UNIVERSITÀ DEGLI STUDI DI PADOVA

Network Science

A.Y. 23/24

ICT for Internet & multimedia, Data science, Physics of data

Homophily and Polarization

an overview

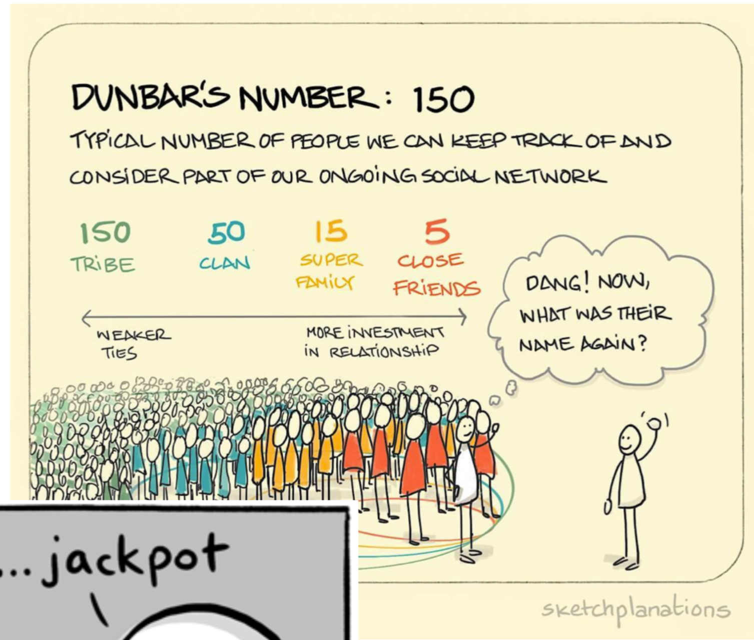
UNIVERSITÀ DEGLI STUDI DI PADOVA

Humans and social media

We have access to an unlimited amount of information, but we follow a limited number of sources

Because we are…

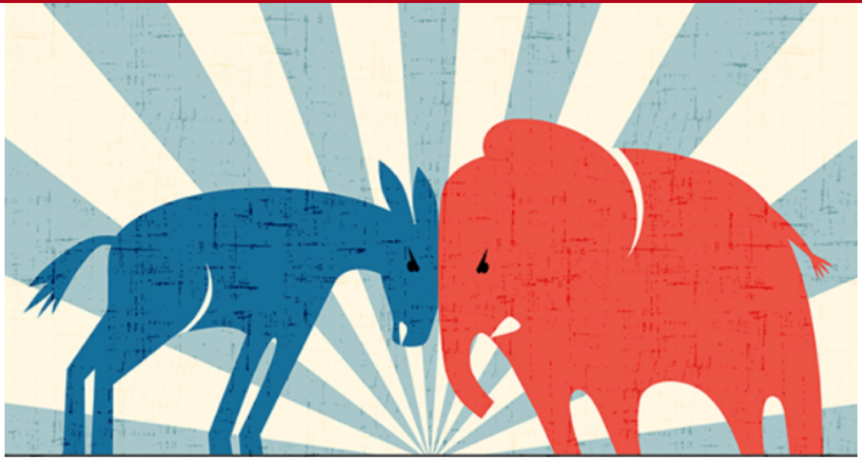
Bounded



Biased

Effects on online behaviour

Polarization

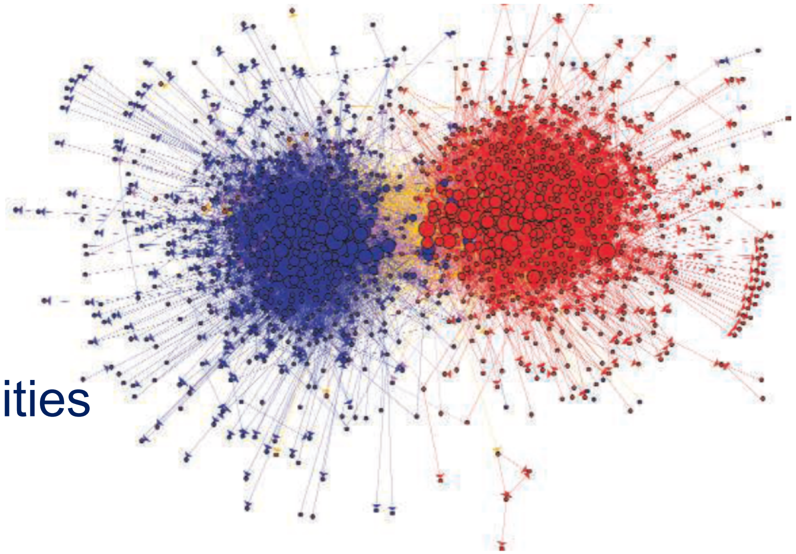


Homophily

Selective exposure

Homophily

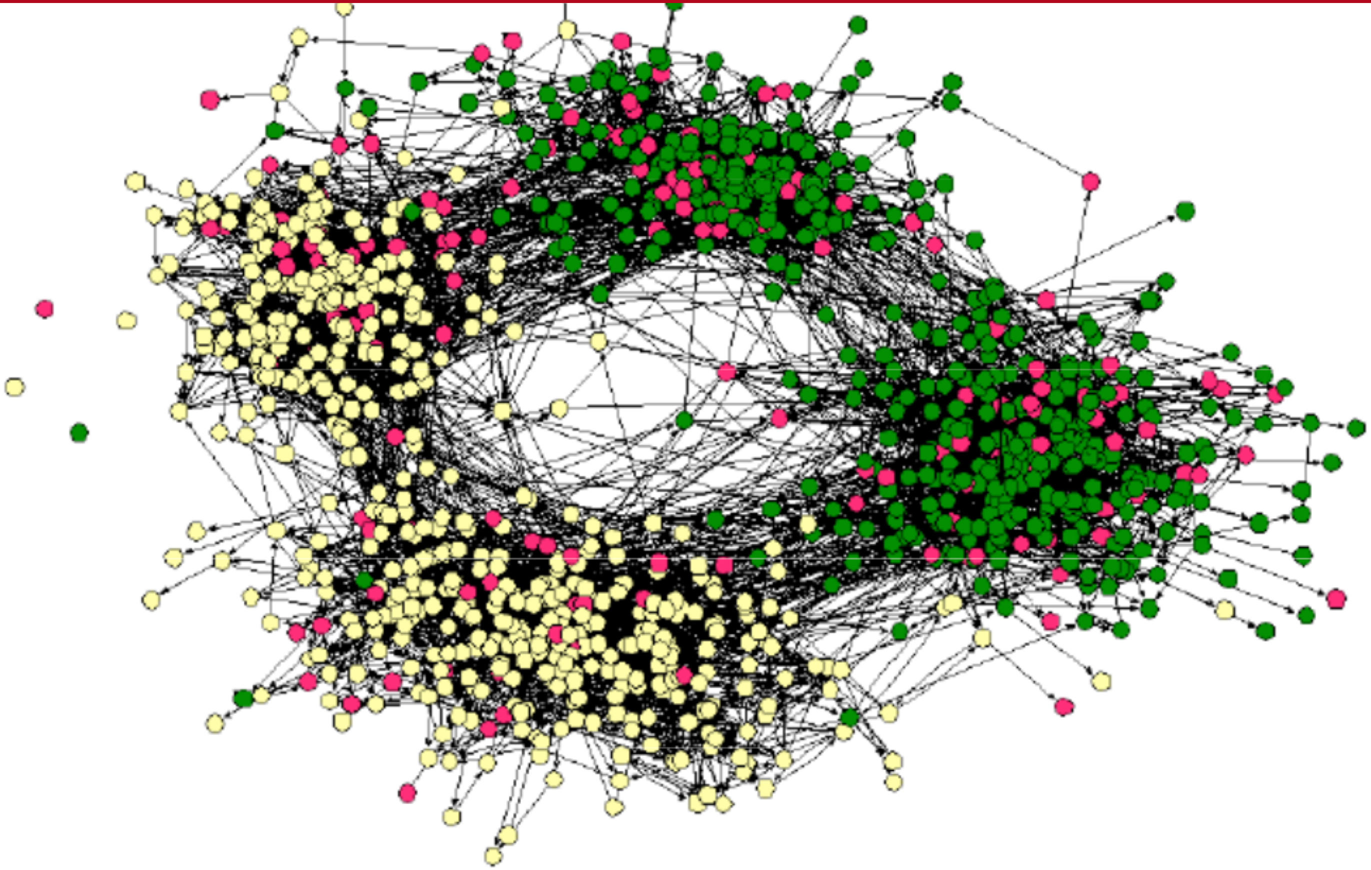
Homophily (from Ancient Greek: *homoû*, 'together' + *philíē*, 'friendship, love') is the tendency of individuals to associate and bond with similar others, as in the proverb "birds of a feather flock together."[1] The presence of homophily has been discovered in a vast array of network studies: over 100 studies have observed homophily in some form or another, and they establish that similarity is associated with connection.[2] The categories on which homophily occurs include age, gender, class, and organizational role.



Political blog communities

Homophily in action

racial segregations

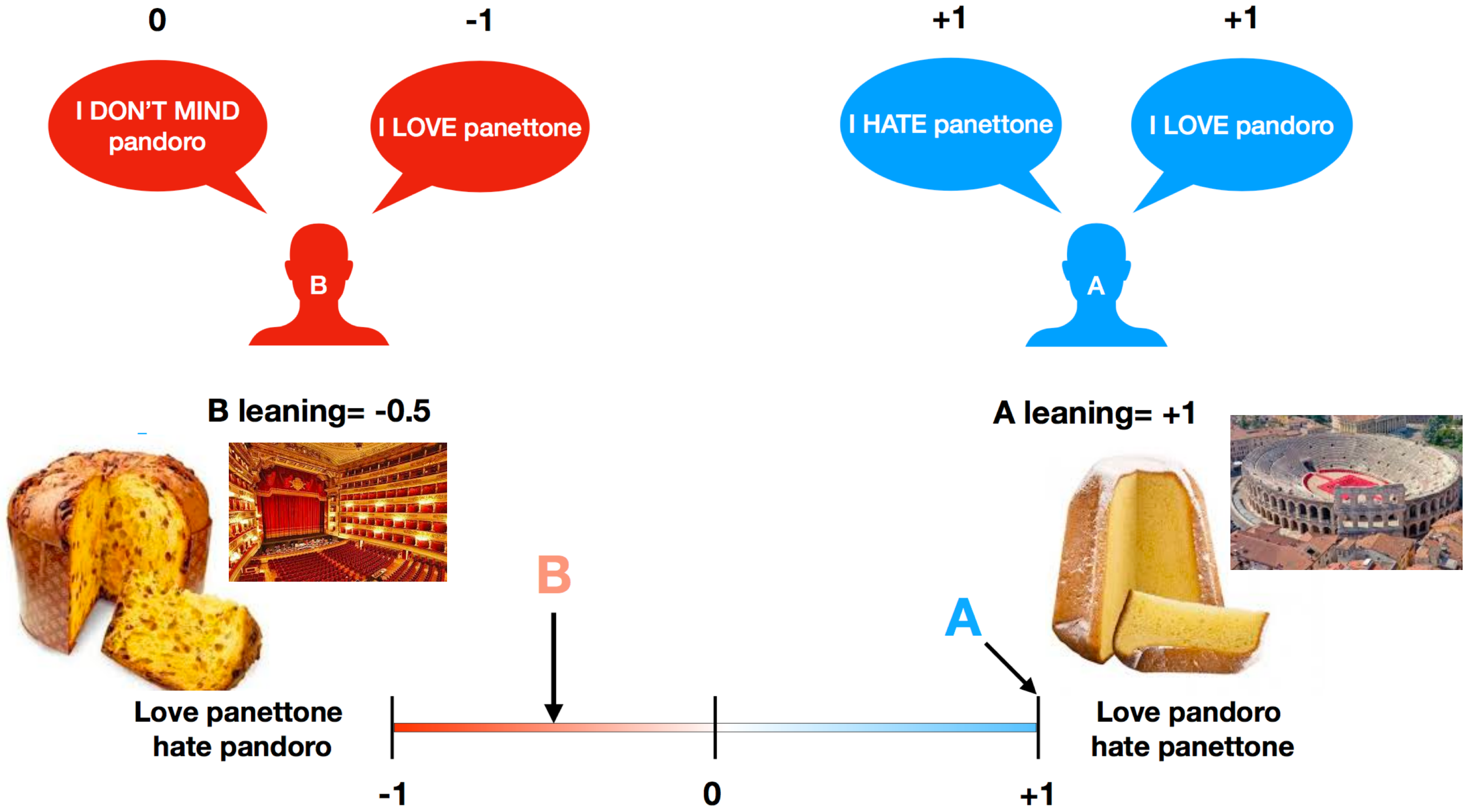


(Easley and Kleinberg, 2010)

Figure 4.1: Homophily can produce a division of a social network into densely-connected, homogeneous parts that are weakly connected to each other. In this social network from a town's middle school and high school, two such divisions in the network are apparent: one based on race (with students of different races drawn as differently colored circles), and the other based on friendships in the middle and high schools respectively [304].

UNIVERSITÀ DEGLI STUDI DI PADOVA

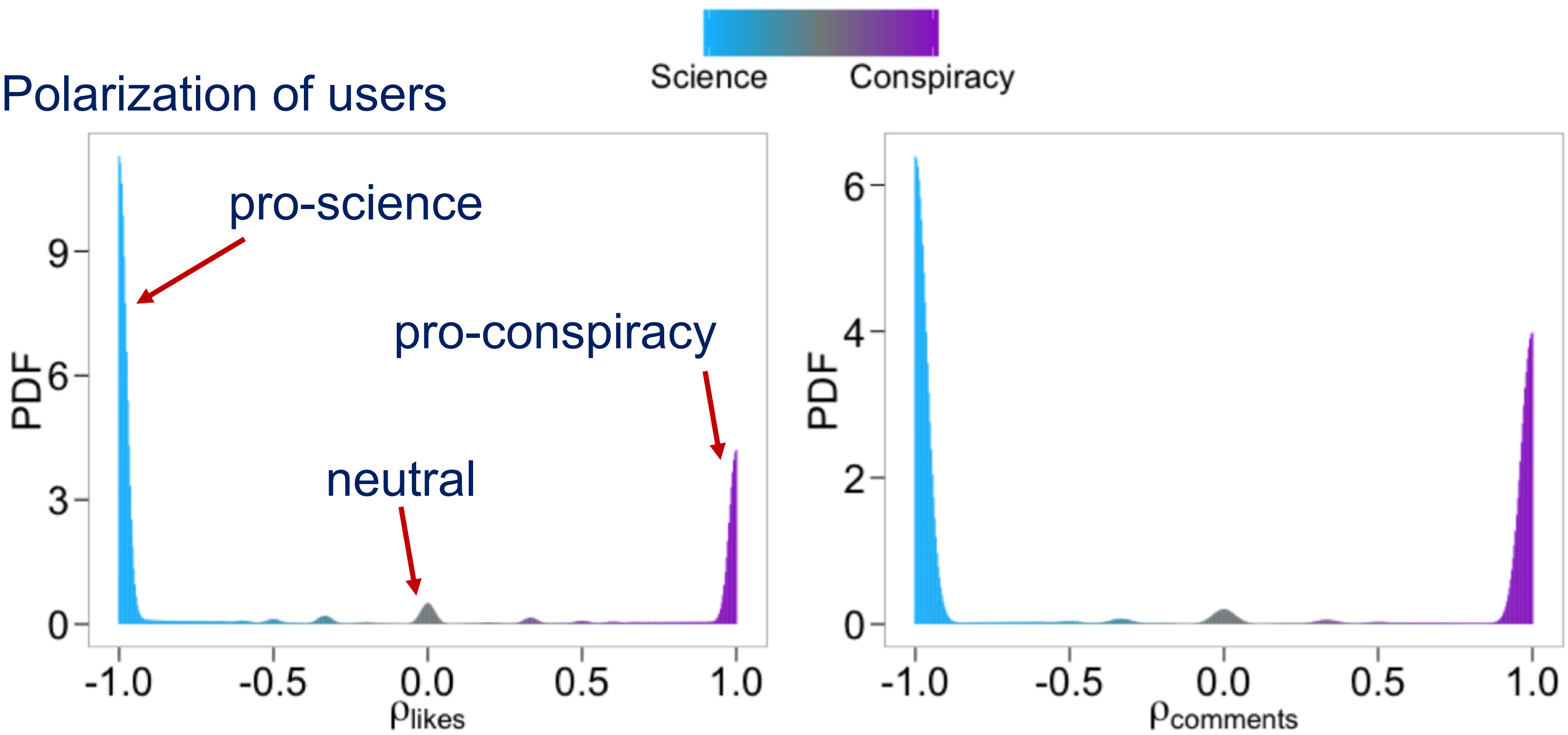
Users leaning
on a controversial topic



Polarization

The extreme segregation of users into homogeneous communities based on their opinion on a controversial topic

Polarization of users

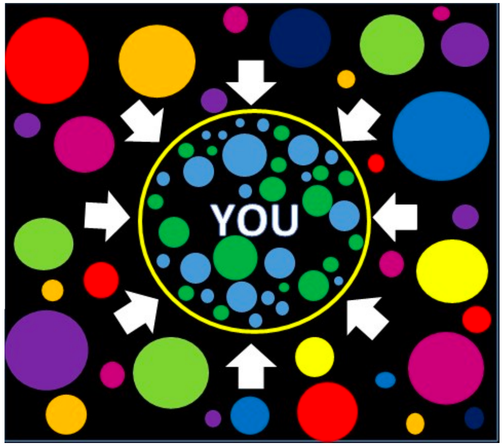


Echo chambers

Echo chamber (media)

From Wikipedia, the free encyclopedia

In news media, an **echo chamber** is a metaphorical description of a situation in which beliefs are amplified or reinforced by communication and repetition inside a closed system and insulates them from rebuttal.[1] By visiting an "echo chamber", people are able to seek out information that reinforces their existing views, potentially as an unconscious exercise of confirmation bias. This may increase social and political polarization and extremism.[2] The term is a metaphor based on the acoustic echo chamber, where sounds reverberate in a hollow enclosure. Another emerging term for this echoing and homogenizing effect on the Internet within social communities, such as Facebook, Instagram, Twitter, Reddit, etc; is cultural tribalism.[3]



Echo chamber

a formalization

Cinelli, Morales, Galeazzi, Quattrociocchi, Starnini (2020)
Echo chambers on social media: A comparative analysis
https://arxiv.org/pdf/2004.09603.pdf

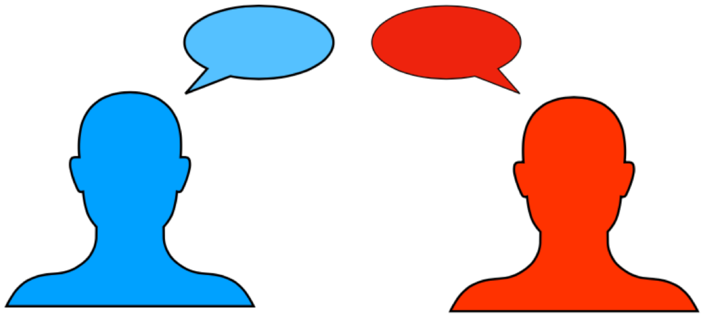
Coexistence of

- opinion polarization with respect to a controversial topic
- homophily in interactions

To assess **polarization**

↓

Measure **individual leaning**

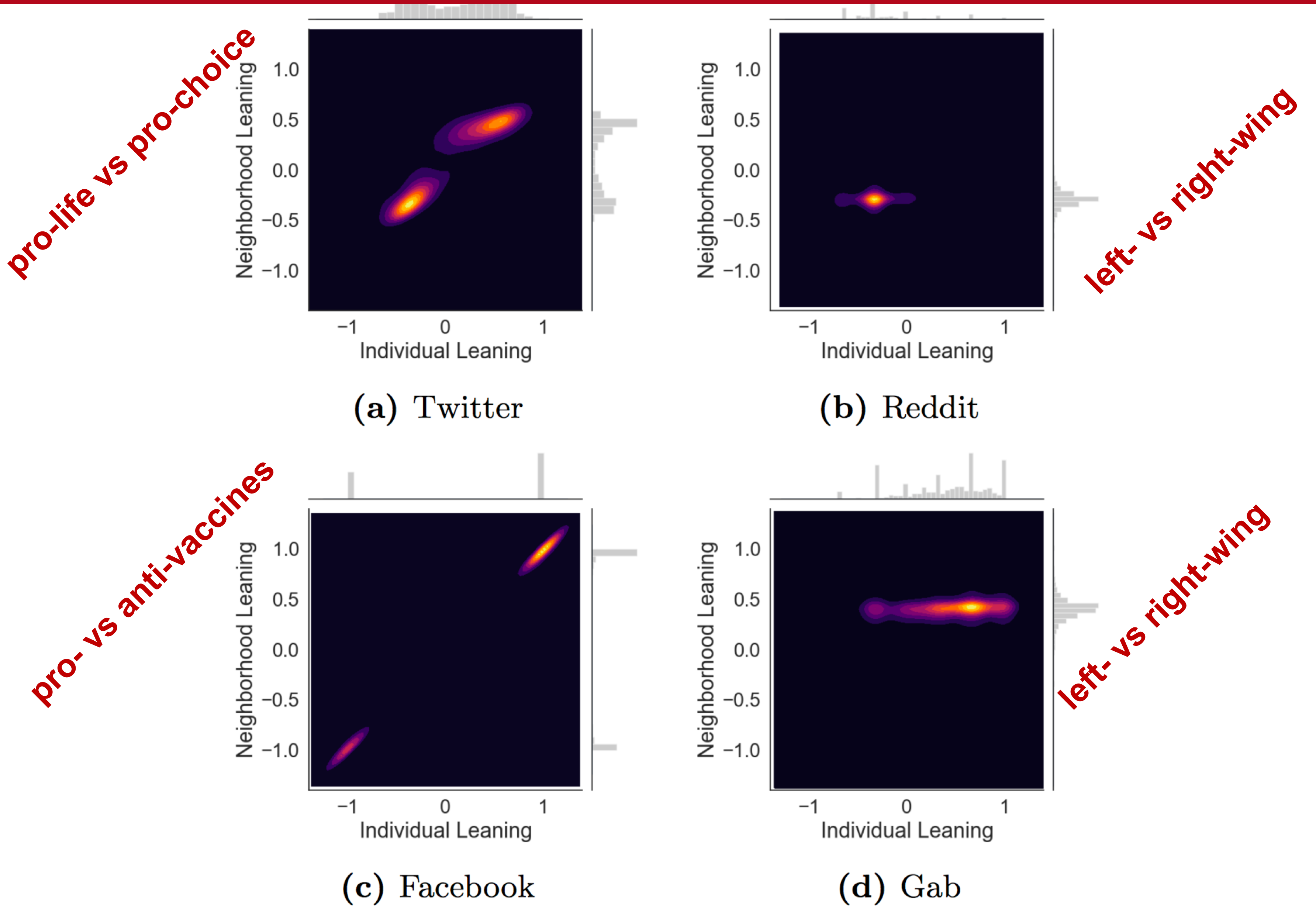


To quantify **homophily**

↓

Build **interaction network**

Echo chamber effect
in social networks



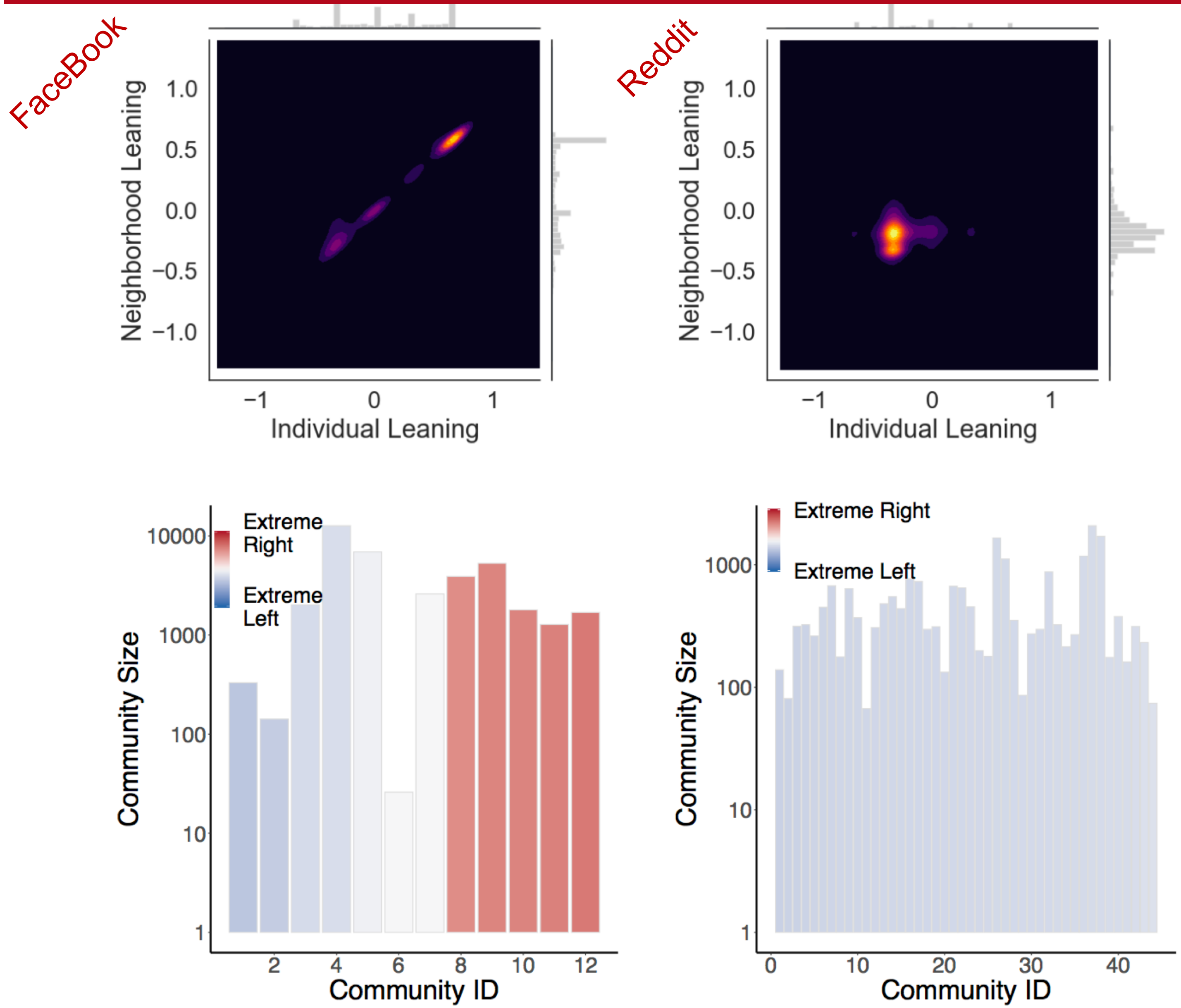
Filter bubble

From Wikipedia, the free encyclopedia

A **filter bubble** – a term coined by internet activist Eli Pariser – is a state of intellectual isolation[1] that allegedly can result from personalized searches when a website algorithm selectively guesses what information a user would like to see based on information about the user, such as location, past click-behavior and search history.[2][3][4] As a result, users become separated from information that disagrees with their viewpoints, effectively isolating them in their own cultural or ideological bubbles.[5] The choices made by these algorithms are not transparent.[6]

The term was coined by internet activist Eli Pariser circa 2010

Filter bubbles
in social networks



- Same Topic: News
- Same leaning assigned to news sources
- Different platforms: Facebook has a strong social feeding algorithm, Reddit has not
- Different characteristics: Facebook shows segregation among groups with different leaning, Reddit has one group

Assortativity

i.e., degree homophily

Correlation between hubs

- In some networks, hubs frequently connect with other hubs

  e.g., celebrity dating, actor networks

- In other cases hubs avoid connections with other hubs

  e.g., methabolic graphs, food webs (predators tend to differentiate their diet)

Assortativity

- Assortative network: high degree nodes connect with each other avoiding low degree nodes (tend to cliques)
- Disassortative network: opposite trend, hubs tend to avoid each other
- Neutral network: one with random wiring, i.e., aside from the (marginal) degree distribution of nodes, there is no correlation

Assortativity

(dis)assortativity quantifies homophily in social networks, e.g., effects like:

- Rich people tend to be friends with each other
- People with the same education tend to hang out together

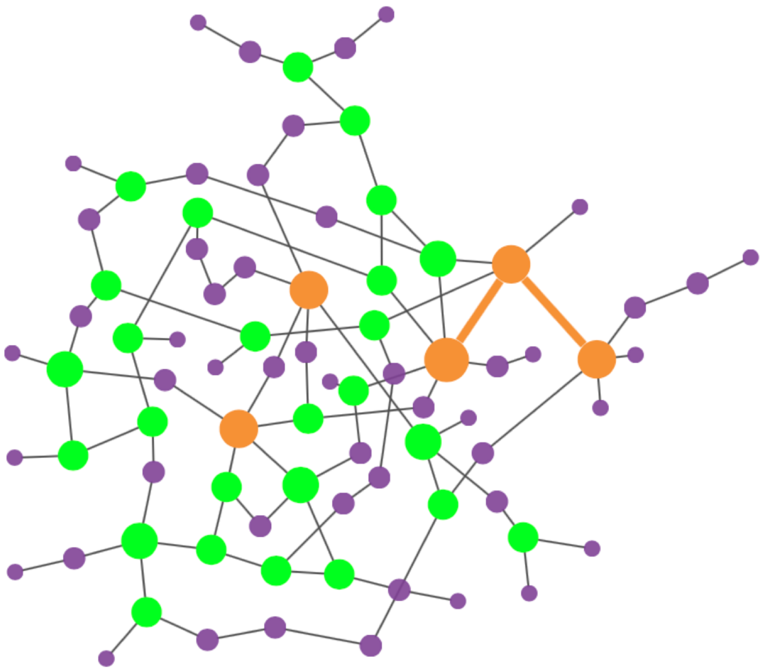
i.e., we expect social networks to be assortative

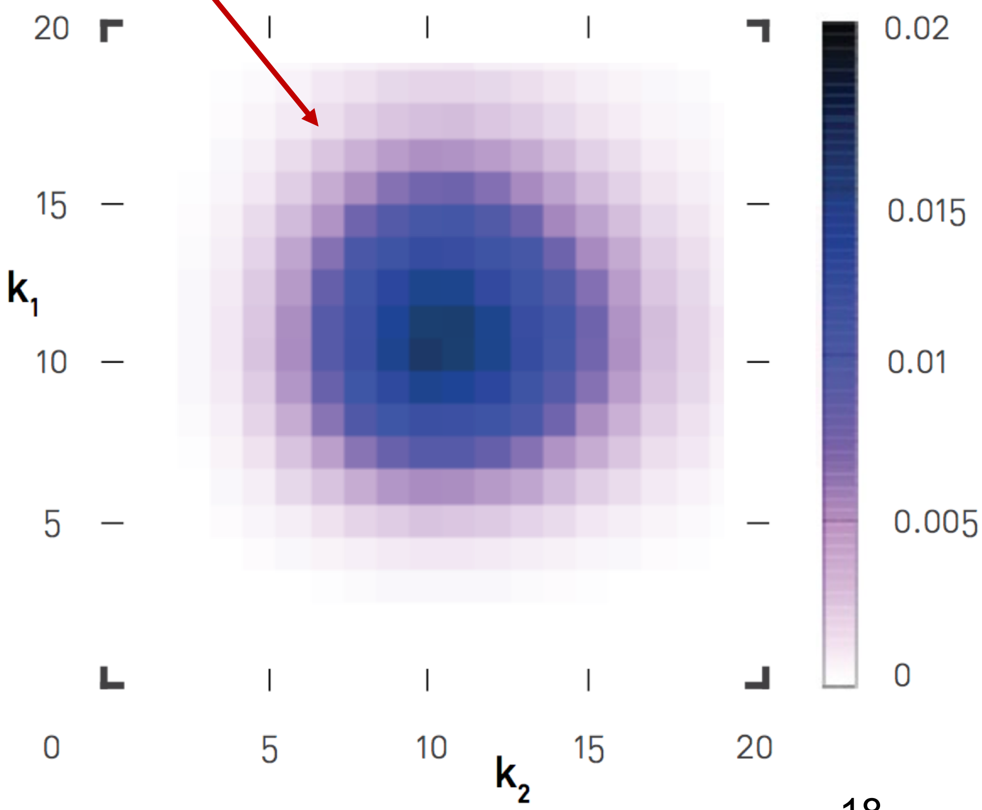
Neutral networks

The degree correlation matrix $E_{k1,k2}$ is visually centred around the average degree

In the neutral case we expect
$E_{k1,k2} = q_{k1}\, q_{k2}$, i.e., independence

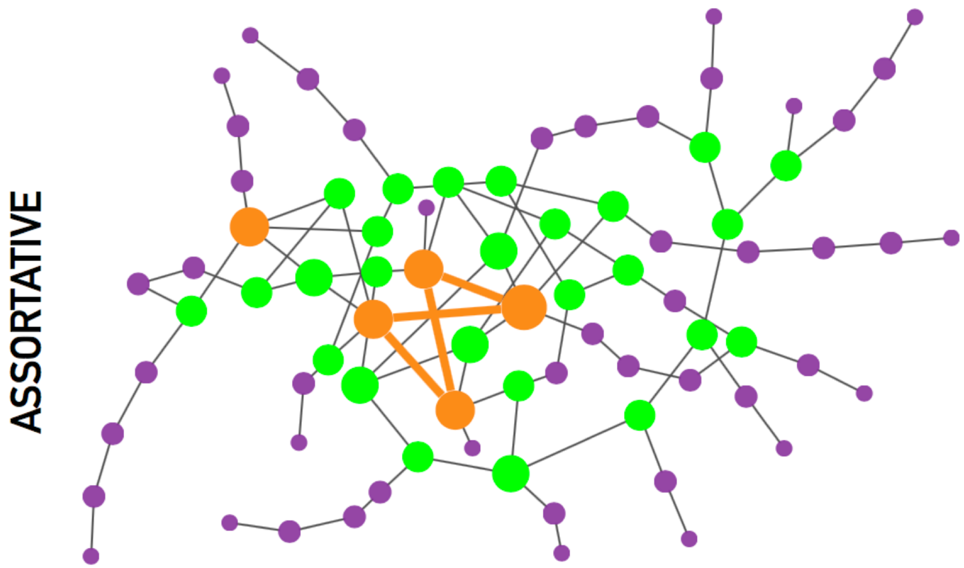
NEUTRAL

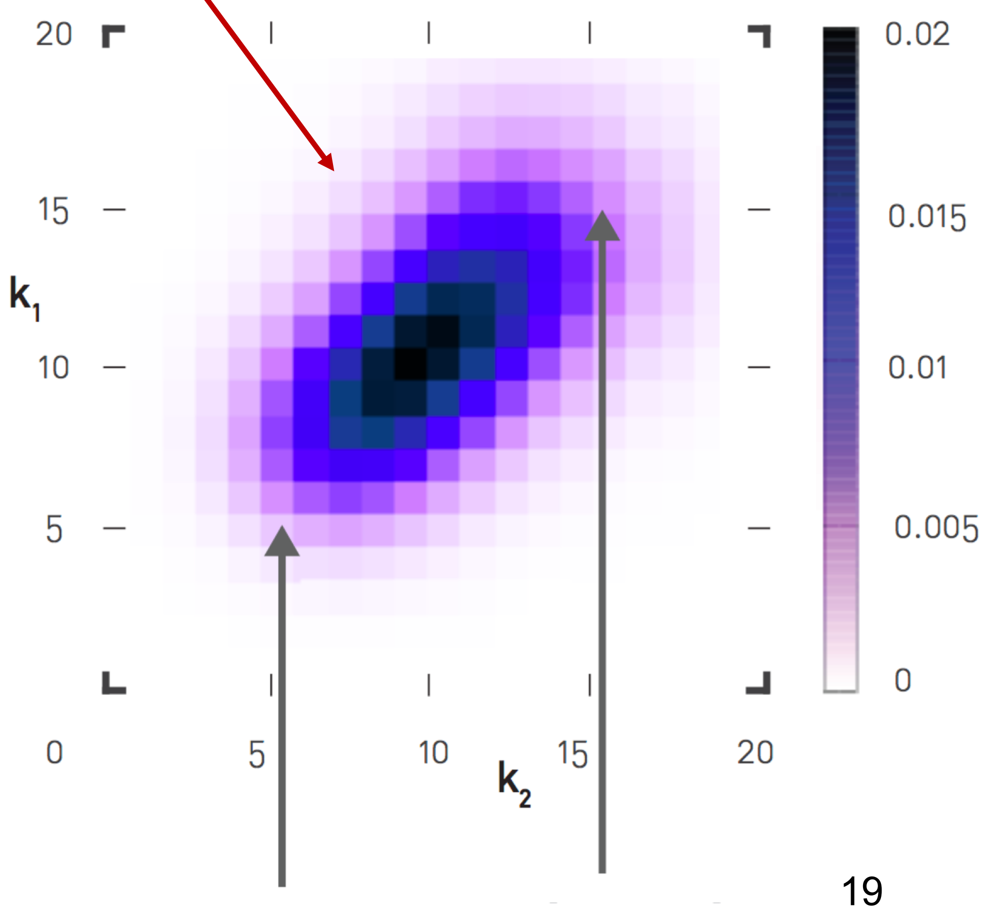




Assortative networks

The degree correlation matrix $E_{k1,k2}$ is turning to the right

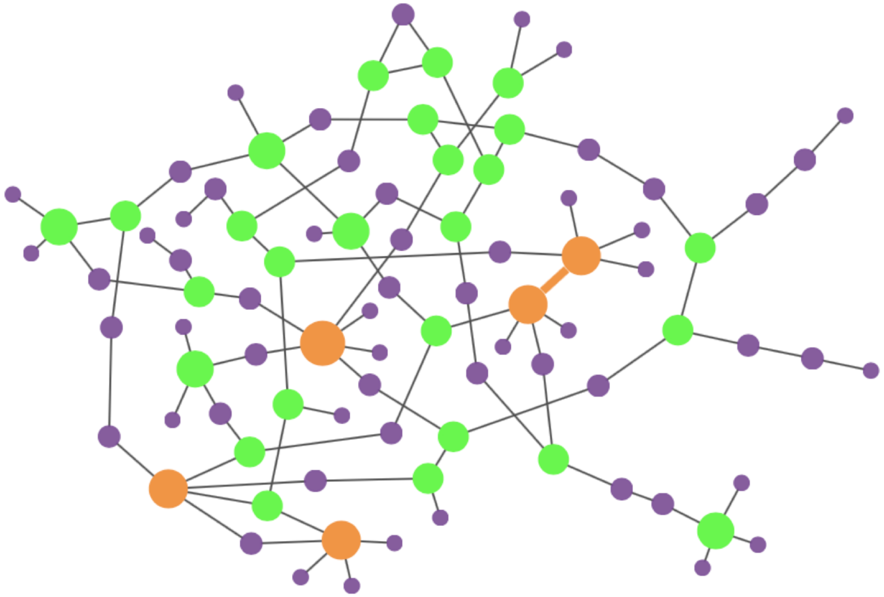


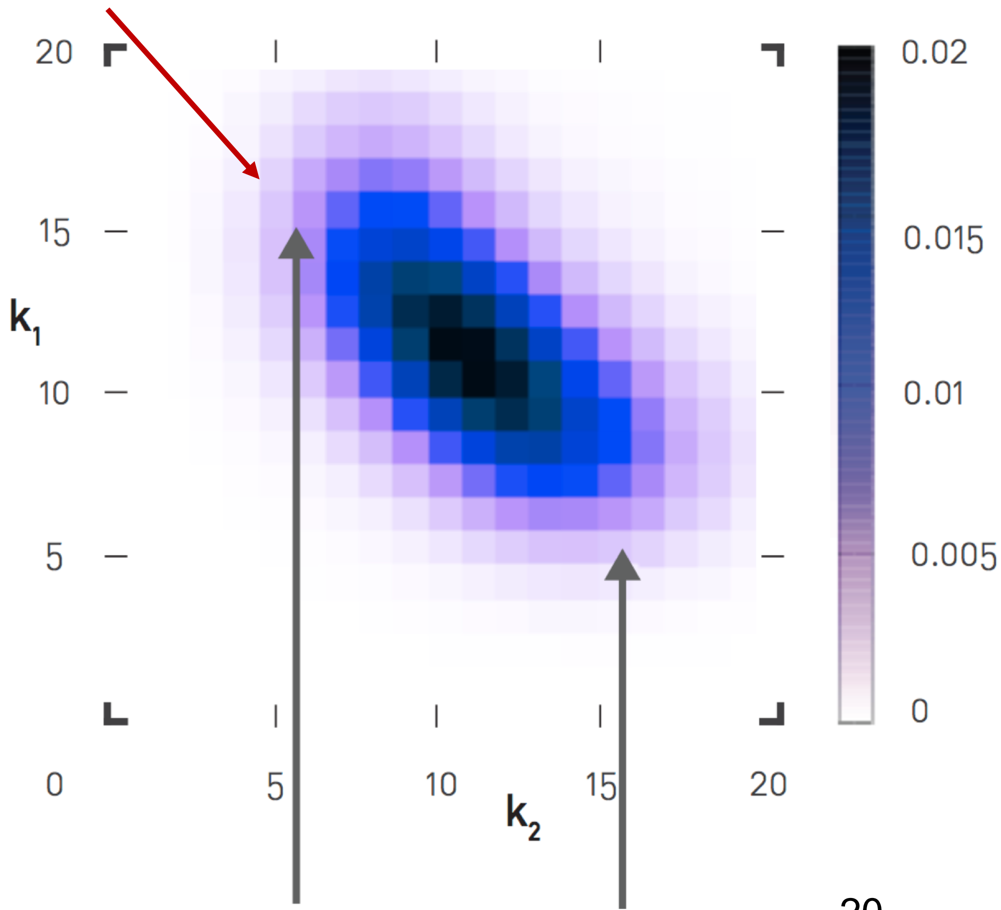


Disassortative networks

The degree correlation matrix
$E_{k1,k2}$ is turning to the left

DISASSORTATIVE

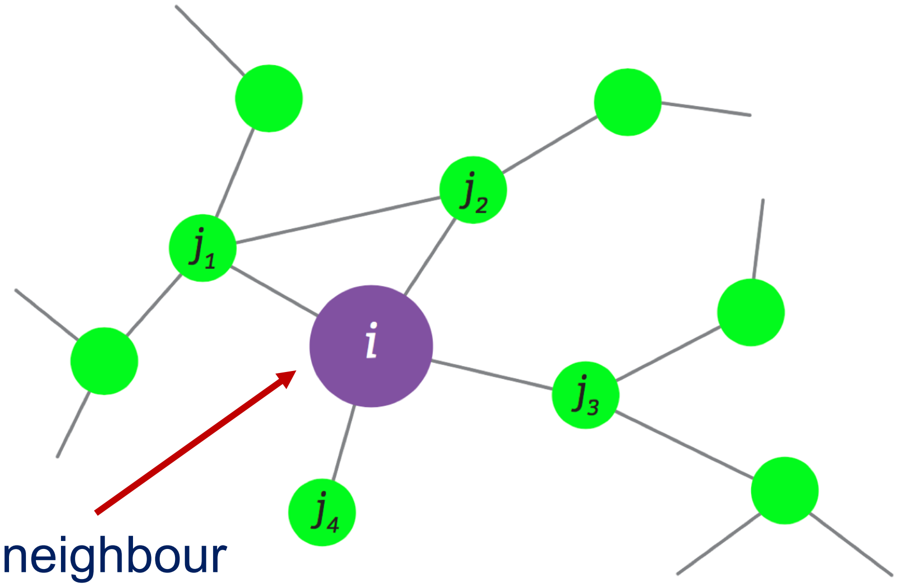




Nearest neighbour degree

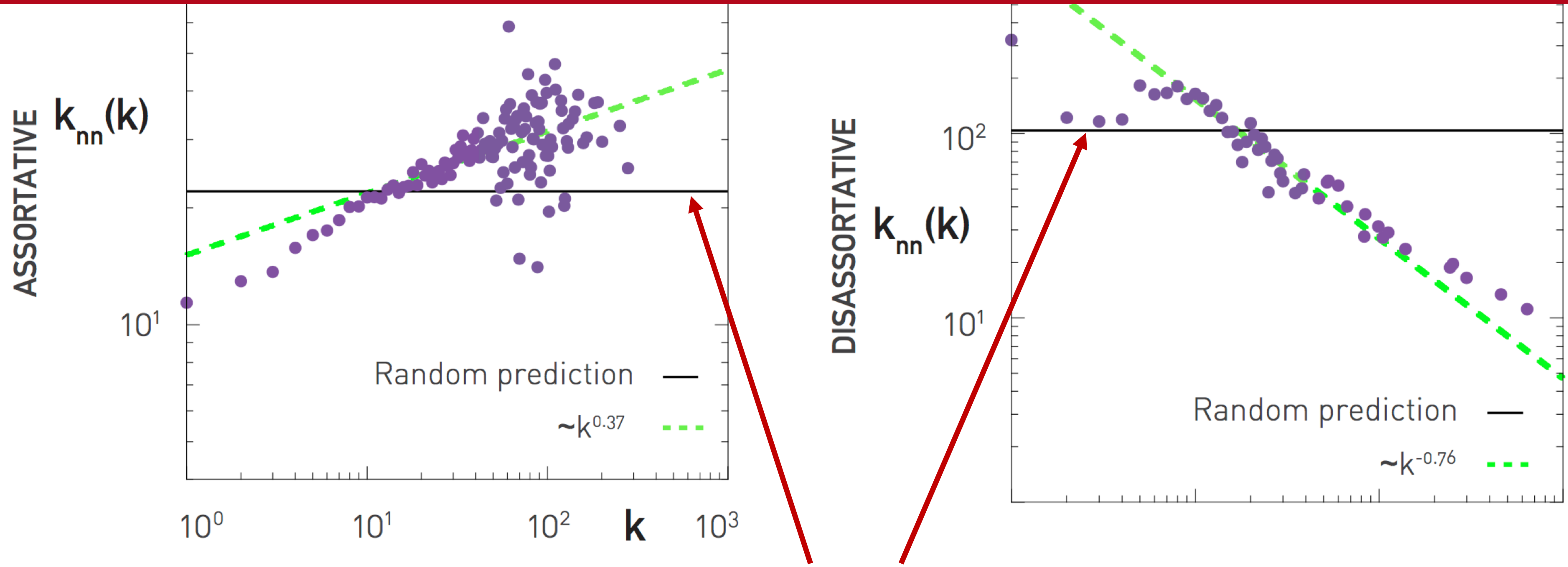
how to simplify plots from 2D to 1D

- Idea : inspect the degrees of the neighbouring nodes (easier than matrices)



average neighbour degree of node *i* is
¼ (4 + 3 + 1 + 3) = 2.75

Nearest neighbour degree plots



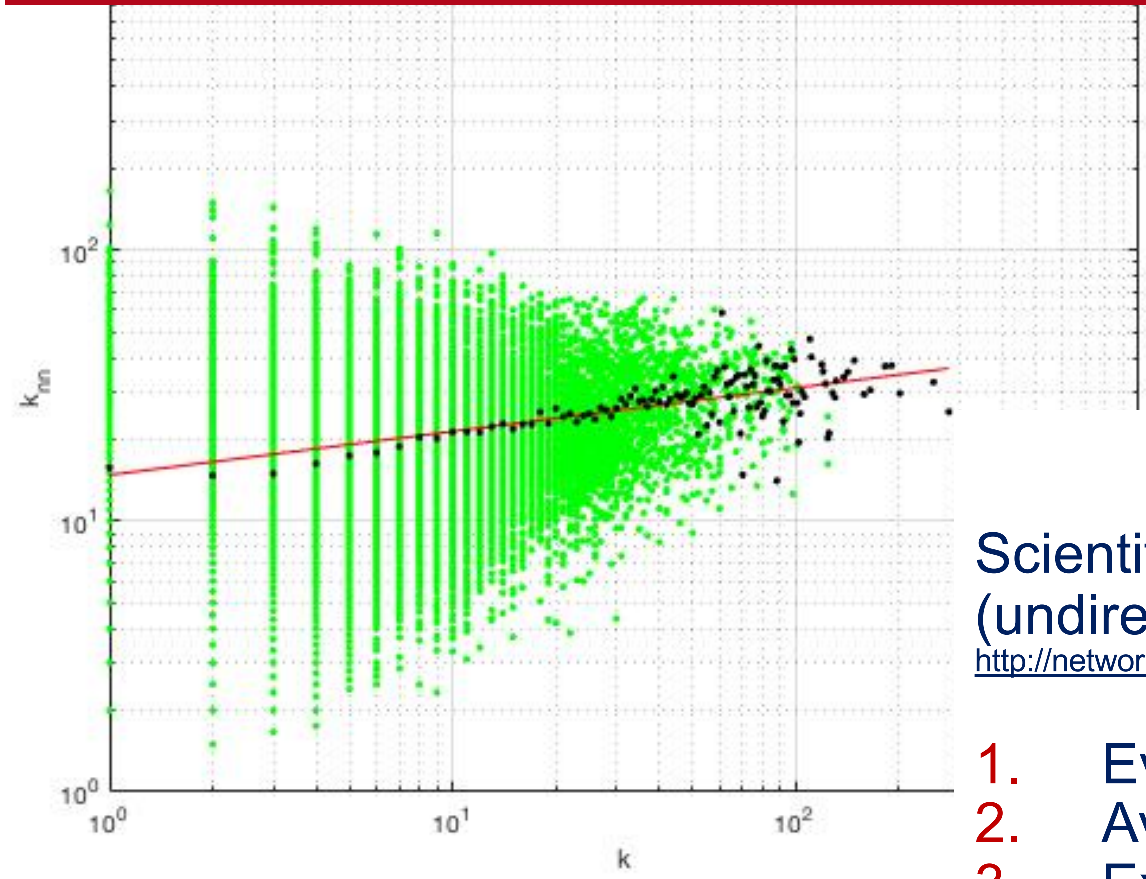
constant = independent of the degree (i.e., random = neutral)

$\ln(k_{nn}) = \mu \ln(k_i)$ → $\mu > 0$ = assortative

$\mu < 0$ = disassortative

A visual example

scientific collaboration network

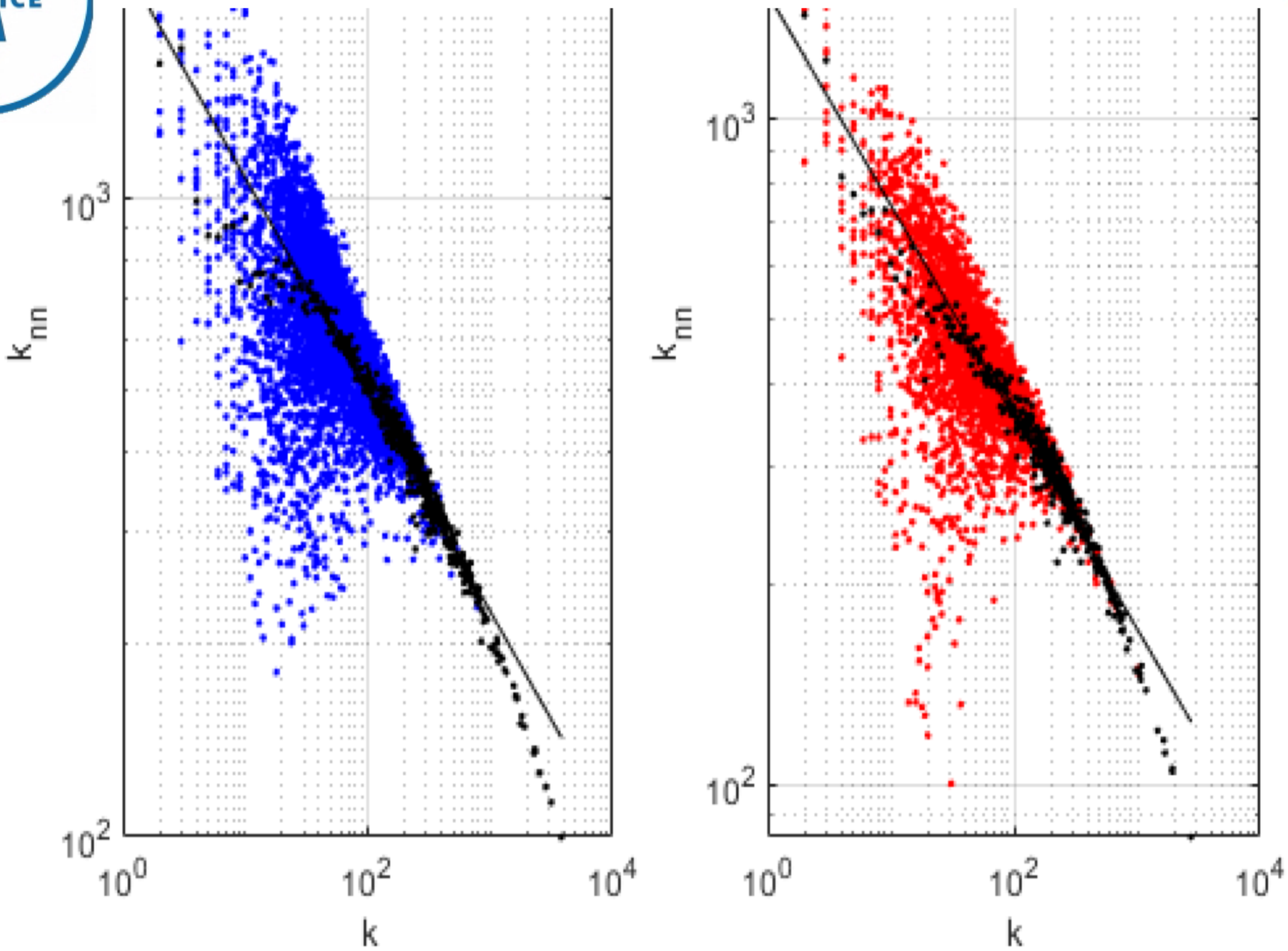


Scientific collaboration network (undirected, assortative)

http://networksciencebook.com/translations/en/resources/data.html

1. Evaluate average neigh. deg. k_{nn}
2. Average w.r.t. k
3. Extract the assortativity value μ=0.16

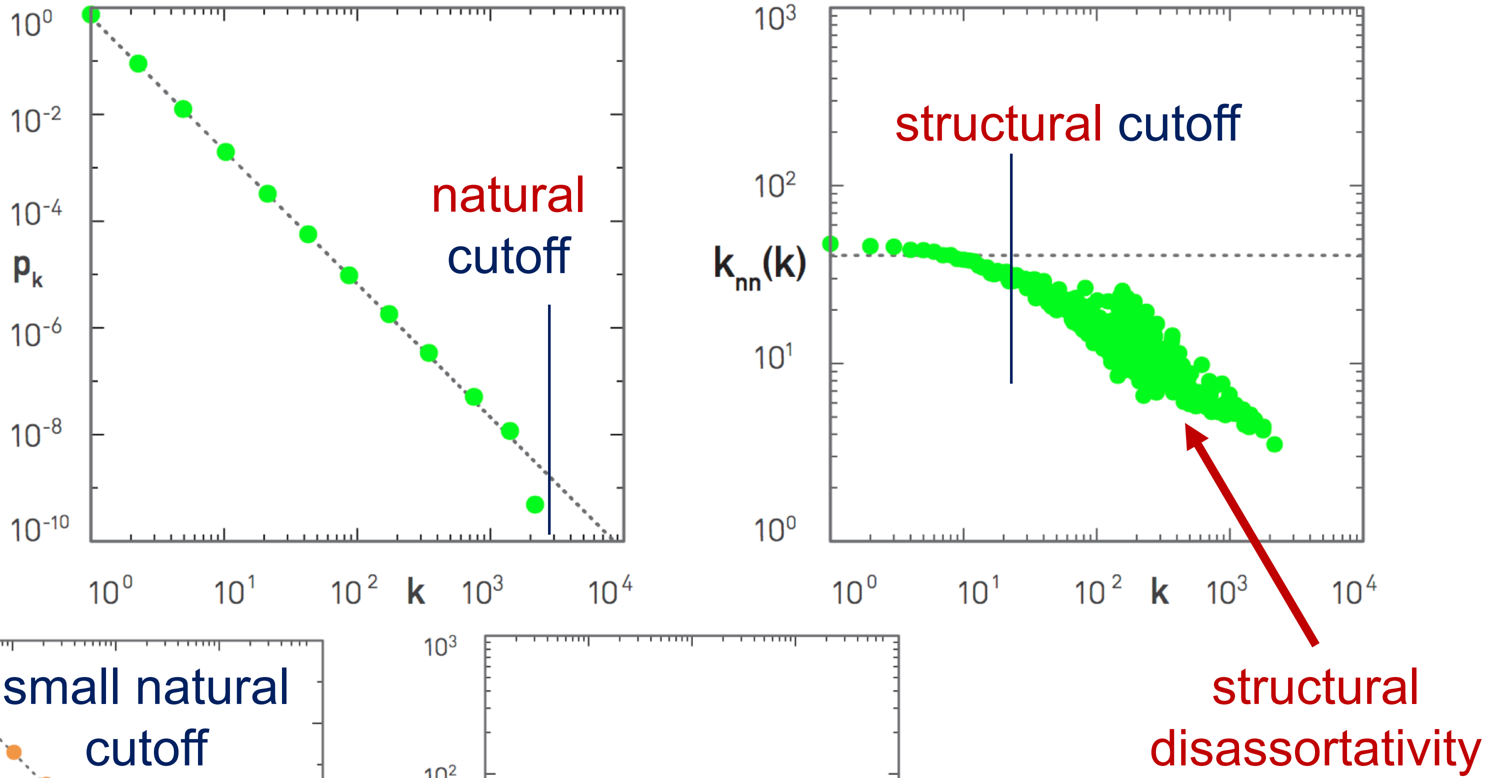
Hashtag network disassortativity
on pro-life/pro-choice data

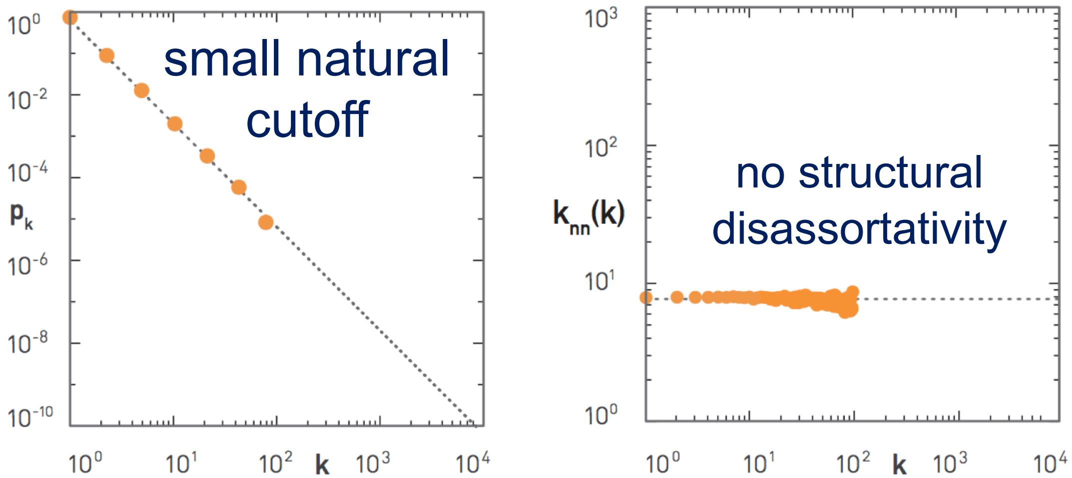


Structural disassortativity

large degrees cannot be supported by a neutral network

(dis)Assortativity can be linked to structural network properties





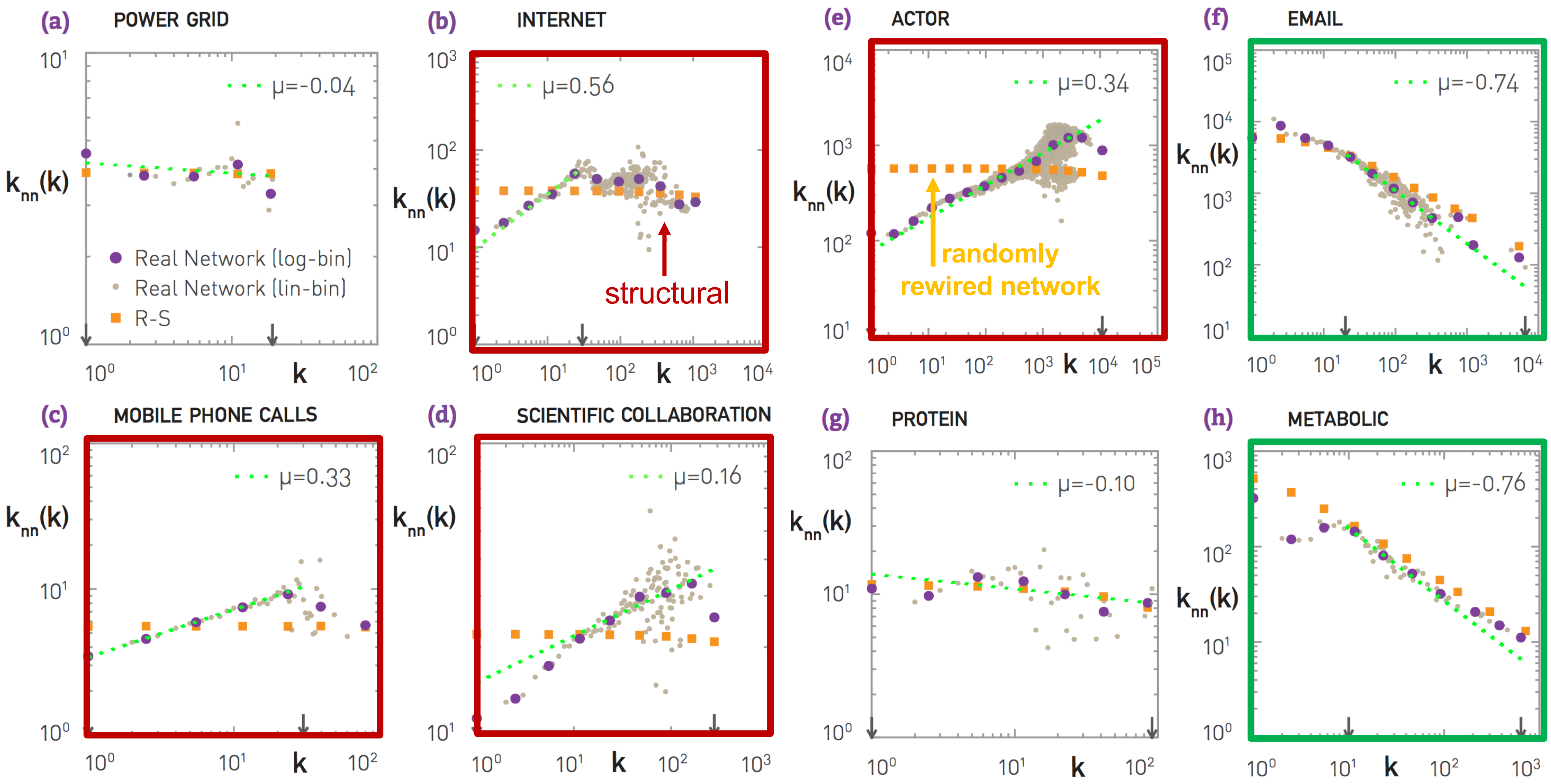
Università degli Studi di Padova

Structural disassortativity in real networks

social networks are assortative, most with a structural cutoff

assortative in red

disassortative in green

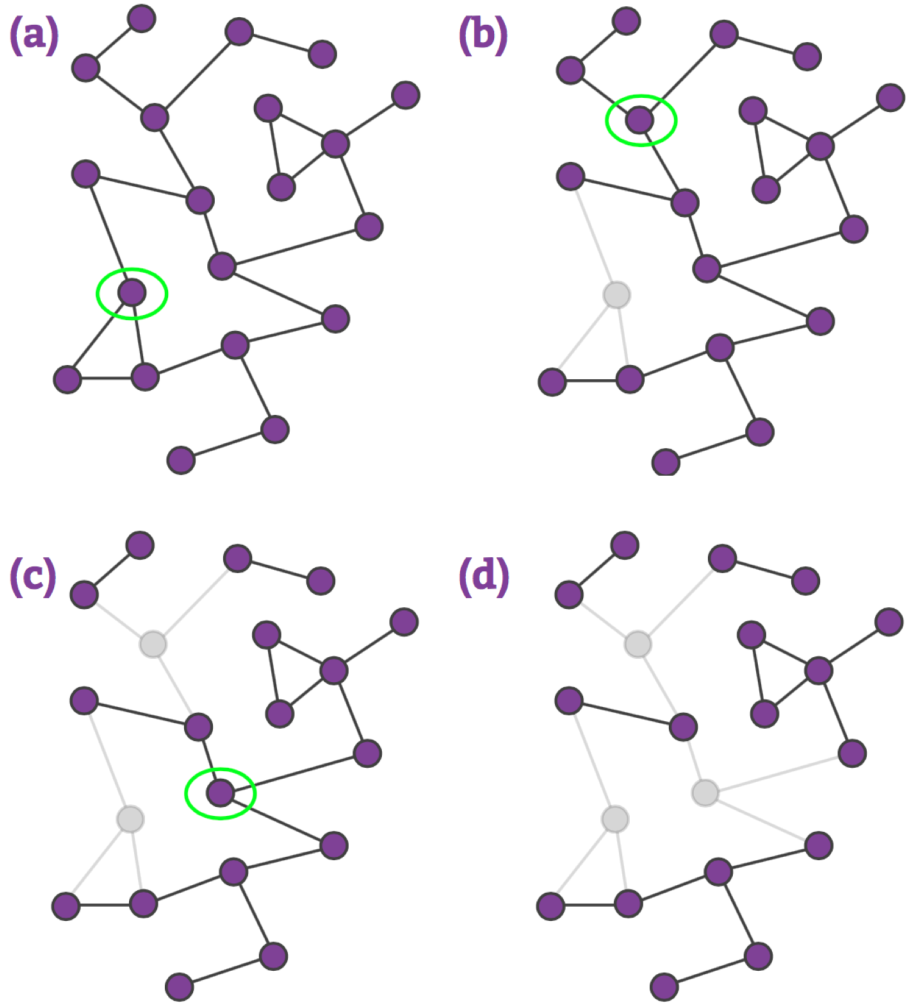


Robustness

of networks to failures

Network robustness

- Would the network still "work" in the presence of missing nodes?
- Failures can lead to either just isolating nodes or breaking the whole network apart
- What is the limit/phase transition?



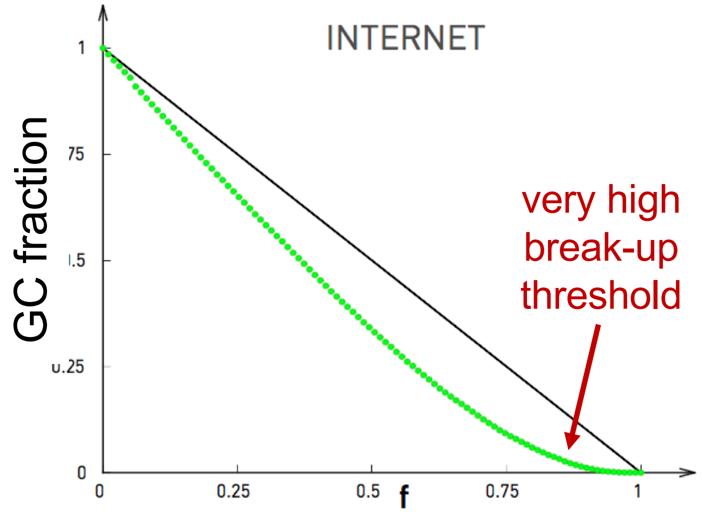
Applications

This can serve to identify:

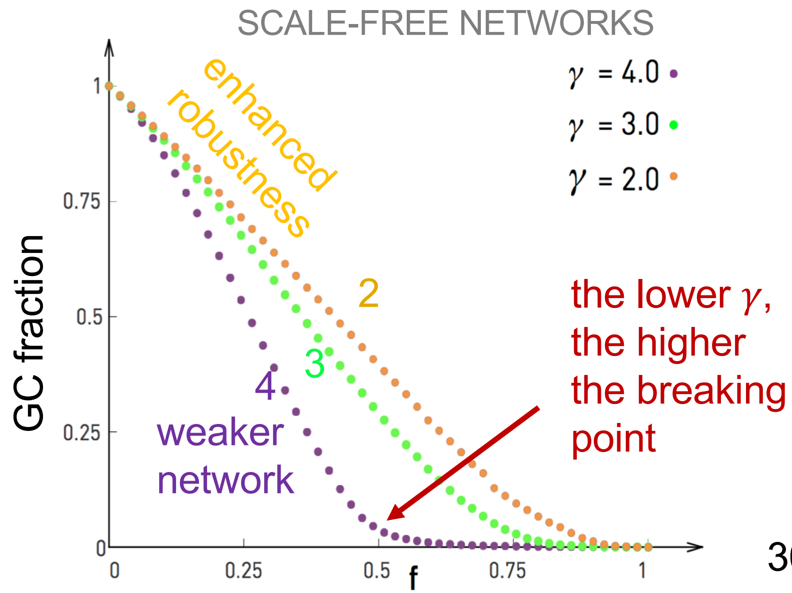
- robustness of air transportation under random strikes
- robustness of social contacts even when someone is off
- possibility of destroying of criminal/terror networks
- eradication of an epidemics
- etc.

Robustness of scale-free networks
under random node removal

- Robustness of the Internet due to scale-free properties
- Nodes linked to the GC after random removal with rate *f* → still large if *f<1*



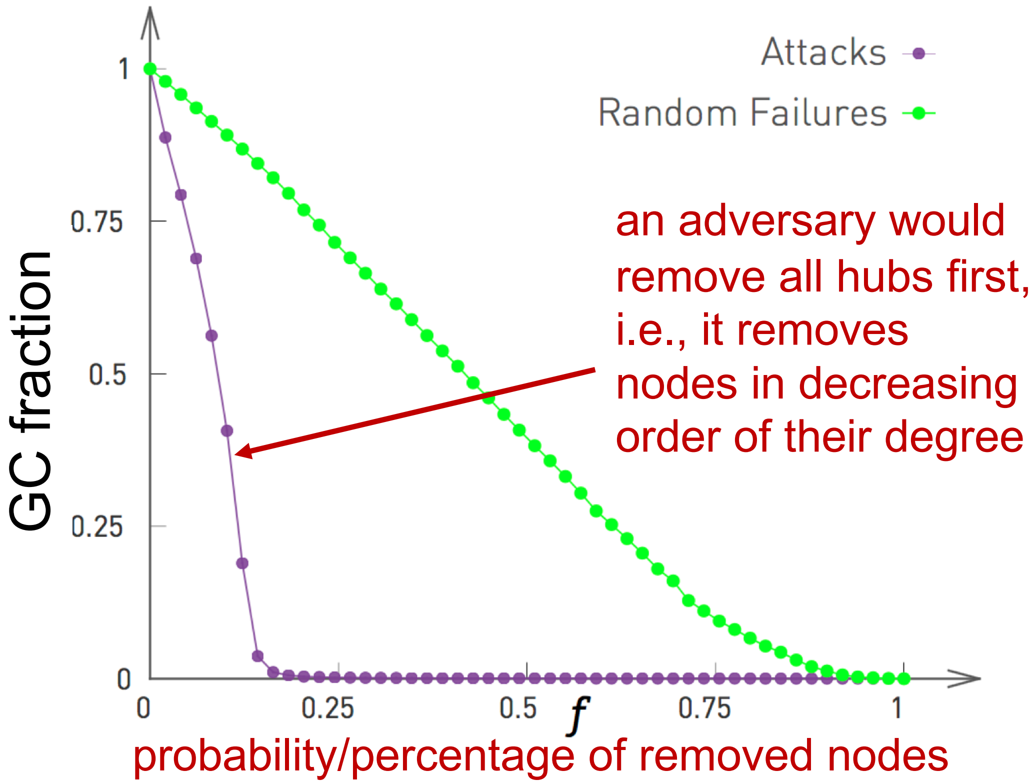
- Experiments aligned with a scale-free model
- Reason: random removal of (many) hubs is very unlikely



Attack tolerance

What if removals are not by chance, but caused by an adversary with sufficient insights on our network?

- Scale-free networks are not very robust to targeted attacks exactly because they have vulnerable hubs
- good news in medicine (vulnerability of bacteria) ☺
- bad news for the Internet ☹

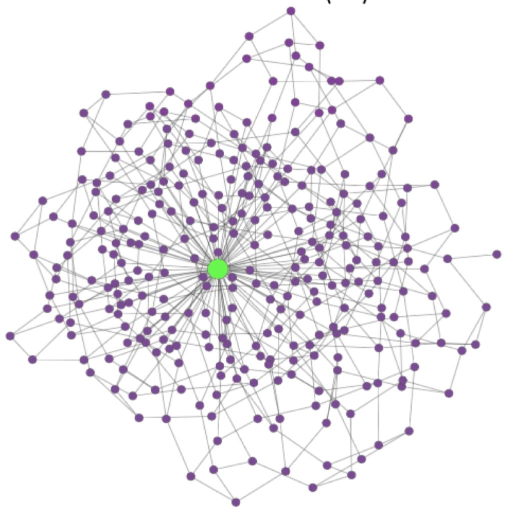


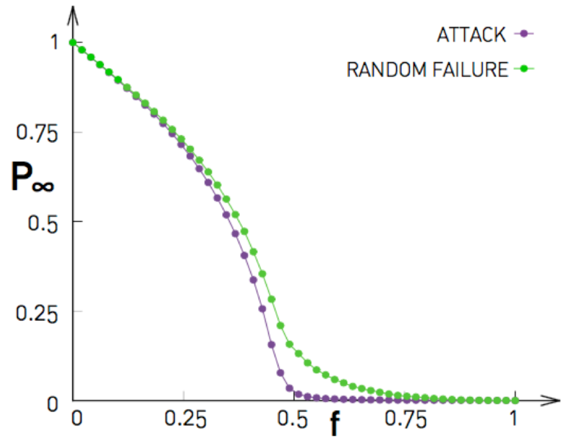
Optimizing robustness
is not an option in real-world networks

The best option is a bimodal distribution

$$p_k = r\,\delta_{k\max} + (1-r)\,\delta_{k\min}$$

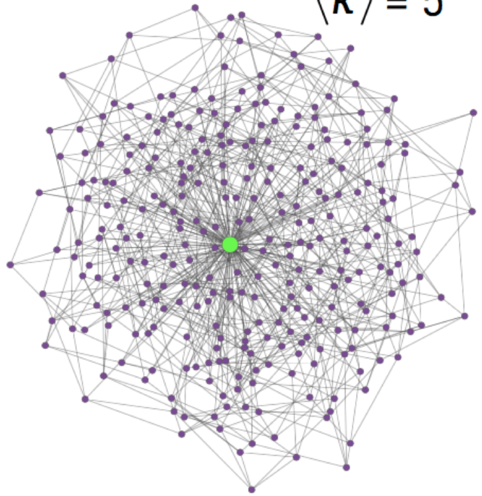
$\langle k \rangle = 3$

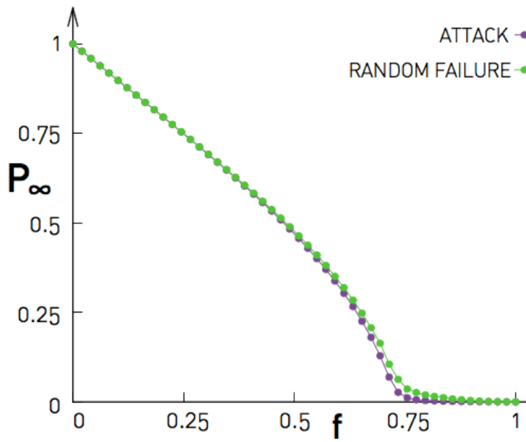




r = 1/N
k_{max} chosen to maximize the breakpoints

$\langle k \rangle = 5$





Example

network analysis of Tweets' sentiment

Salvatore Romano, Alberto Zancanaro, Enrico Lanza, Carlo Facchin

robustness of original network to positive node removal