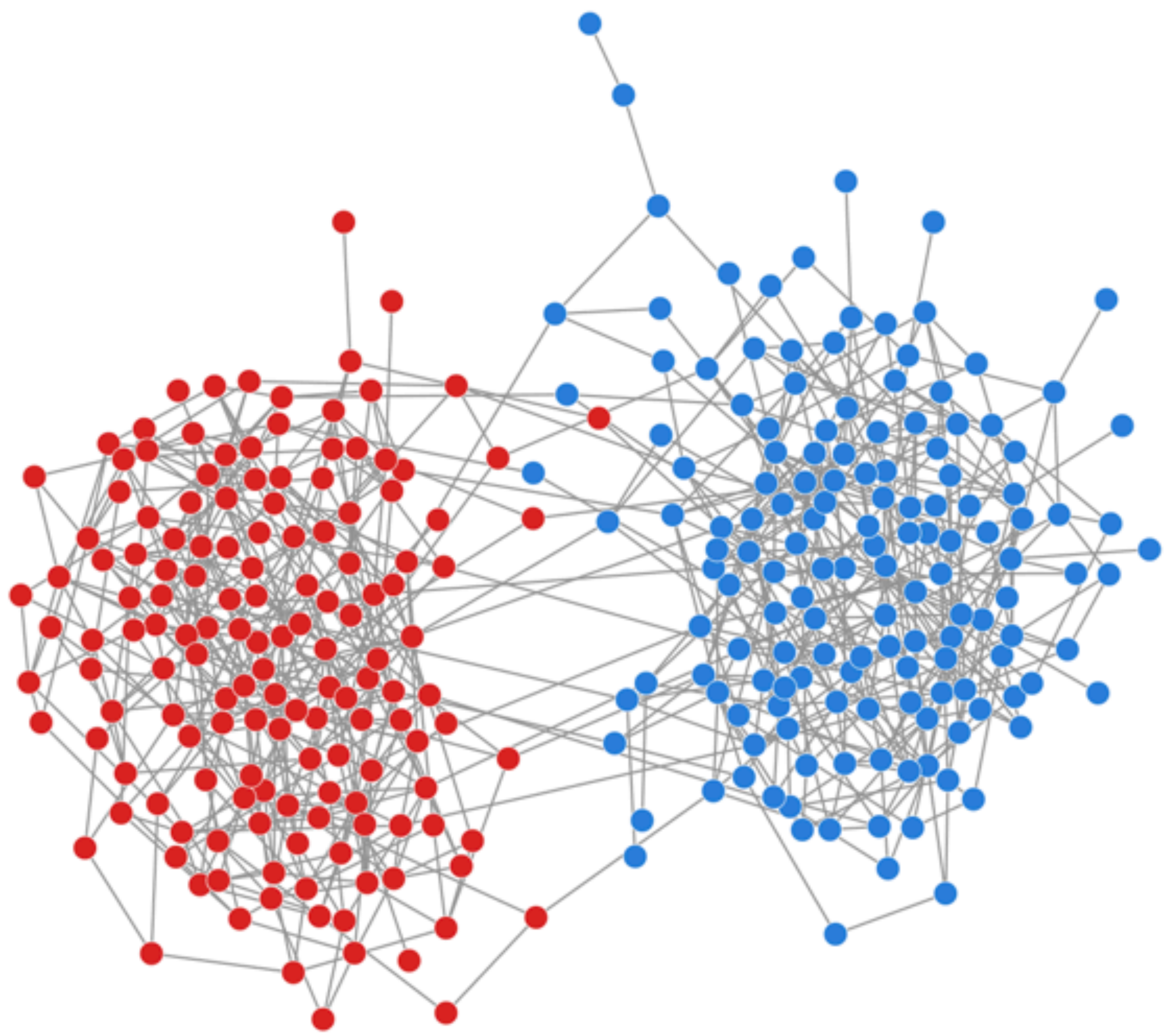
$$\begin{cases} p_{\text{in}} & \text{i, j in the same community} \\ p_{\text{out}} & \text{i, j in different community} \end{cases}$$

$$p_{\text{in}} > p_{\text{out}}$$

Community structure can be prescribed







Holland, Paul W., Kathryn Blackmond Laskey, and Samuel Leinhardt. "Stochastic blockmodels: First steps." Social networks 5.2 (1983): 109-137.

Generalized stochastic block models

Specify node characteristics (race, age, community, etc) in a vector $\mathbf{z}_i$

Probability that two nodes are connected is a function of $\mathbf{z}_i$, $\mathbf{z}_j$:

$$P(A_{ij} = 1) = f(\mathbf{z}_i, \mathbf{z}_j)$$

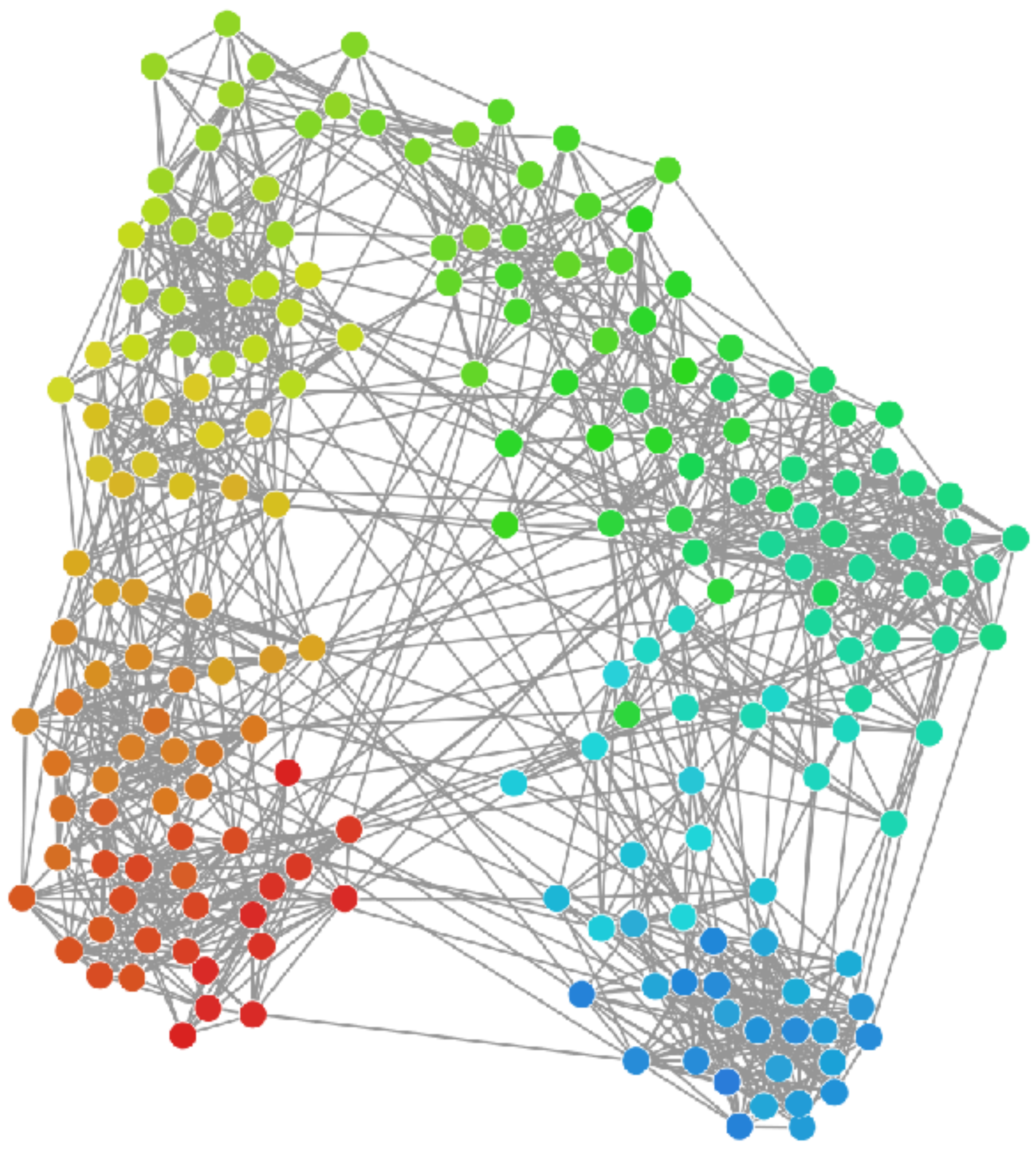
Very flexible; characteristics can be:

Intended nodal degree (Chung-Lu model)

Community index (previous slide)

Demographical metadata

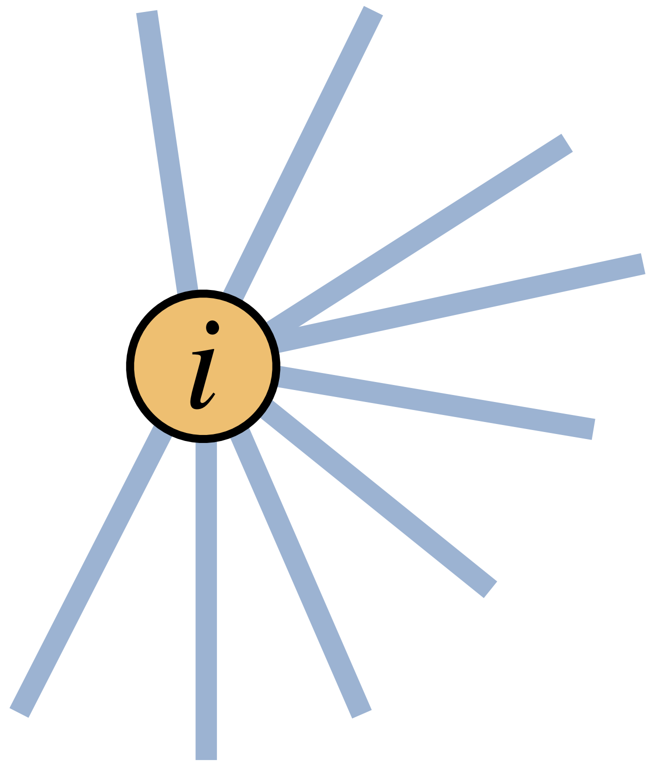
Dynamical parameters (e.g., infectivity for epidemic models, frequency for oscillator networks)



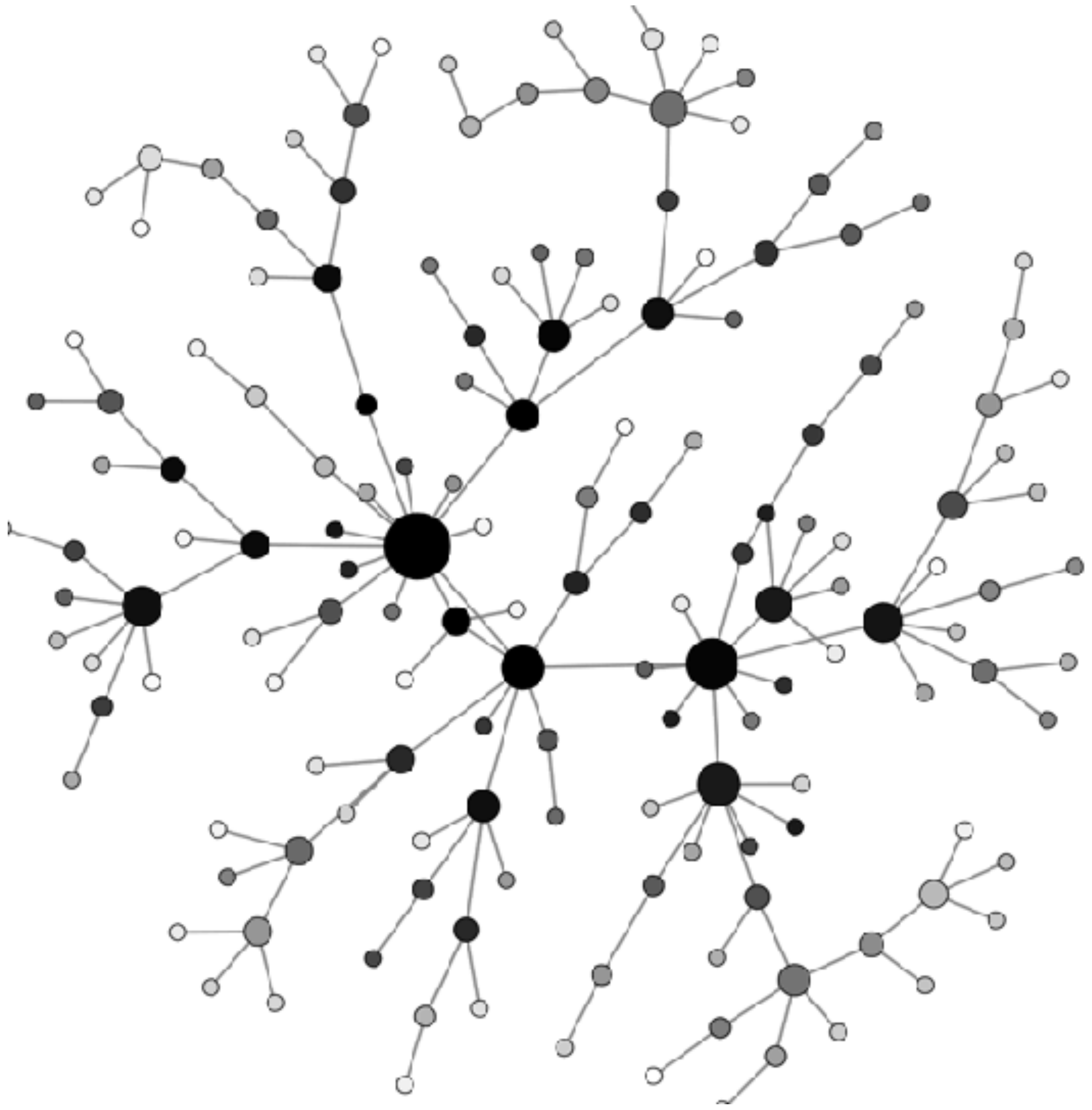
Growing network models

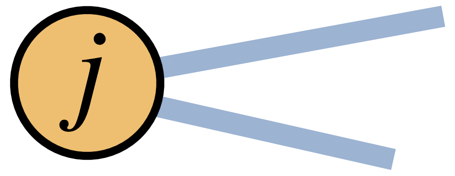
Preferential attachment mechanism: originally proposed by Price, popularized by Barábasi-Albert

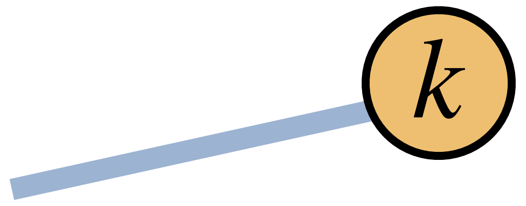
Nodes join the network and connect preferentially to nodes that have a high degree already



First-mover advantage







Price, Derek J. De Solla. "Networks of scientific papers: The pattern of bibliographic references indicates the nature of the scientific research front." Science 149.3683 (1965): 510-515.

Barabási, Albert-László, and Réka Albert. "Emergence of scaling in random networks." science 286.5439 (1999): 509-512.

Evolving network models

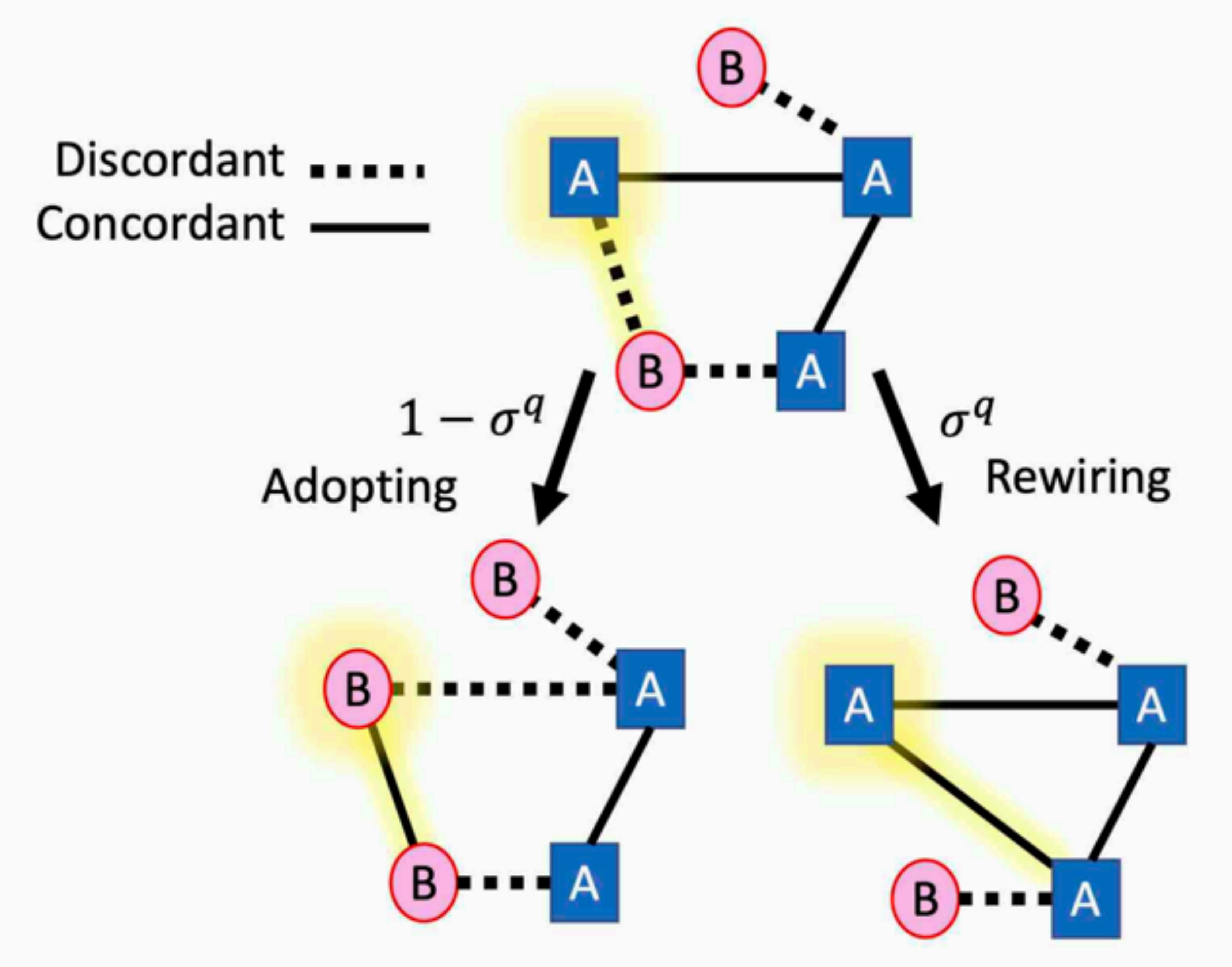
Time-evolving network models could be very relevant for social justice modeling

(an example will be shown in Part II)

Example: network and opinions coevolving

Fitting in and breaking up: A nonlinear version of coevolving voter models

Yacoub H. Kureh and Mason A. Porter
Phys. Rev. E **101**, 062303 – Published 2 June 2020



Talk outline

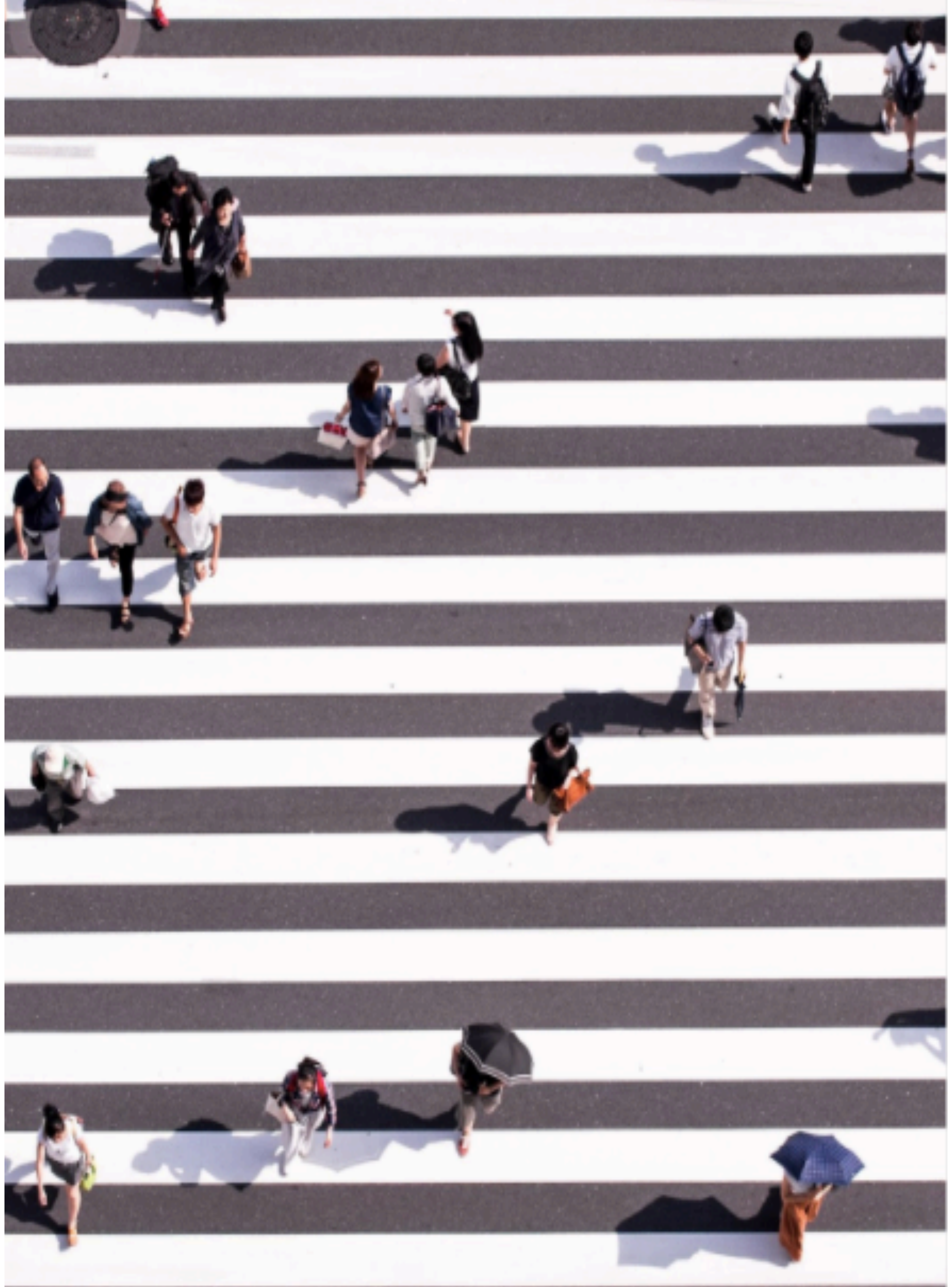
Part 1

- Motivation: how are **networks useful** for **social justice** analysis?
- ✓ Basic network **properties**
- ✓ Network **models**
- Beyond networks

Part 2

- Examples

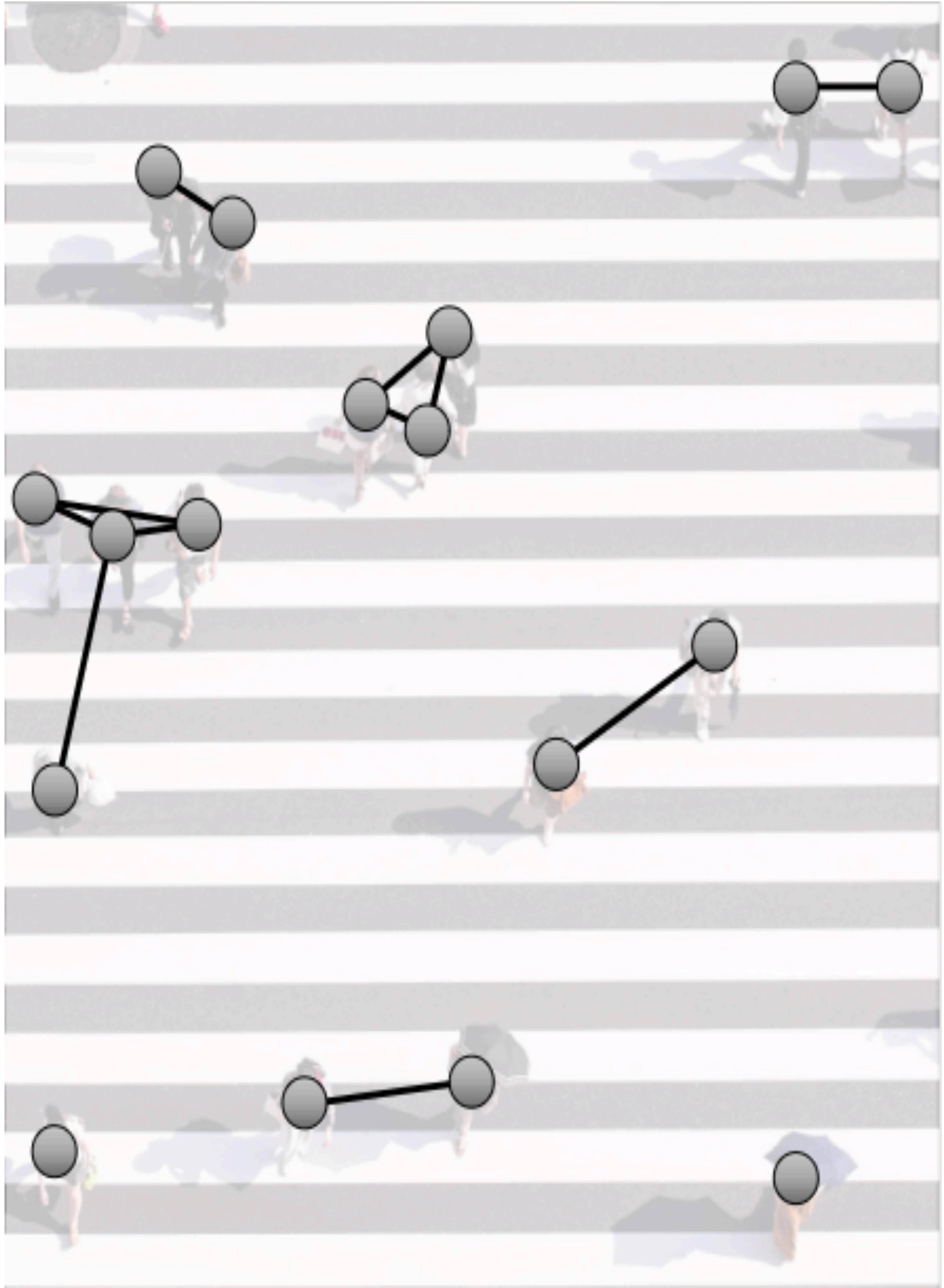
Beyond networks



Slide from Nicholas Landry

Networks

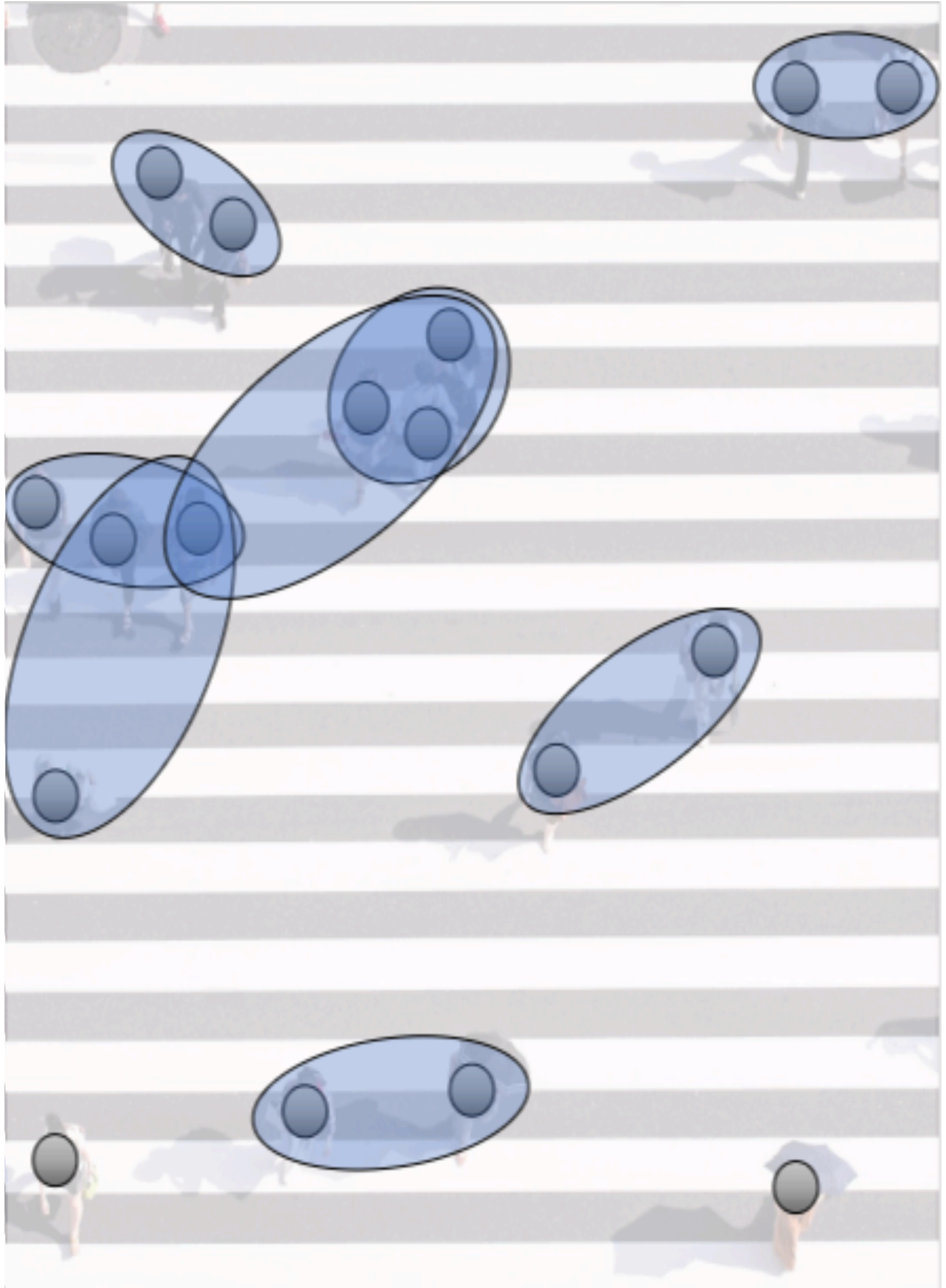
Nodes and links



Slide from Nicholas Landry

Hypergraphs

Nodes and hyperedges

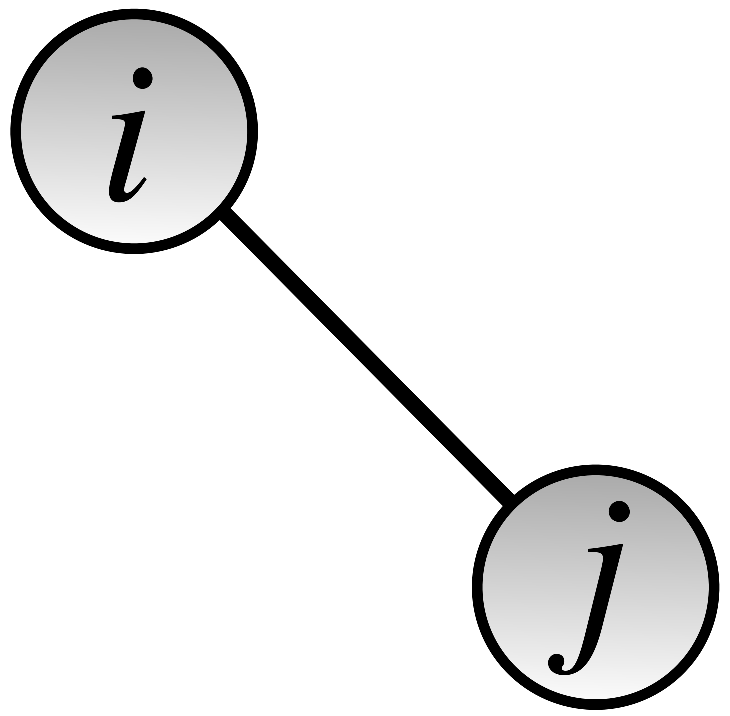


Slide from Nicholas Landry

What is a hypergraph?

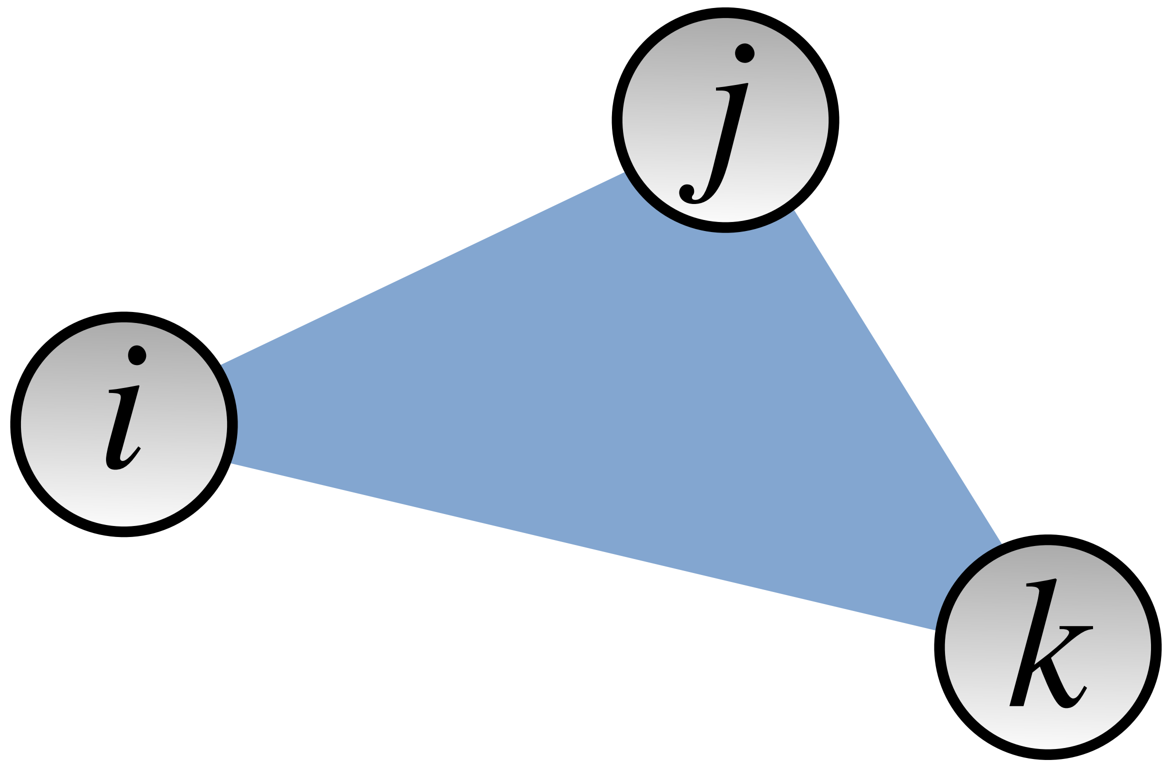
Network

Nodes: {1,2,3,…N},
Edges: pairs of nodes



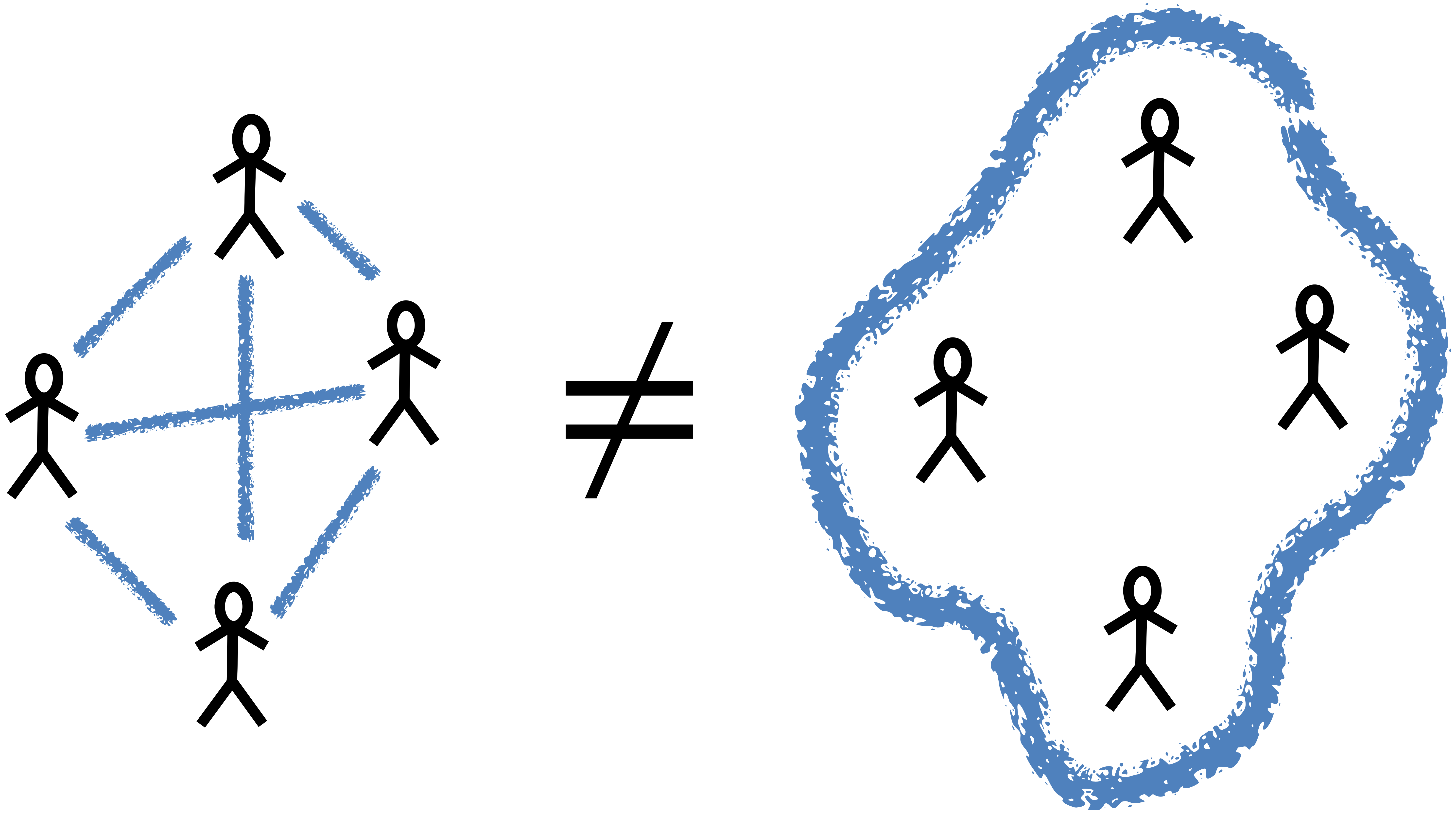
Hypergraph

Nodes: {1,2,3,…N},
Hyperedges: sets of nodes



Hypergraphs

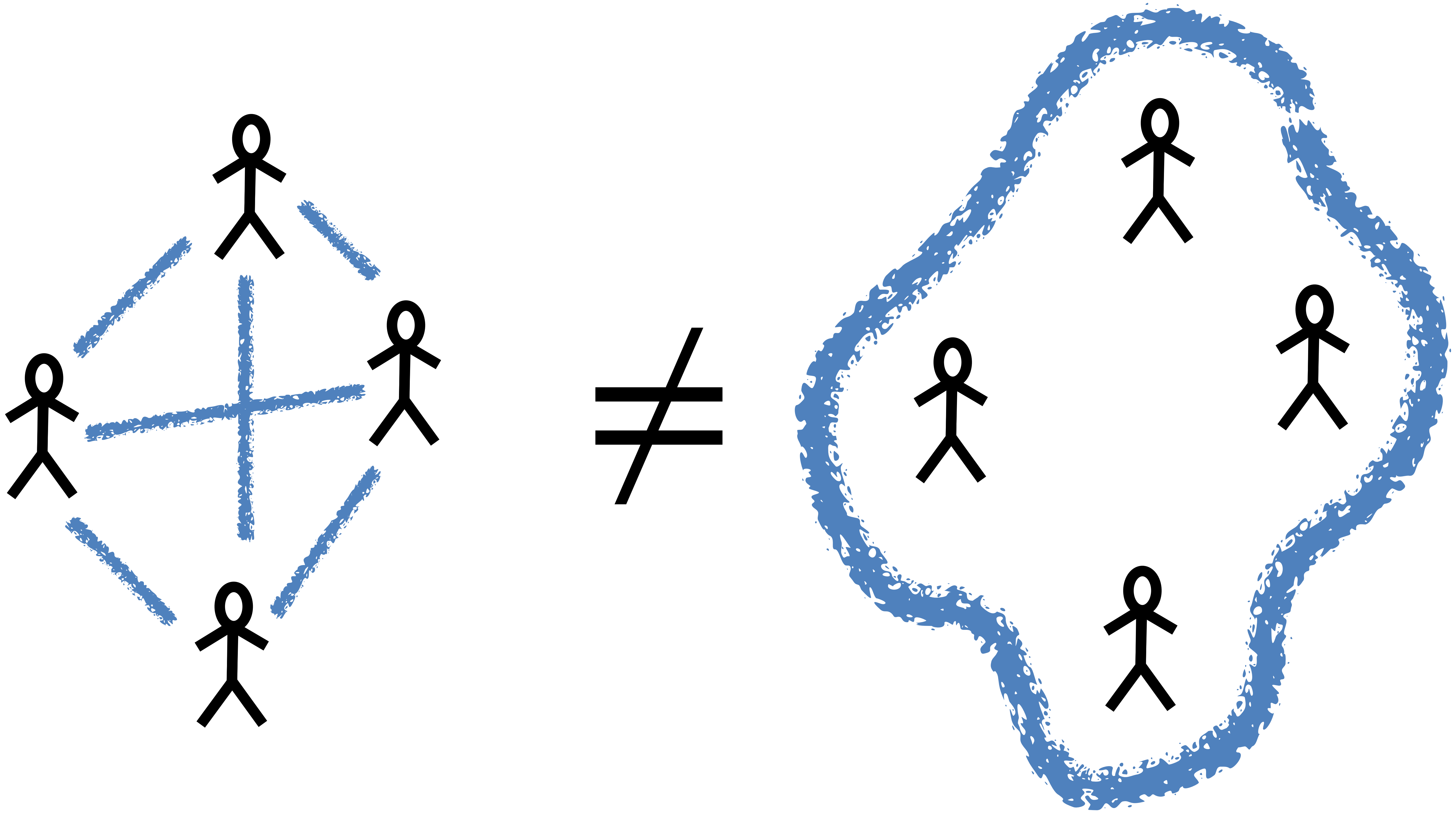
Opinion models



Multiple one-on-one discussions are not
equivalent to one group discussion

Hypergraphs

Scientific collaborations



Multiple pairwise joint publications are not equivalent to a single joint publication.

Dynamics on Hypergraphs

Review: Battiston et al., Physics Reports (2020)

Synchronization:

P. S. Skardal, A. Arenas, PRL (2019)
A. P. Millán, J.J. Torres, G. Bianconi (2020)

Opinion models:

L. Neuhäuser et al., arXiv 2004.00901
A. Hickok et al., SIAM JAPD, (2021)

Contagion models

I. Iacopini, G. Petri, A. Barrat, V. Latora, Nature Communications (2019)
B. Jhun, M. Jo, B. Kahng, Journal of Statistical Mechanics (2019).
G. F. de Arruda, G. Petri, Y. Moreno, PRR (2020).

Part II

Examples of simple spreading processes on networks

Demonstrations of network creation and manipulation in matlab and python

Example of diversity/productivity modeling