What can we learn about the homo sapiens from massive multiplayer online games?

Stefan Thurner

MEDICAL UNIVERSITY OF VIENNA
SANTA FE INSTITUTE
IIASA
COMPLEXITY SCIENCE HUB VIENNA

www.complex-systems.meduniwien.ac.at   www.santafe.edu

https://drive.google.com/file/d/1i9dpsx_Bc4fYuYwD-LO-4wmtZ2zjRz44/view?usp=sharing_eil&ts=5b505057

MEDICAL UNIVERSITY OF VIENNA
SANTA FE INSTITUTE
IIASA
COMPLEXITY SCIENCE HUB VIENNA

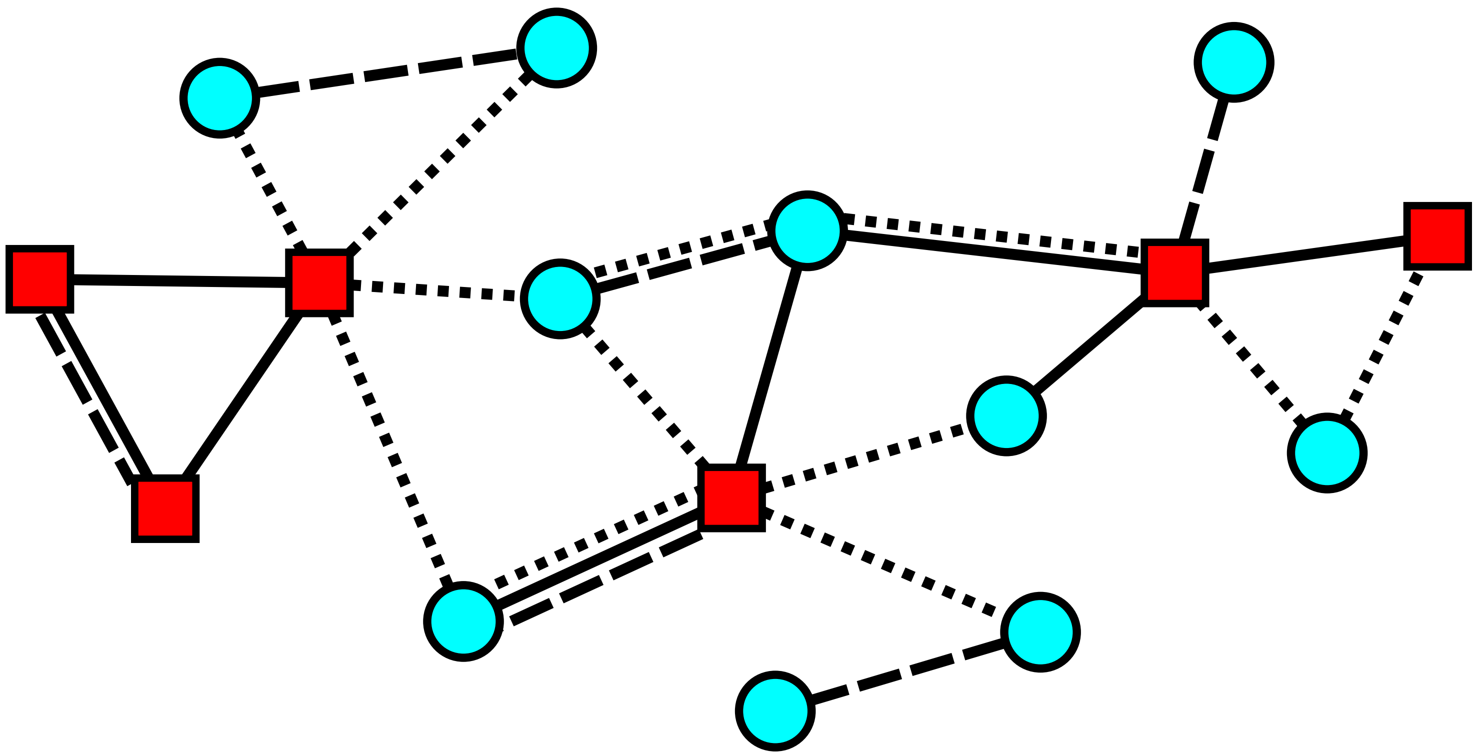
Let's do it again ...

... transform social science into an experimental science

quantitative, predictive, testable

What is a society?

Societies are co-evolving multiplex networks



- Multiplex network, $M_{ij}^{\alpha}(t)$
- Nodes i characterized by states, $\sigma_i^{\beta}(t)$

Co-evolving multiplex networks = complex system

Formally!

$$\frac{d}{dt}\sigma_i^{\alpha}(t) \sim F\left(M_{ij}^{\alpha}(t), \sigma_j^{\beta}(t)\right)$$

and

$$\frac{d}{dt}M_{ij}^{\alpha}(t) \sim G\left(M_{ij}^{\alpha}(t), \sigma_j^{\beta}(t)\right)$$

can not **solve** this – but can watch it (not analytic but algorithmic)

- States of individuals are observable (big data)
- Networks are observable (big data)

with Michael Szell, Peter Klimek, Benedikt Fuchs, Renaud Lambiotte, Vito Latora, Roberta Sinatra, Didier Sornette, Olesya Mryglod, Yurko Holovatch, Bernat Corominas-Murtra, Maximilian Sadilek

Social Networks **32** (2010) 313
PNAS **107** (2010) 13636
Advances of Complex Systems **15** (2012) 1250064
PLoS ONE **7** (2012) e29796
Scientific Reports **2** (2012) 457
Scientific Reports **3** (2013) 1214
New J of Physics (2013) 063008
PLoS ONE **9** (2014) e103503
Scientific Reports **4** (2014) 6526
PLoS ONE **9** (2014) e103503
Physica A **419** (2015) 681 PNAS **112** (2015) 5348

supported by EC FP7 projects LASAGNE and MULTIPLEX
Austrian Science Fund FWF

Our habitat: the PARDUS universe

www.pardus.at

- massive multiplayer browser game
- 480,000 registered, 16,000 active players
- online since Sep 2004
- free, optional 5$/month

What do avatars do?

"Economic life"

- Production, distribution and trade
- Make profit from economic activity
- Spend money on ships, equipment, drugs, . . .

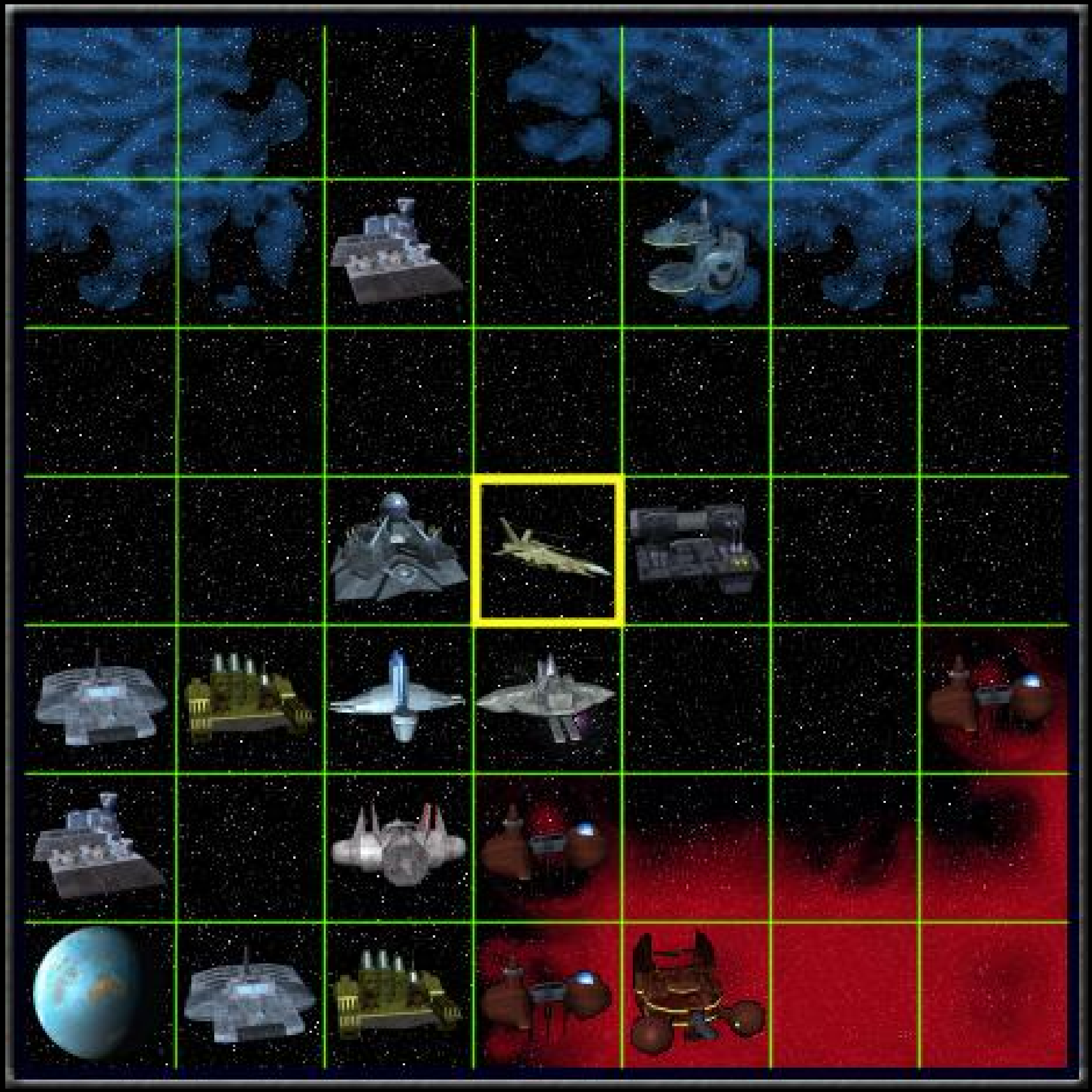
"Social life"

- Chats, forums, private messages
- Accumulate status

"Exploratory life" ("Science")

- Explore universe
- Fight space monsters...

The universe



Emergence of highly structured social behavior

- Hierarchical groups
- Political parties
- Organized attacks, e.g. wars
- Cartels, banks, courts
- “Science”

MEDICAL UNIVERSITY OF VIENNA

SANTA FE INSTITUTE

IIASA

COMPLEXITY SCIENCE HUB VIENNA

What do we know?

Everything !

- All actions of all (480.000) players at any point in time
- First time: **complete information** on an entire **human society**

MMO games – lab for socio-economic behavior

- high-frequency data from role-playing online games
- no observer-bias

all the information is there: node properties ($\sigma_i(t)$) and links ($M_{ij}^{\alpha}(t)$)

MEDICAL UNIVERSITY OF VIENNA
SANTA FE INSTITUTE
IIASA
COMPLEXITY SCIENCE HUB VIENNA

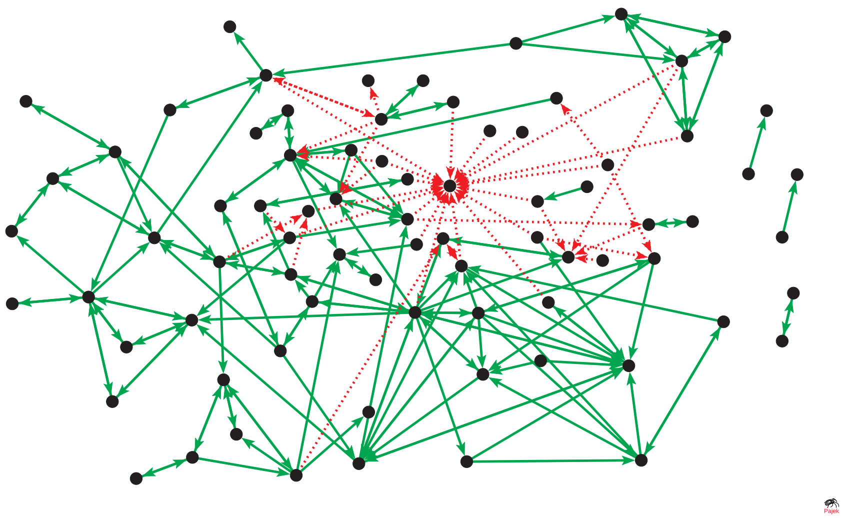
What to do if we know everything?

Experimental Soc Science

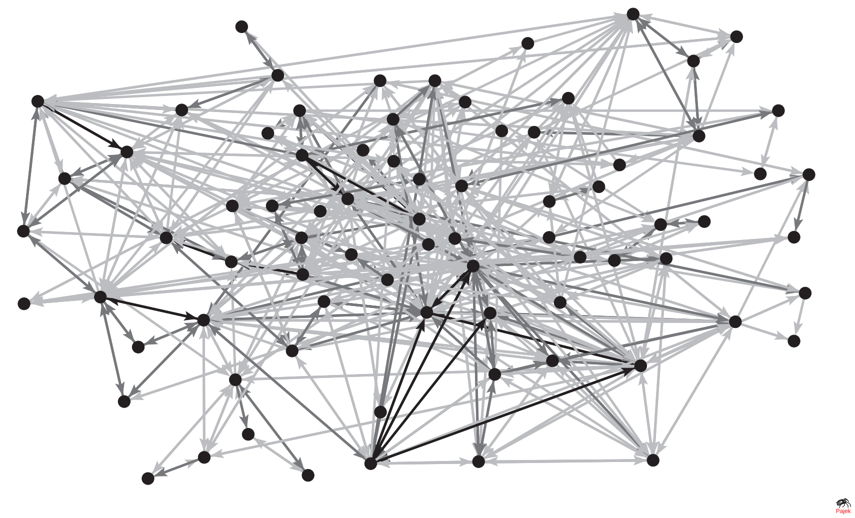
I How do people interact?

Measuring dynamics of social networks

Friend network, Enemy network

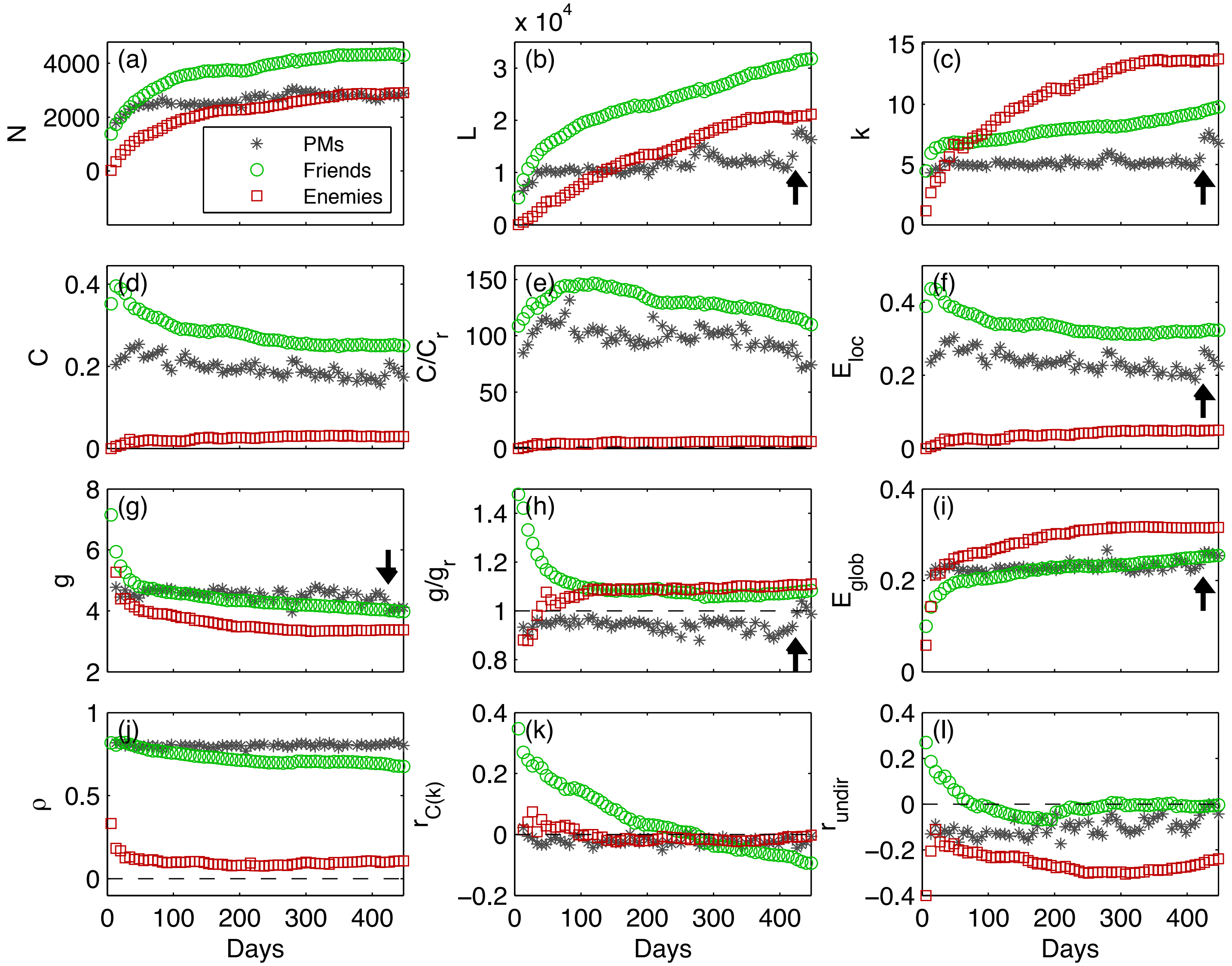


Communication network



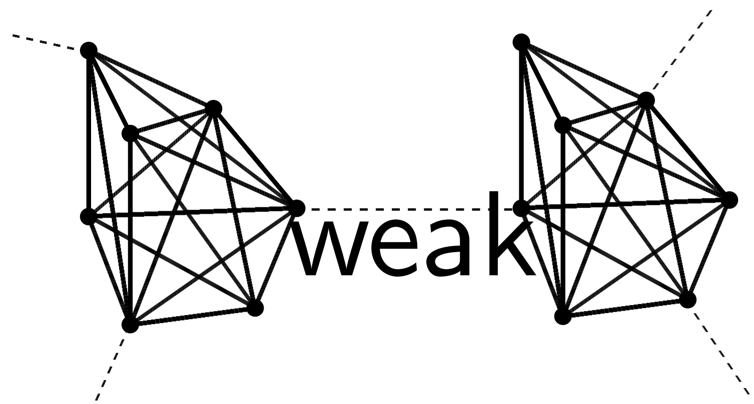
similar: trading, attack, production, transportation, bounty networks, ...

Evolution of social network properties

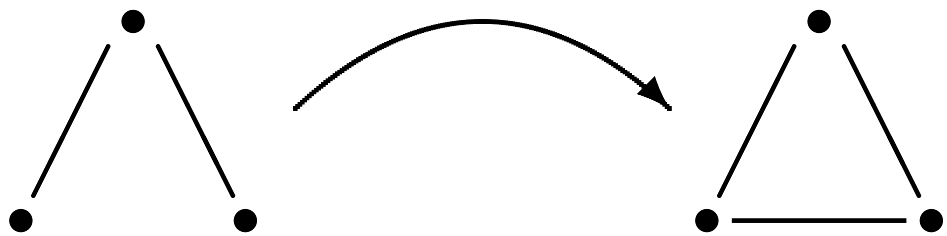


Testing classical sociological hypotheses

- **Weak ties hypothesis:** Communities are connected by weak ties



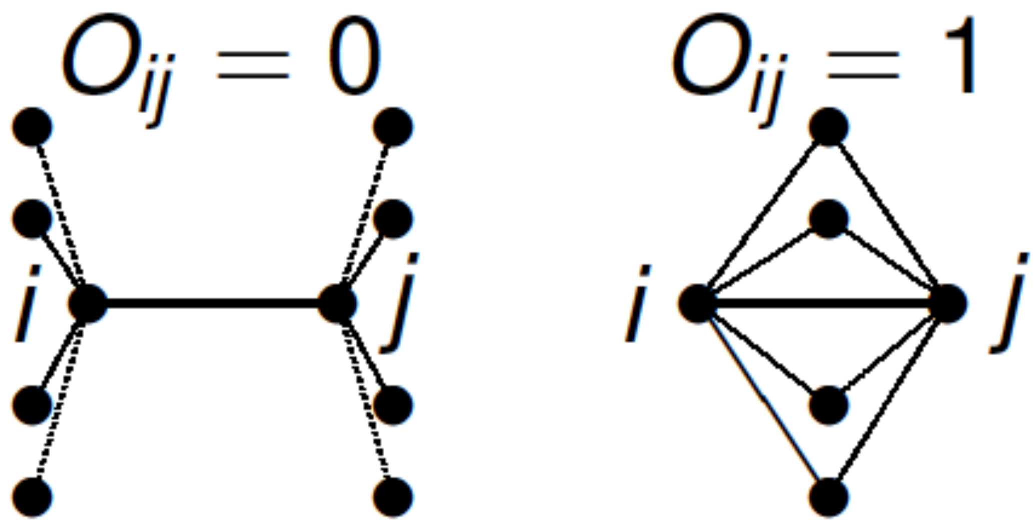
- **Triadic closure**: two friends of mine tend to become friends themselves



Weak ties hypothesis – experimentally justified?

(Granovetter 1973): *the degree of overlap of two individual's friendship networks varies* **directly** *with the strength of their tie to one another*

$$O_{ij} := \frac{n_{ij}}{(k_i - 1) + (k_j - 1) - n_{ij}}$$

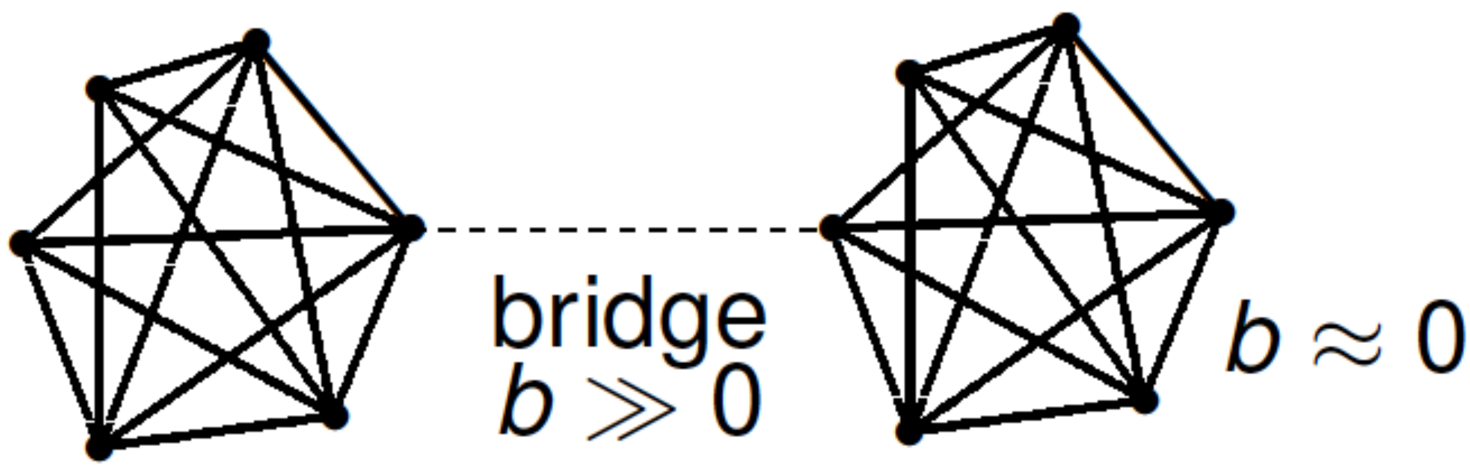


strength $\equiv$ number of exchanged messages w

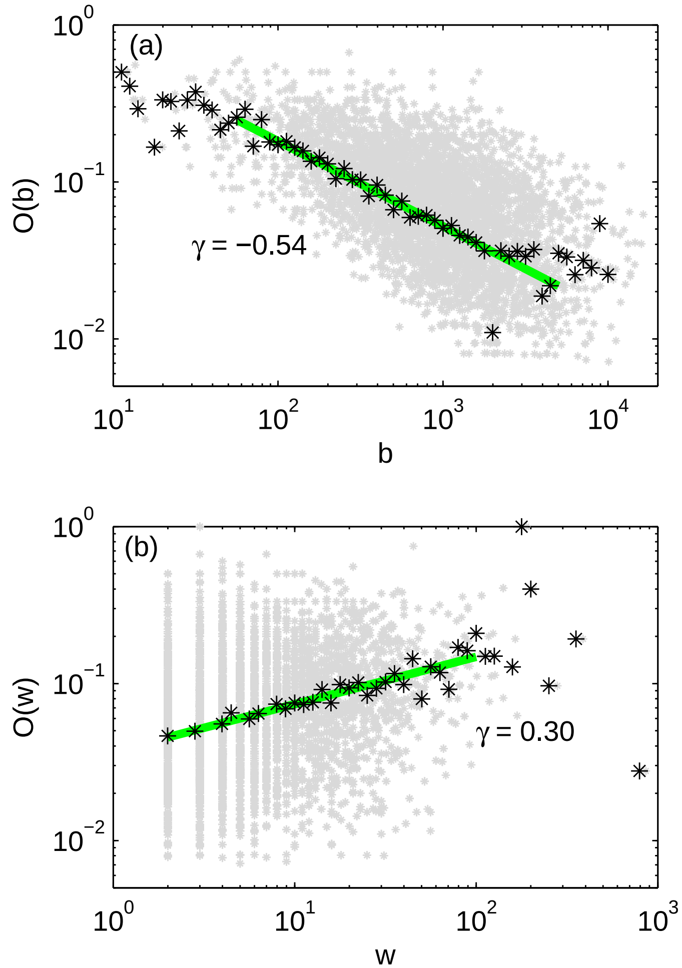
"No strong tie is a bridge" (Granovetter 1973)

measure for bridge: link-betweenness

$$b_{ij} := \sum_{m \in \mathcal{N}} \sum_{n \in \mathcal{N} \setminus \{m\}} \frac{\rho_{mn}(l_{ij})}{\rho_{mn}}$$



Weak ties hypothesis seems valid !



How strong do people interact?

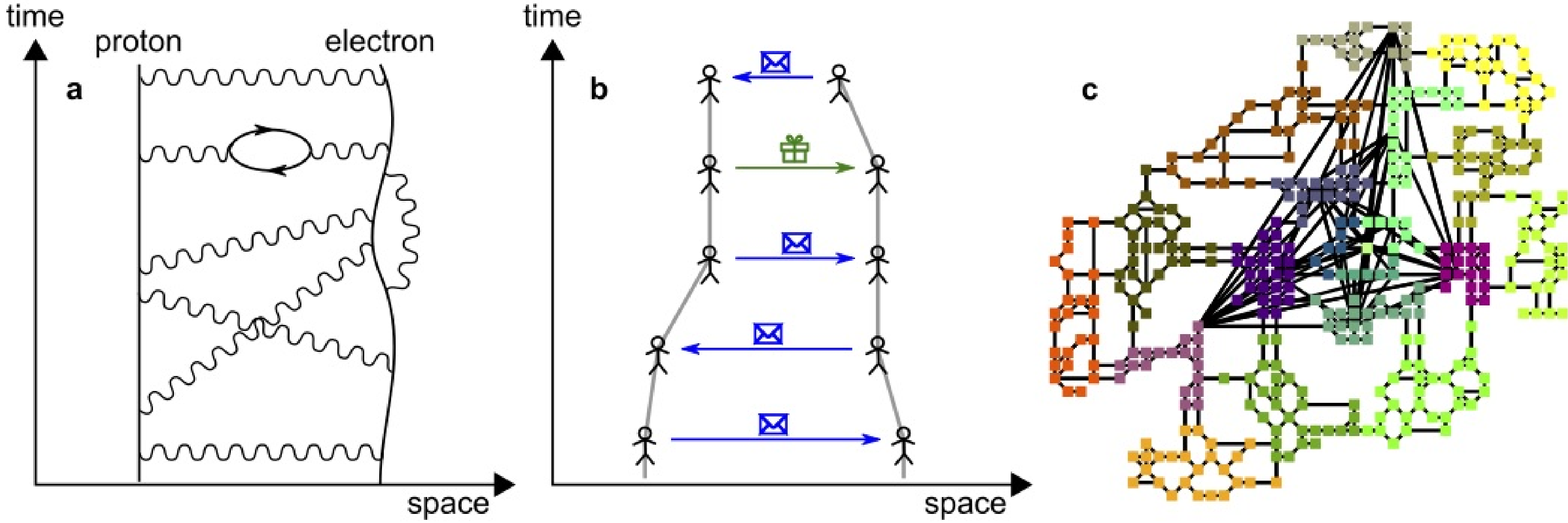
$$b = \frac{1}{O^2} \qquad \text{and} \qquad w = O^3$$

→ **Social "Kepler" law:**

$$w^2 = \left(\frac{1}{b}\right)^3$$

Strength of individual relation is determined by entire network

Physical forces between avatars

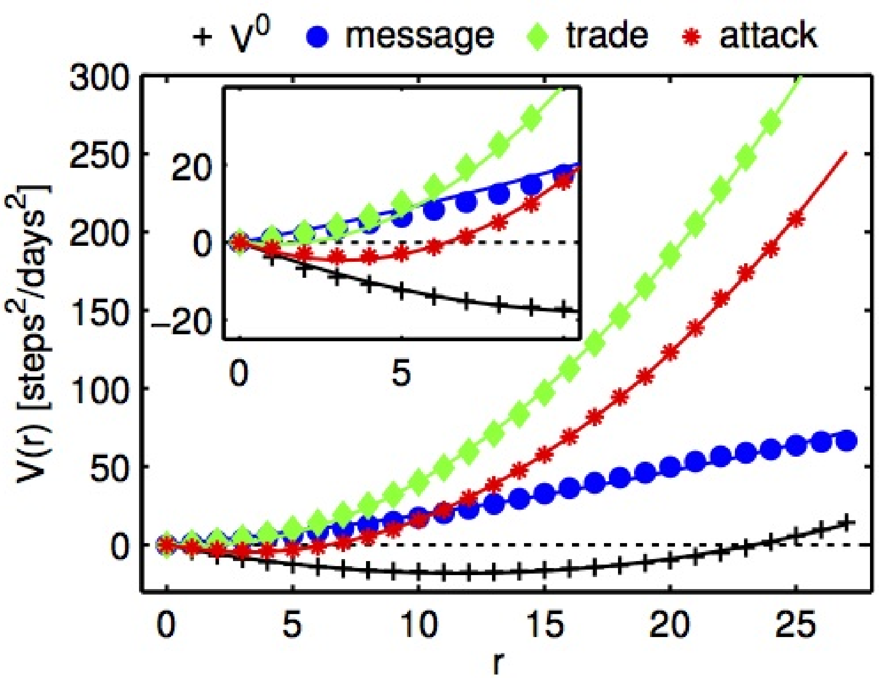


Forces in physics: exchange of gauge bosons

Humans: exchange of messages, objects, money, hostility, ...

Measure acceleration as a function of exchange

$$F = m\ddot{x} = -\frac{d}{dr}(V - V_0)$$

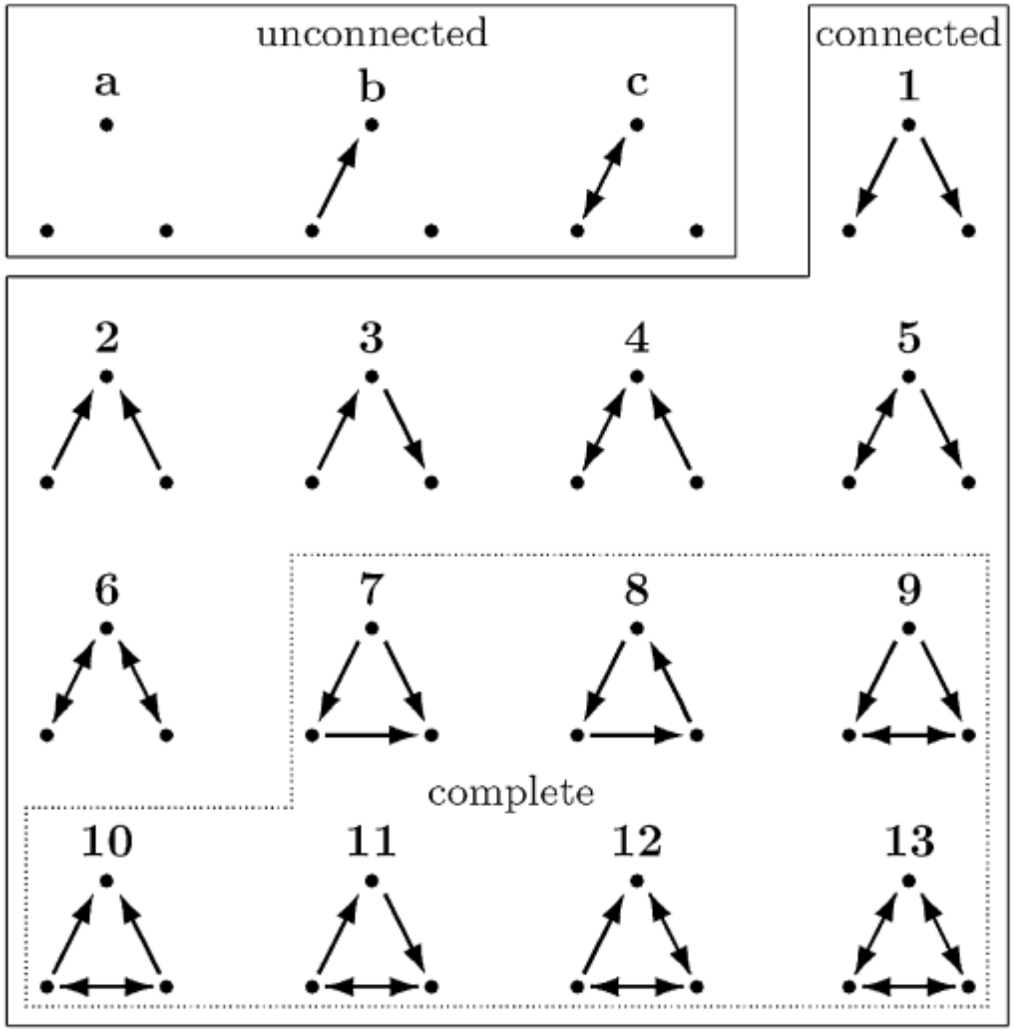


Attractive forces due to exchange of messages and trade

Repulsive and attractive force from hostile actions ('hit and run' force)

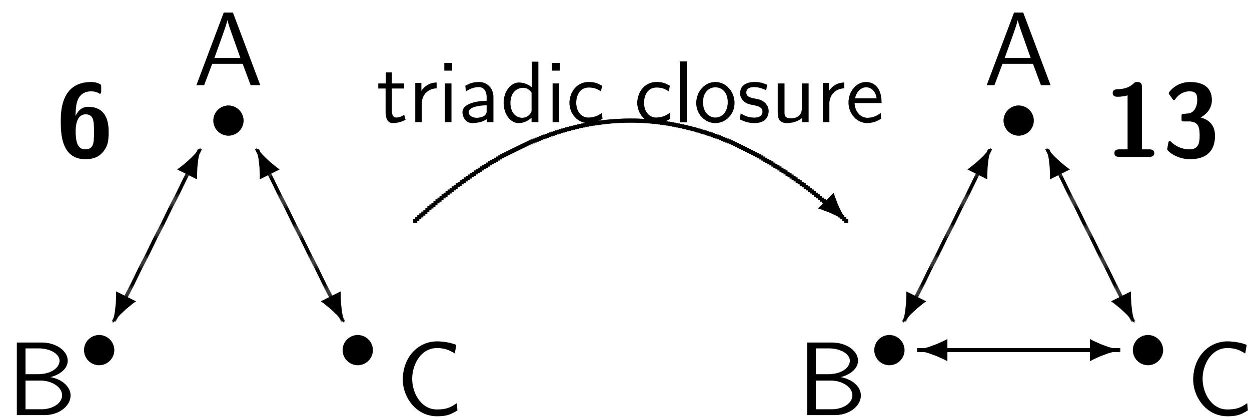
II How do people organize?

Triadic closure



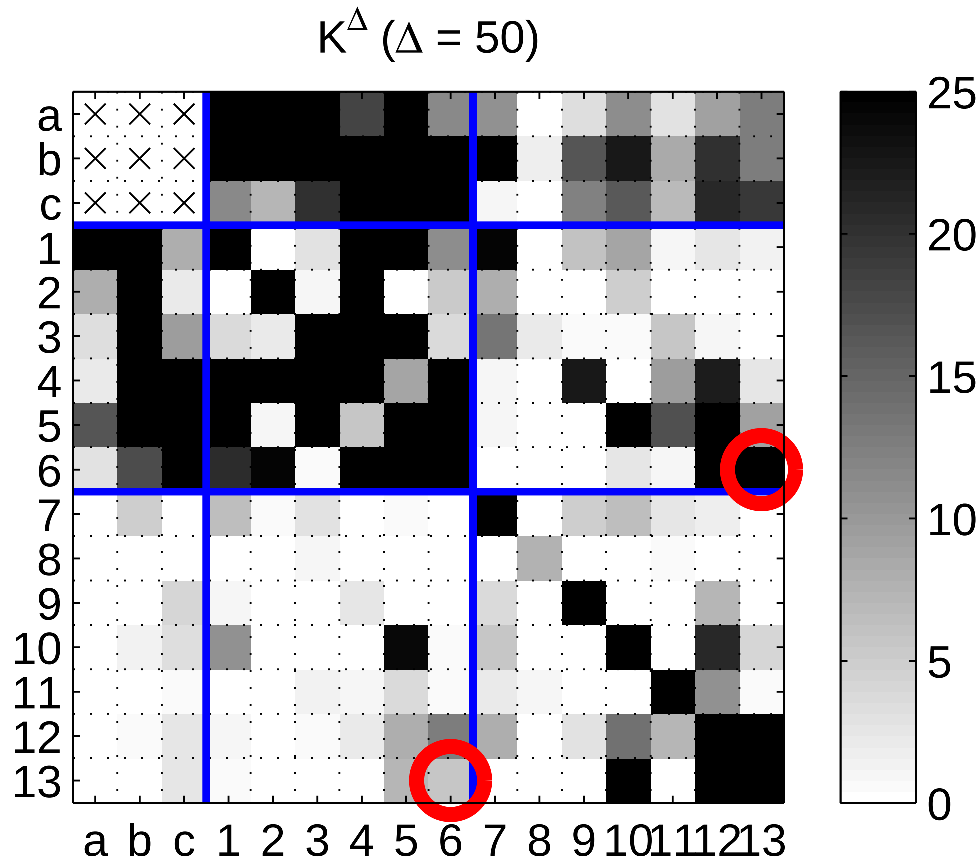
Triadic closure – dynamics of microscopic building blocks

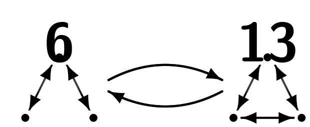
(Granovetter 1973): The triad which is most *unlikely* to occur, [. . .] is that in which A and B are strongly linked, A has a strong tie to some friend C, but the tie between C and B is absent



In proper English: **Expect over-representation of complete triads**

Transition probabilities between triads

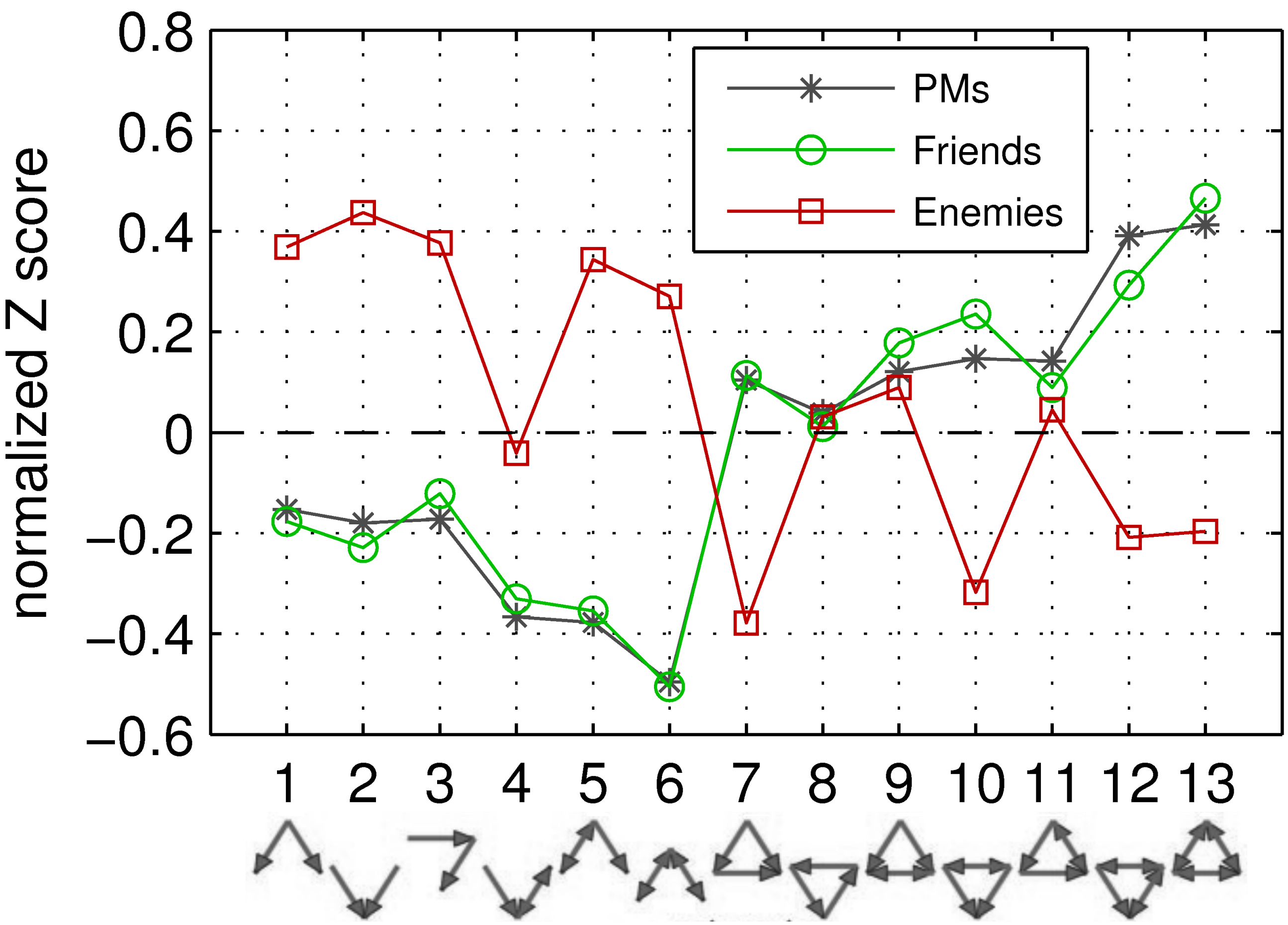




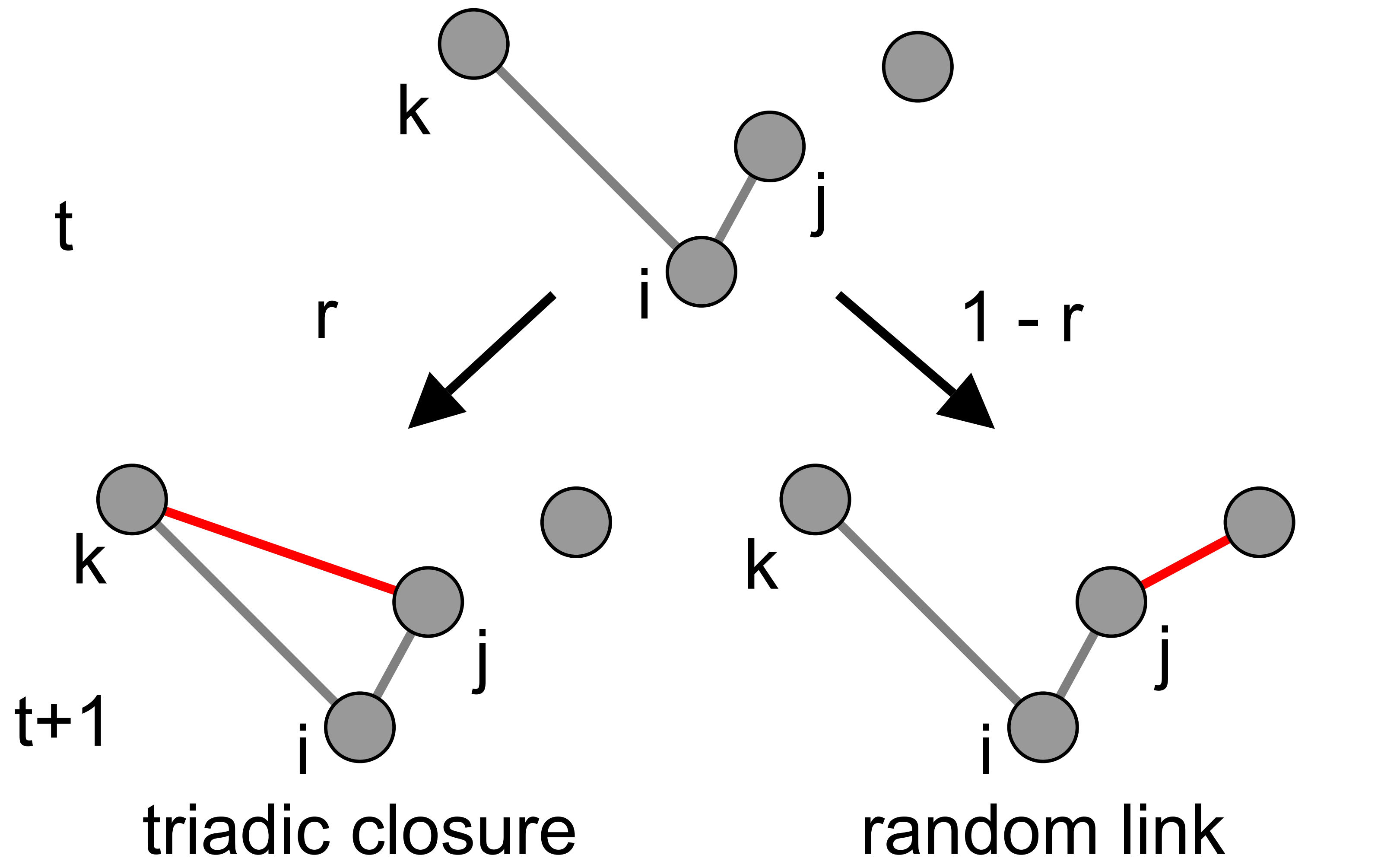
- Explicit quantitative evidence for triadic closure
- Dynamic modeling has to incorporate transition probabilities

Triad significance profile

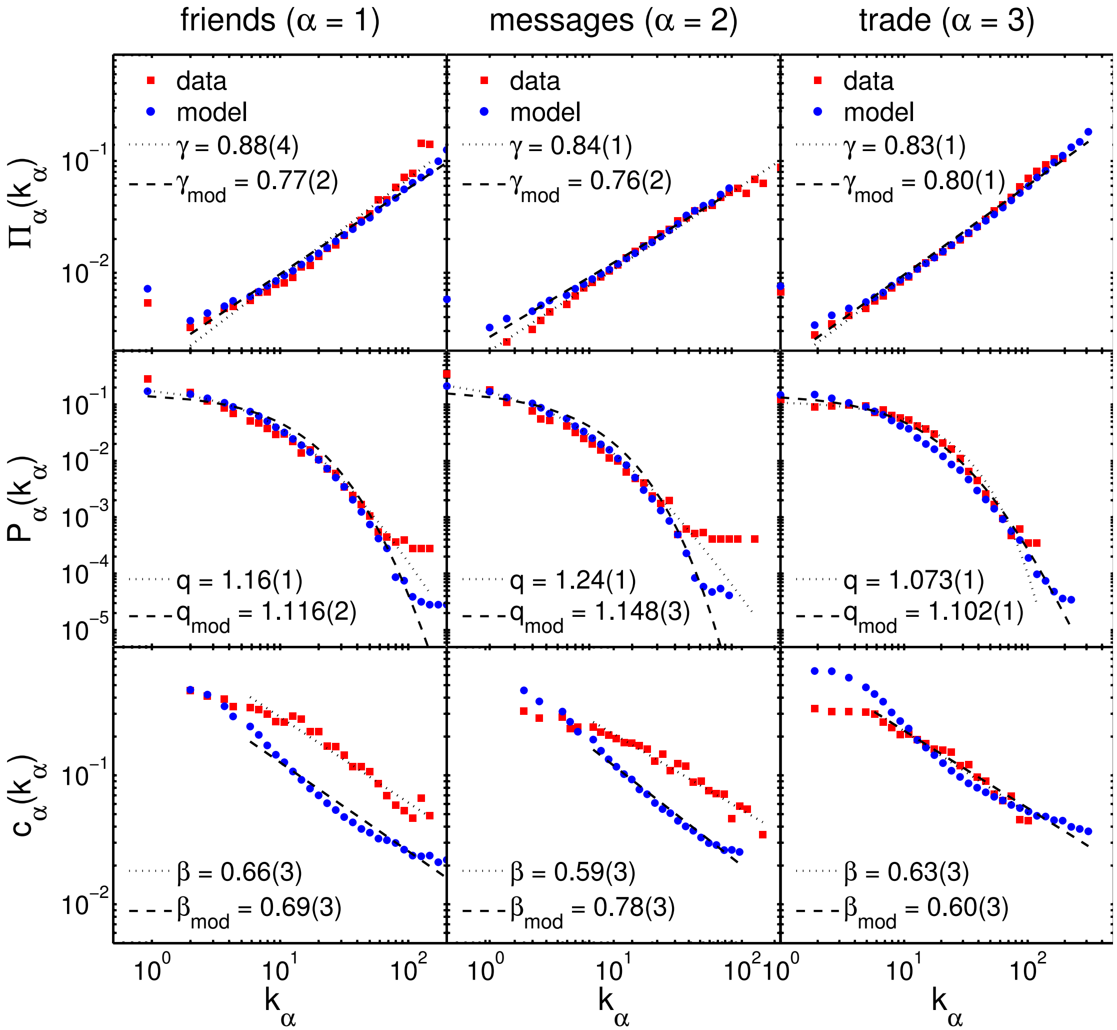
Z-score: measures over/under-representation of triads w.r.t. random graph



Taking triadic closure seriously

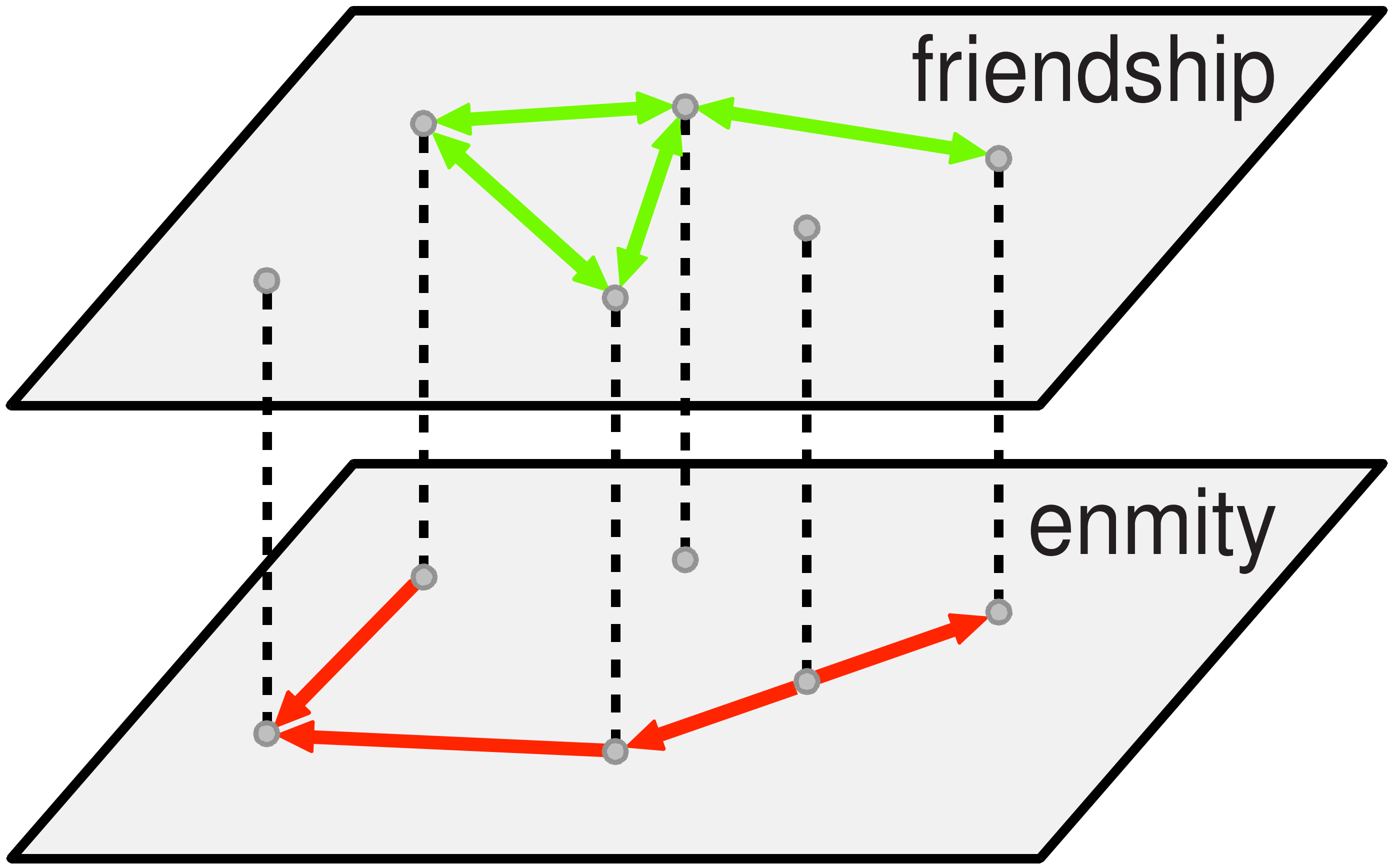


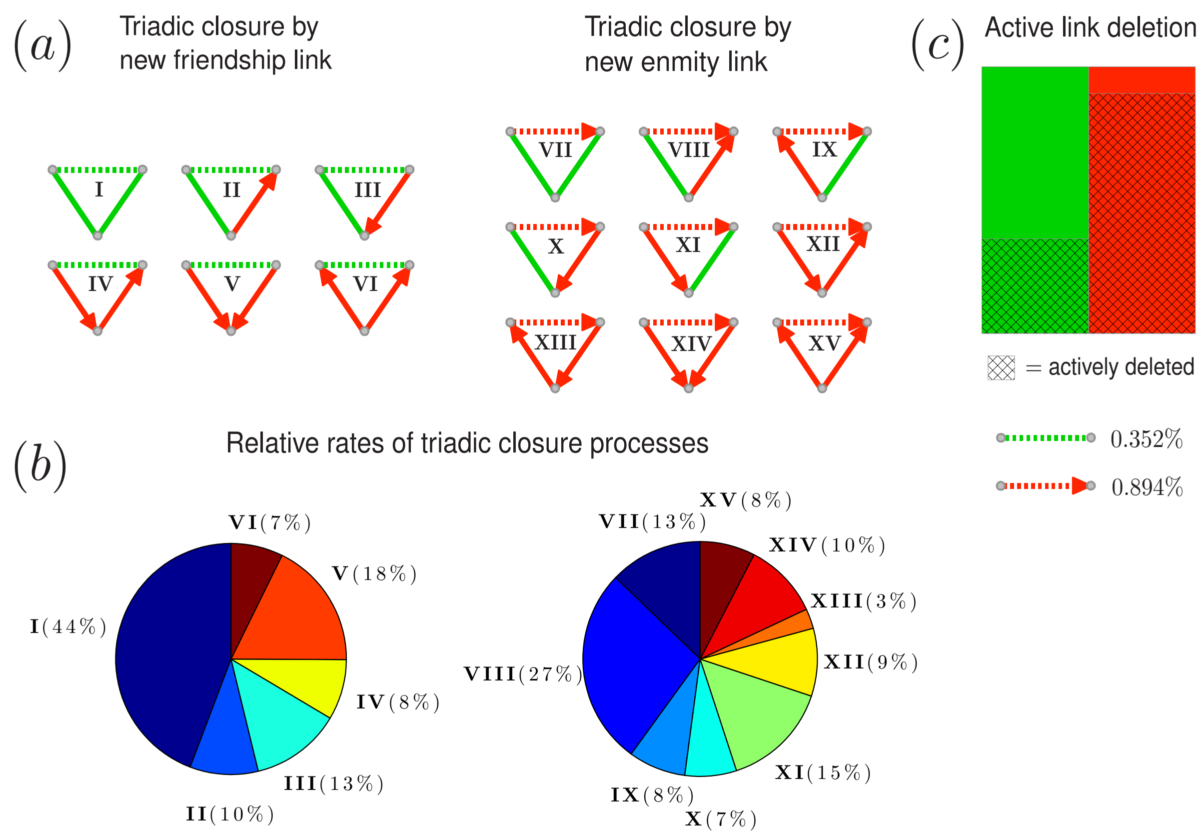
Understand social multiplex network!



- **Input:** transition rates and birth-death rates, as measured in game

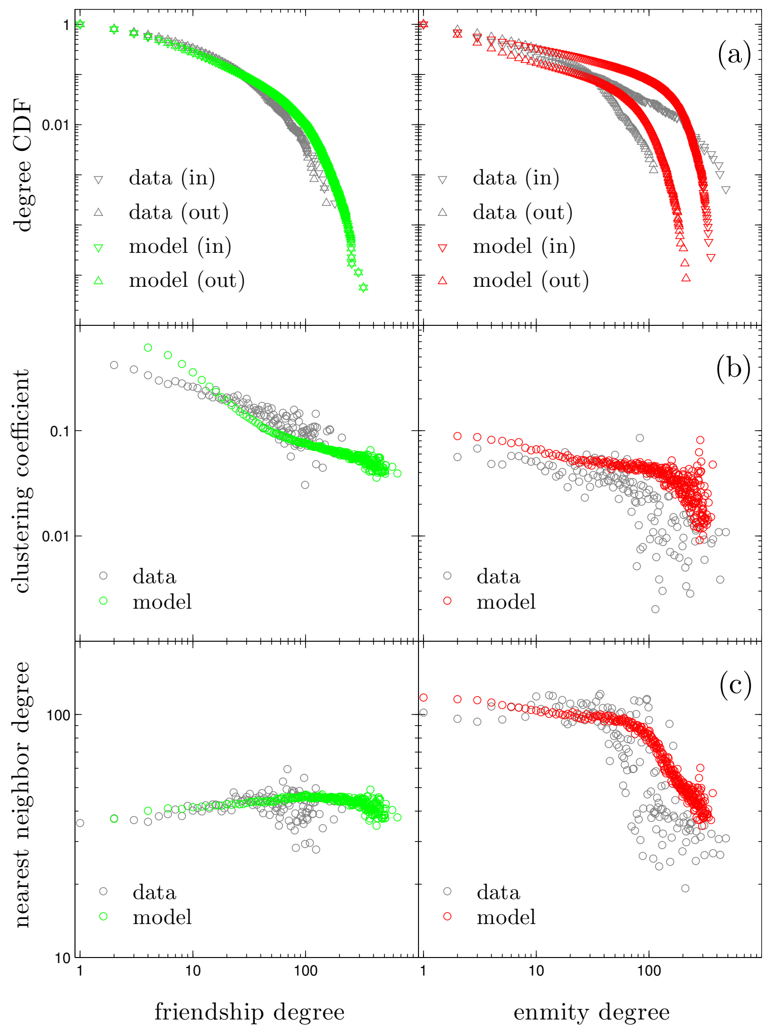
Friends and enemies—how does this work?




(a)
Triadic closure by new friendship link
I
II
III
IV
V
VI
Triadic closure by new enmity link
VII
VIII
IX
X
XI
XII
XIII
XIV
XV
(c)
Active link deletion
= actively deleted
0.352%
0.894%
(b)
Relative rates of triadic closure processes
VI(7%)
V(18%)
I(44%)
IV(8%)
III(13%)
II(10%)
XV(8%)
VII(13%)
XIV(10%)
XIII(3%)
XII(9%)
VIII(27%)
XI(15%)
IX(8%)
X(7%)

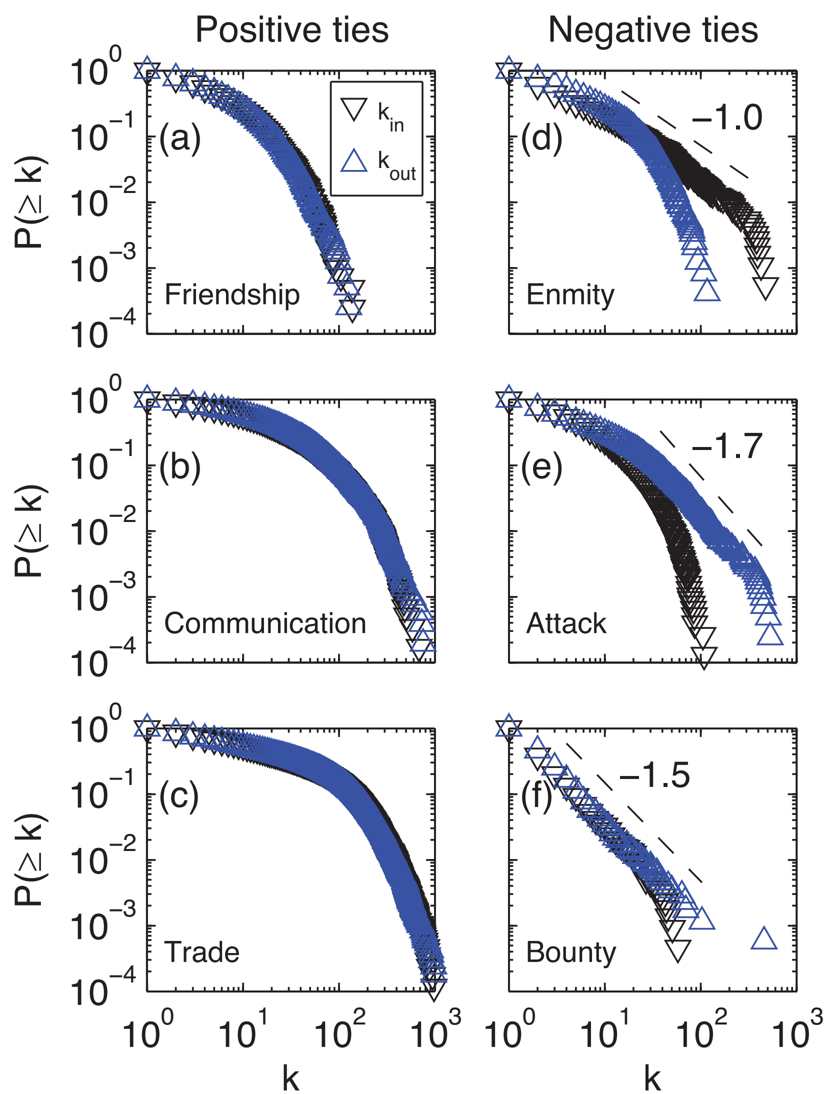
Two-layer triadic closure model

1. With p, add friendship link
   (a) **Friendship TC.** With s_f, pick node randomly and connect two randomly selected friends
   (b) With $1 - s_f$ (if no friendship triad is closed), pick node randomly and connect one of its friends with randomly chosen node
2. With $1 - p$ add an enmity link to the network:
   (a) **Mixed TC.** With s_e, pick node with $k_f \geq 1$ and enmity out-degree $k_e^{out} \geq 1$ randomly, and randomly select one neighbor in each layer. Connect friend with enemy s.t. new link points towards the enemy
   (b) With $1 - s_e$, pick node randomly and connect one of its enemies with any node such that new link points towards enemy
3. **Active friendship link deletion.** With r_f^*, pick node and remove one of its friendship links
4. **Active enmity link deletion.** With r_e^*, pick node and remove one of its outgoing enemity links
5. **Node turnover.** With q, pick node and remove it from the network with all its links. Introduce new node and link it to two randomly selected nodes (one friendship link and one incoming enmity link). Continue with timestep $t + 1$.



degree CDF
1
0.01
(a)
data (in)
data (out)
model (in)
model (out)
clustering coefficient
0.1
0.01
(b)
data
model
nearest neighbor degree
100
10
(c)
1
10
100
friendship degree
enmity degree

Negative ties are power laws – positive are not



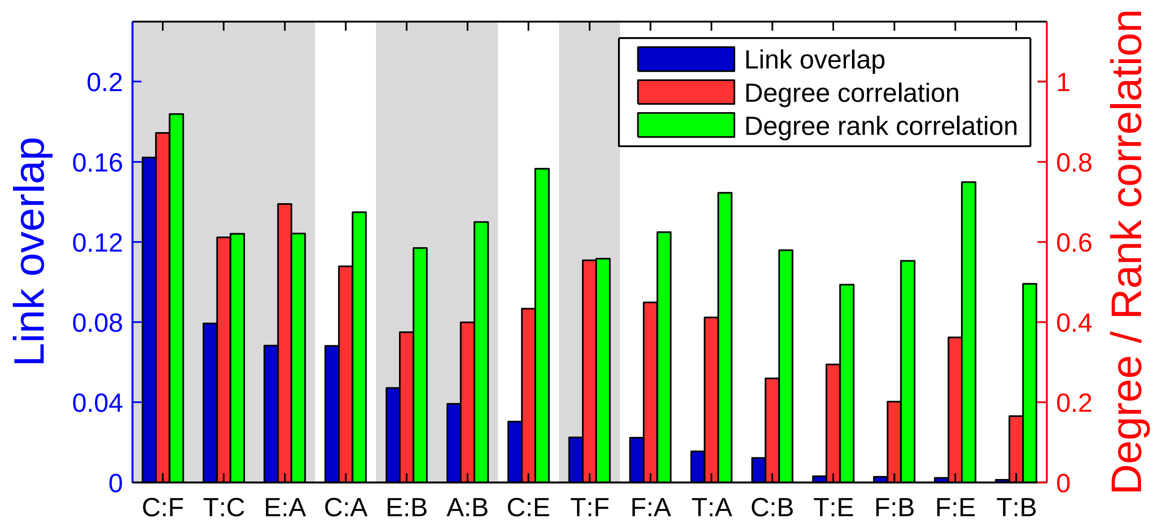
Network-network interactions

- Link overlap: $O_{\alpha\beta} = \frac{1}{2}\sum_{ij} M_{ij}^{\alpha} M_{ij}^{\beta}$

measures overlap between networks α and β

- Degree correlation: $\rho(k_\alpha, k_\beta) = \frac{E\left[(k_\alpha - \bar{k}_\alpha)(k_\beta - \bar{k}_\beta)\right]}{\sigma_{k_\alpha}\sigma_{k_\beta}} \in [-1, 1]$

people who have high / low degrees in one NW tend to have high / low degrees in another



Network-network correlations: interpretation

• **Communication – Friendship:** players who communicate with many (few) others tend to have many (few) friends. High overlap: friends tend to talk with each other

• **Enmity – Attack:** High correlation: aggressors or victims of aggression tend to be involved in many hostile relations. High overlap: enemies tend to attack each other – or attacks are likely to lead to enemy markings

• **Communication – Attack:** High correlation: players who communicate with many (few) tend to attack or be attacked by many (few) players. Relatively high overlap: there communication between players who attack each other. Aggression is not anonymous.

• **Friendship – Attack:** Relatively high correlation: players with many (few) friends attack or are attacked by many (few) others. Low overlap: these attacks tend to not take place between friends, or fighting players do not become friends

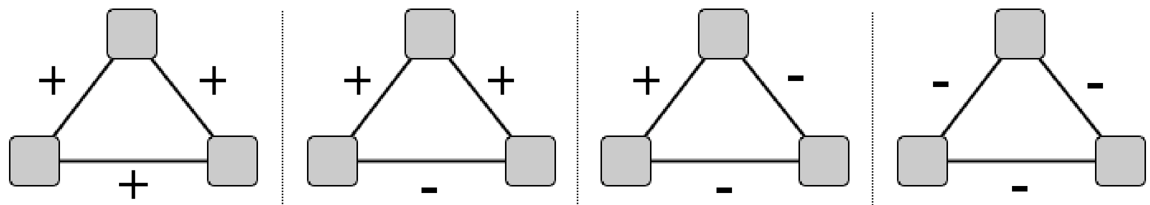
• **Friendship – Enmity:** Correlation is substantial: players who are socially active tend to place both positive as well as negative links. Vanishing overlap: absence of ambivalent relations. Friends are never enemies.

• **Trade – Other:** Trade activities correlate with all other social activities, however to moderate extent. Overlap strongly suppressed: which could be due to the large number of links in trade networks

• **Bounty – Other:** Correlations insignificant for all NWs, with exception of trade networks: players who are good in trade have a preference to act out negative sentiments

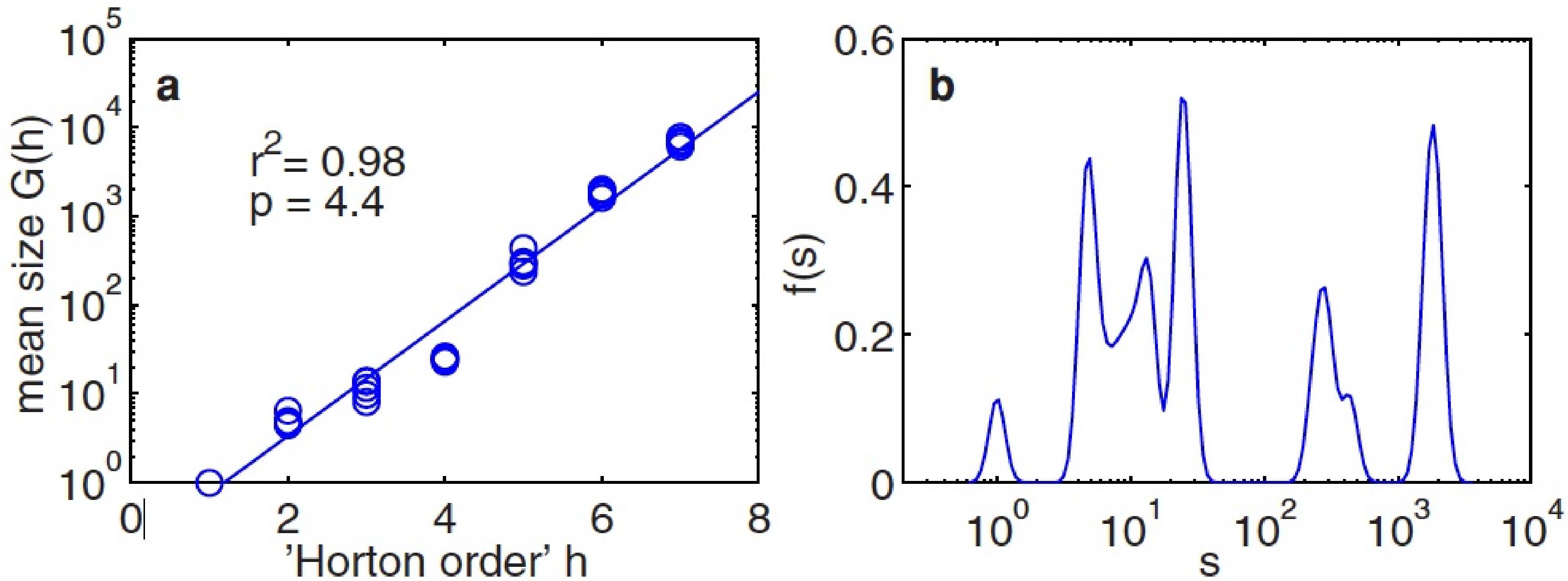
Measuring Social Balance

Signed graphs: friend + ; enemy –



	+ + +	+ + -	+ - -	- - -
Strong formulation of balance	B	U	B	U
Weak formulation of balance	B	U	B	B
N_Δ	26,329	4,428	39,519	8,032
$N_{\Delta,r}$	10,608	30,145	28,545	9,009
z	71	-112	47	-5

Avatars organize in multiples of four



Dunbar numbers?

III Behavioral Code

Human behavioral code

8-letter alphabet of actions

A $\cdots$ attack $\cdots$ bad
B $\cdots$ bounty $\cdots$ bad
C $\cdots$ communicate $\cdots$ good
D $\cdots$ delete friend $\cdots$ bad
E $\cdots$ make enemy $\cdots$ bad
F $\cdots$ make friend $\cdots$ good
T $\cdots$ trade $\cdots$ good
X $\cdots$ delete enemy $\cdots$ good

other actions: move, produce, work, do nothing, ...

Multiplex networks as interacting sequences

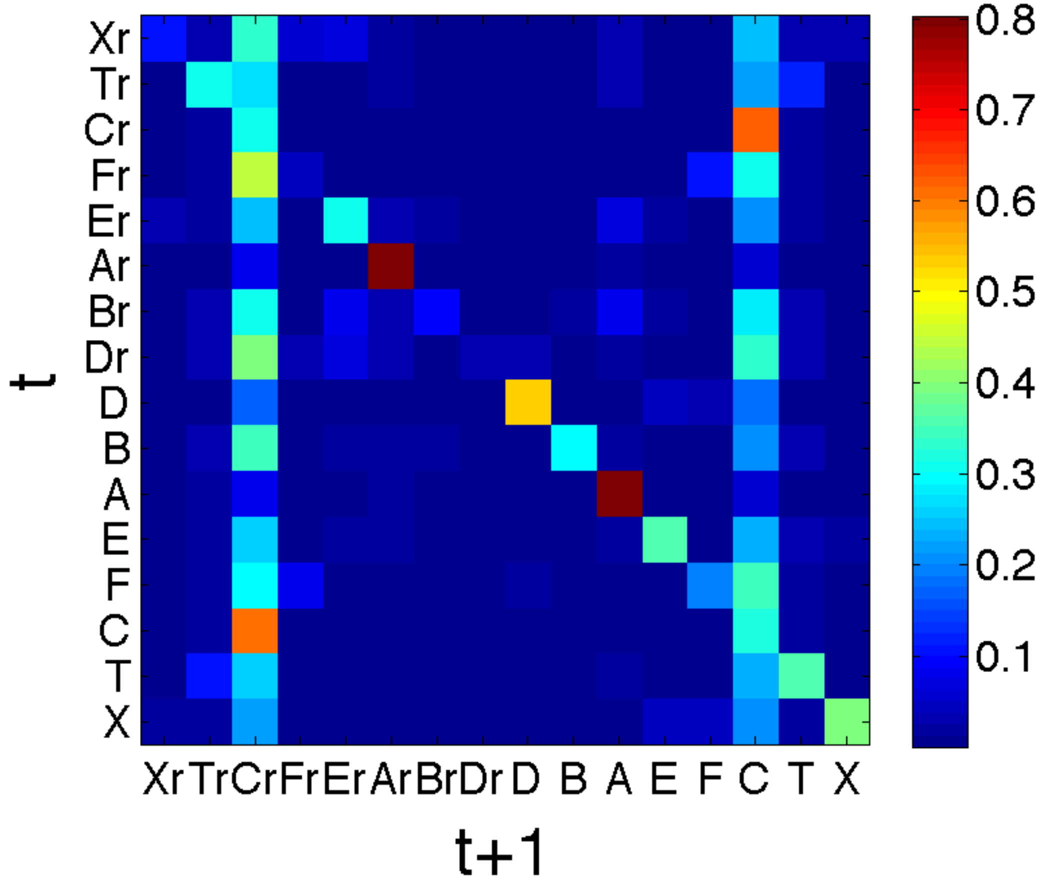
Player 146 ...AAAAAACTT EEX FTTTTTX CCCTTTTT AC...

Player 199 ...CCA BCAAAAATTA AACCCCCBX CFFFF...

Player 701 ...CCCCTTTT TCTCT FF CXXTT CCCCC TTT ...

Player 199 all ...CCAABCAAAAATTTAAATCCCCCCTBXCCFFFF.

Behavioral code and transition probabilities



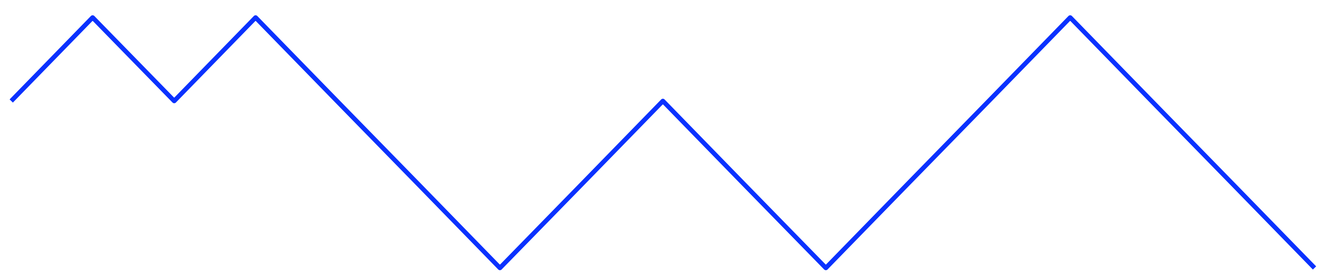
Probability of negative action following a positive received action? $p = 0.03$
Probability of negative action following a negative received action? $p = 0.22$

Homo sapiens gets up to 10 times more aggressive - if treated badly

“Good vs evil – good will prevail” (G.W. Bush)

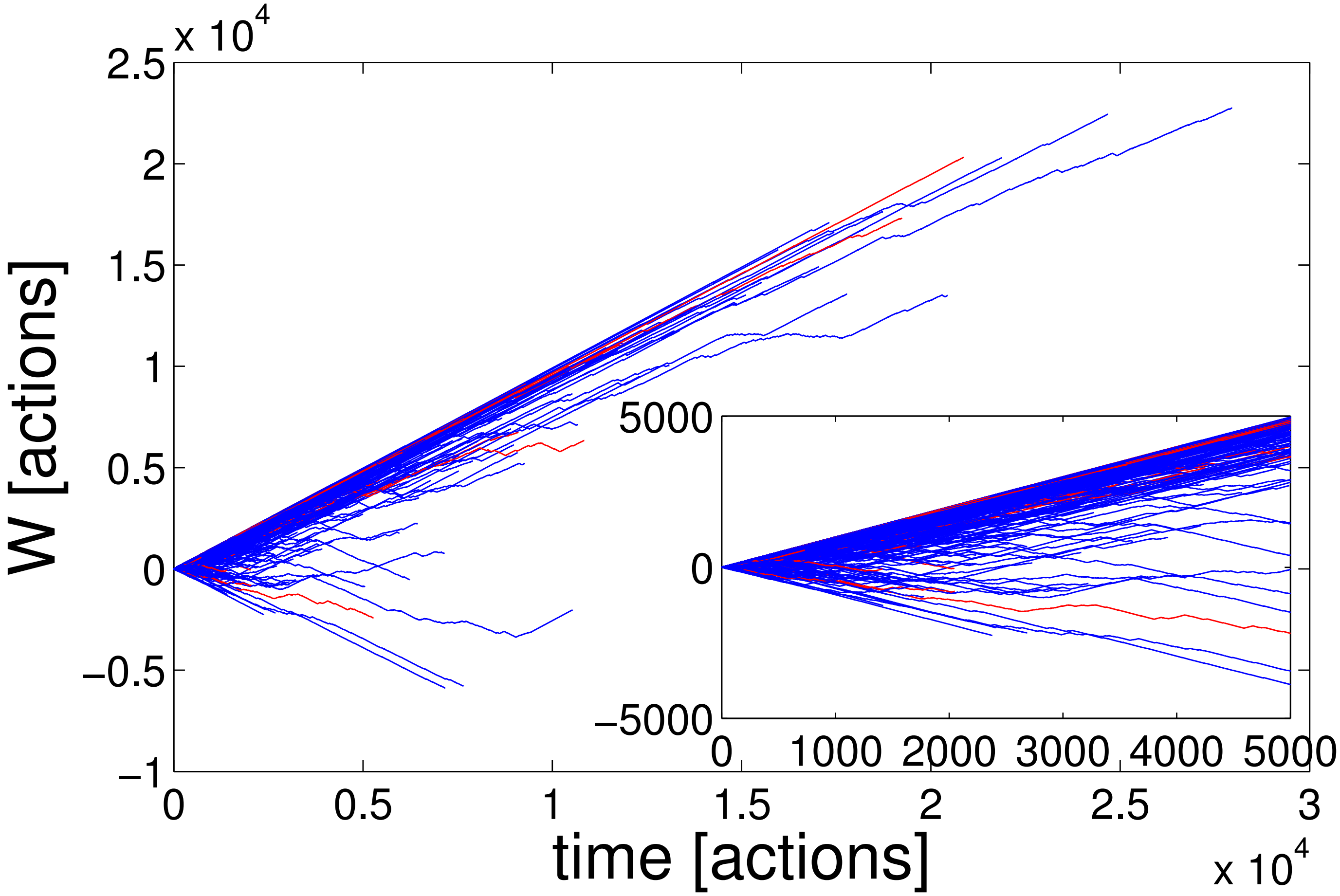
C A C A A E C C E A C C C A A A

+ − + − − − + + − − + + + − − −

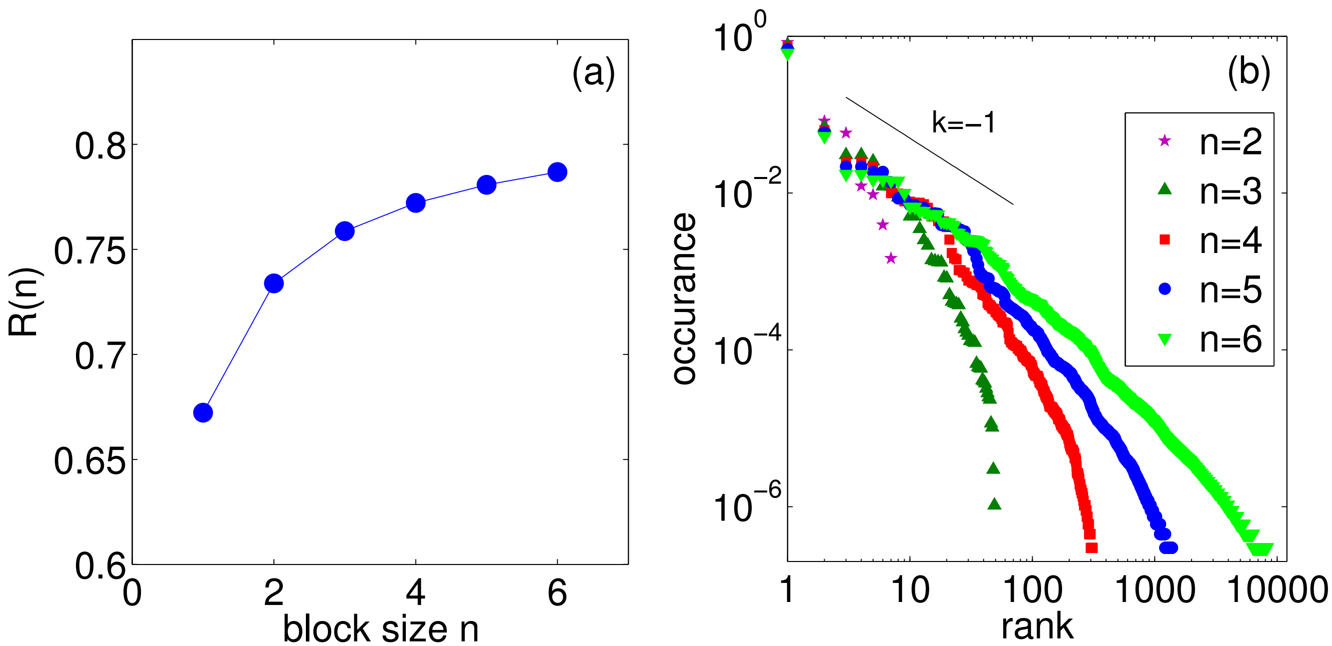


interpret action sequence as random walks

Worldlines of players



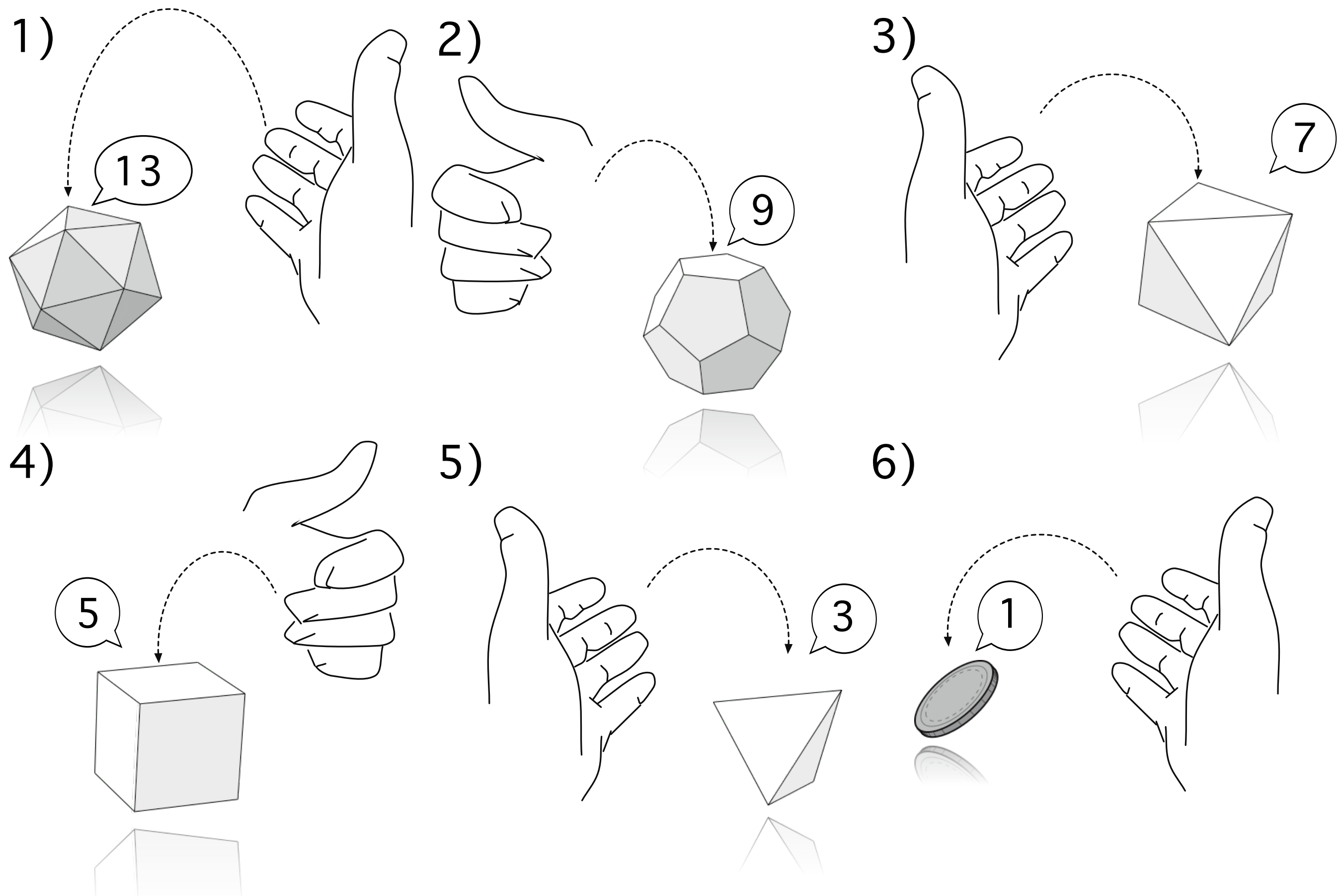
Zipf law in human behavioral code



$$R^{(n)} = 1 + \frac{1}{2n} \sum_{i}^{8^n} P_i^{(n)} \log P_i^{(n)}$$

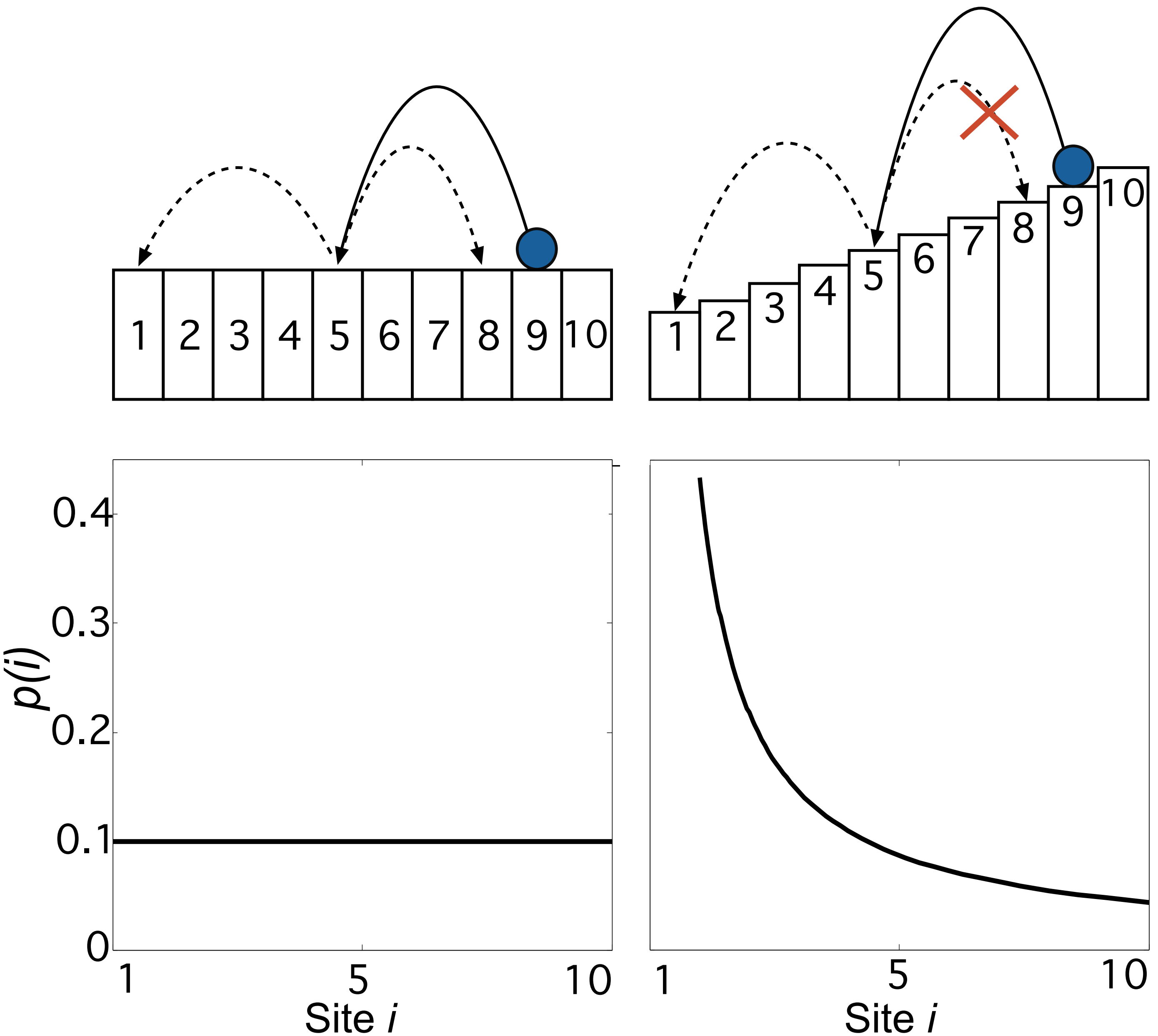
P_i probability for specific n letter word. Uncorrelated sequences: $R^{(n)} = 0$
Only one letter being used: $R^{(n)} = 1$

Statistics of path-dependent processes



Sample-space reduces and has a nested structure

$$\Omega_1 \subset \Omega_2 \subset \ldots \subset \Omega_N \subset \Omega$$


1 2 3 4 5 6 7 8 9 10
1 2 3 4 5 6 7 8 9 10
p(i)
0.4
0.3
0.2
0.1
0
1 5 10
Site i
1 5 10
Site i

This gives exact Zipf law !

Probability to visit site i

$$p(i) = \frac{1}{i}$$

Proof by induction

Let $N = 2$. There are two sequences ϕ: either ϕ directly generates a 1 with $p = 1/2$, or first generates 2 with $p = 1/2$, and then a 1 with certainty. Both sequences visit 1 but only one visits 2. As a consequence, $P_2(2) = 1/2$ and $P_2(1) = 1$.

Now suppose $P_{N-1}(i) = 1/i$ holds. Process starts with dice N, and probability to hit i in the first step is $1/N$. Also, any other j, $N \geq j > i$, is reached with probability $1/N$. If we get $j > i$, we get i in the next step with probability $P_{j-1}(i)$, which leads to a recursive scheme for $i < N$, $P_N(i) = \frac{1}{N}\left(1 + \sum_{i<j\leq N} P_{j-1}(i)\right)$. Since by assumption $P_{j-1}(i) = 1/i$, with $i < j \leq N$ holds, some algebra yields $P_N(i) = 1/i$.

The role of driving rates

restart process before it reaches bottom, with probability $1-\lambda$

$$p(i) = i^{-\lambda}$$

$(1-\lambda)$ is **driving rate** for SSR process

The role of driving – result is exact too

Clearly, $p^{(\lambda)}(i) = \sum_{j=1}^{N} P(i|j)\, p^{(\lambda)}(j)$ holds, with

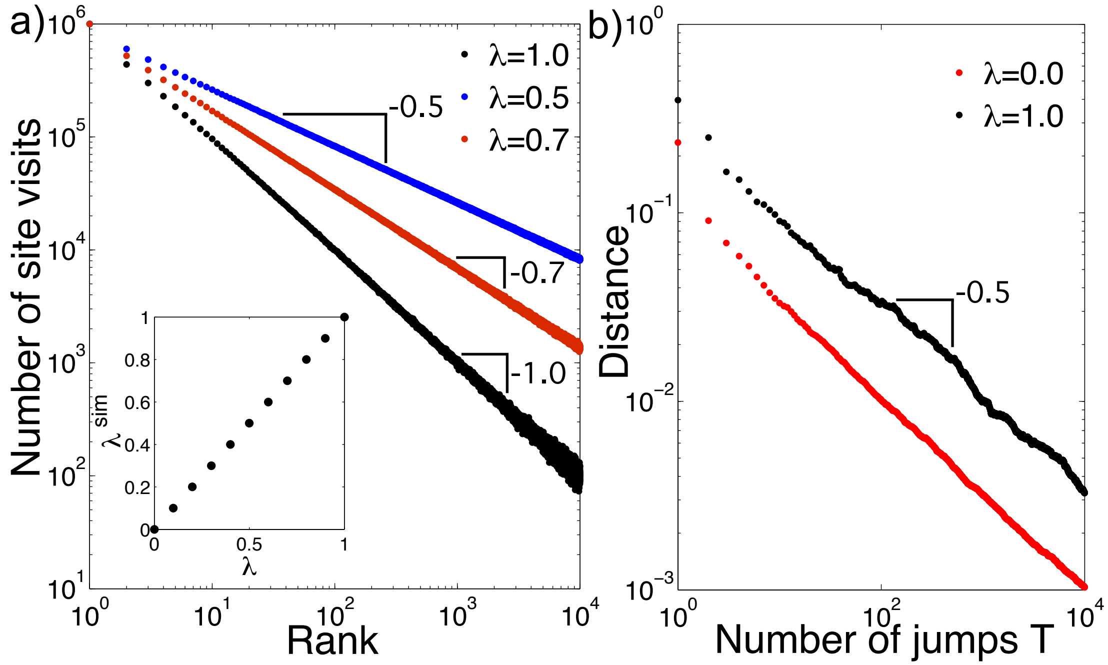
$$P(i|j) = \begin{cases} \frac{\lambda}{j-1} + \frac{1-\lambda}{N} & \text{for } i < j \quad (SSR) \\ \frac{1-\lambda}{N} & \text{for } i \geq j > 1 \quad (RW) \\ \frac{1}{N} & \text{for } i \geq j = 1 \quad (restart) \end{cases}$$

We get $p^{(\lambda)}(i) = \frac{1-\lambda}{N} + \frac{1}{N} p^{(\lambda)}(1) + \sum_{j=i+1}^{N} \frac{\lambda}{j-1}\, p^{(\lambda)}(j)$

to recursive relation $p^{(\lambda)}(i+1) - p^{(\lambda)}(i) = -\lambda \frac{1}{i} p^{(\lambda)}(i+1)$

$$\begin{aligned} \frac{p^{(\lambda)}(i)}{p^{(\lambda)}(1)} &= \prod_{j=1}^{i-1} \left(1 + \frac{\lambda}{j}\right)^{-1} = \exp\left[-\sum_{j=1}^{i-1} \log\left(1 + \frac{\lambda}{j}\right)\right] \\ &\sim \exp\left(-\sum_{j=1}^{i-1} \frac{\lambda}{j}\right) \sim \exp\left(-\lambda \log(i)\right) = i^{-\lambda} \end{aligned}$$

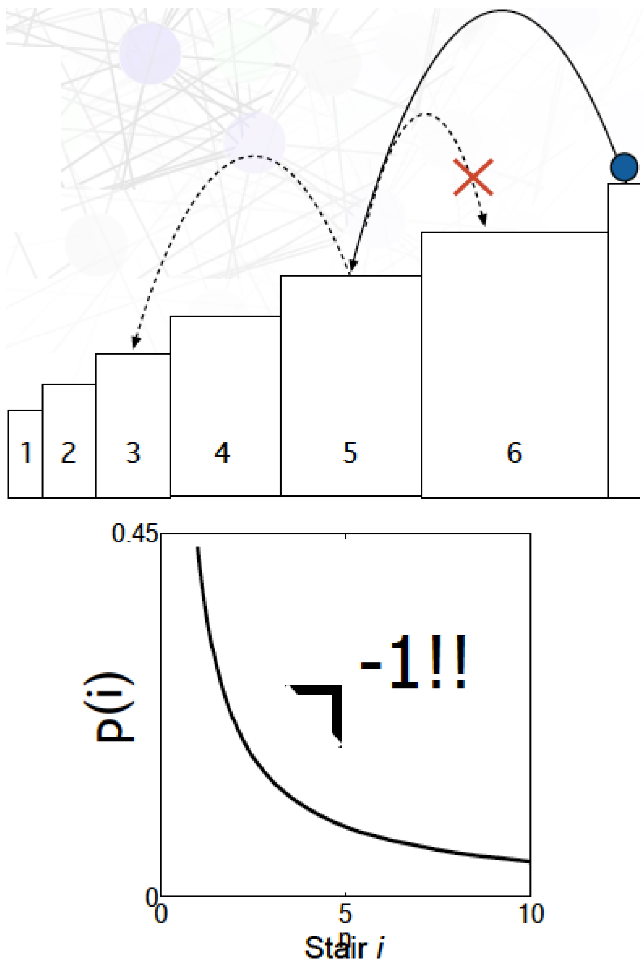
History-dependent processes with driving



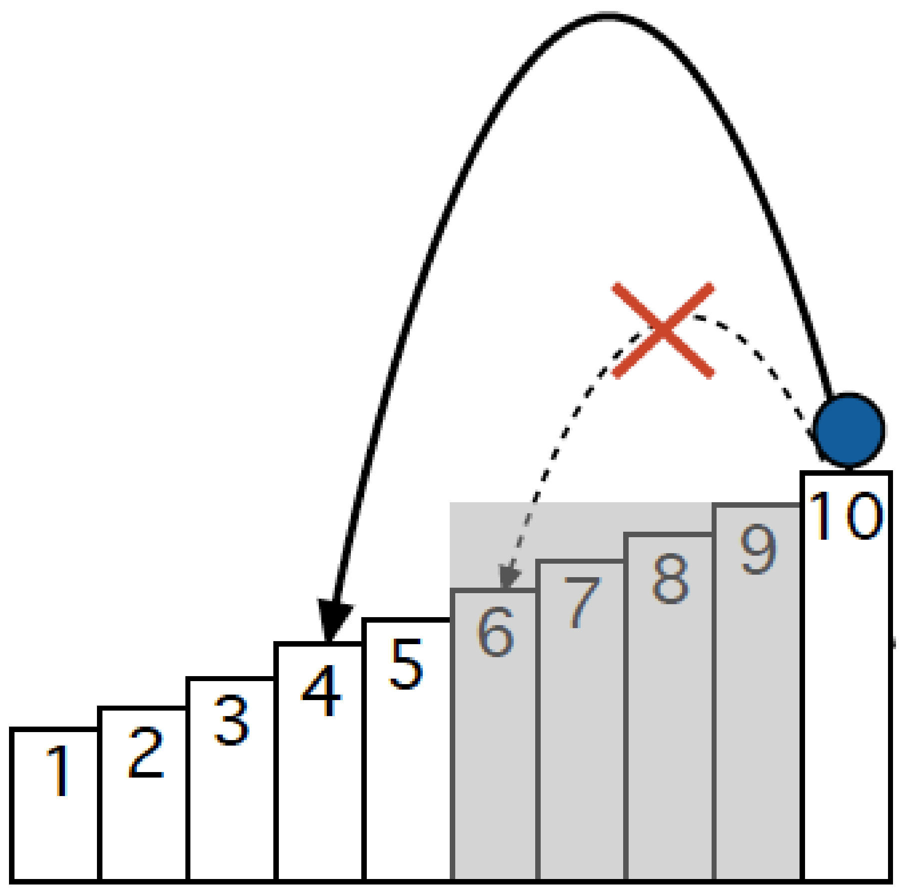
How fast does power law converge to its limiting distribution?

Answer: same convergence speed as central-limit theorem for iid processes (Berry-Esseen theorem)

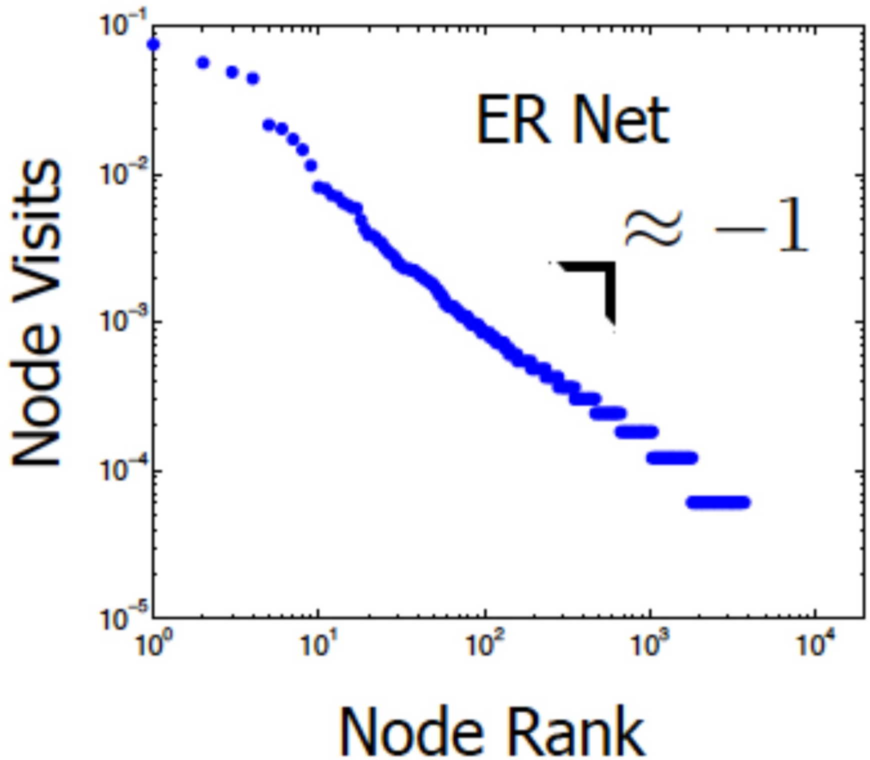
Zipf law is invariant under priors



Zipf law is remarkably robust – accelerate SSR

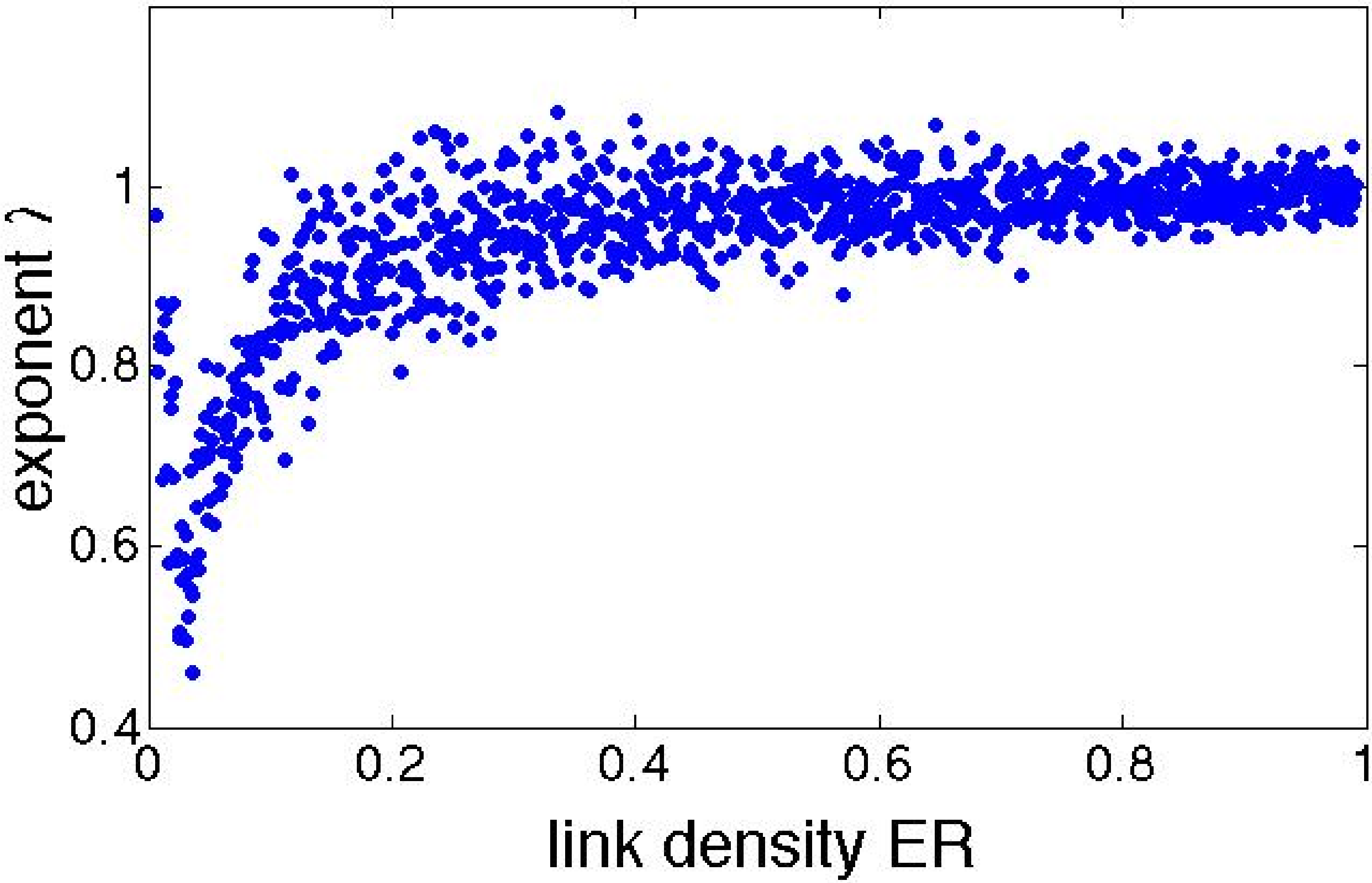


All diffusion processes on DAG are SSR



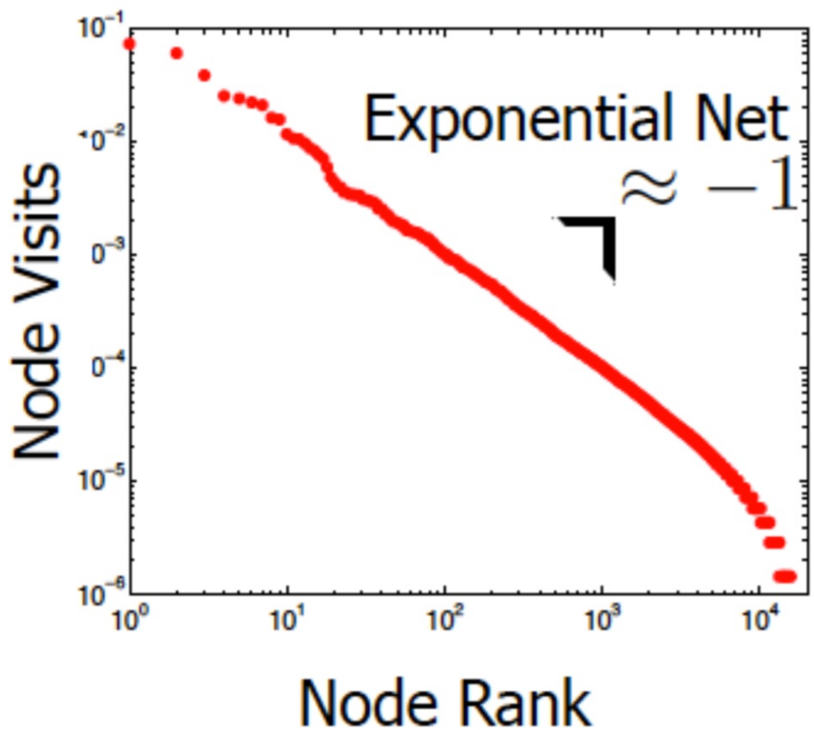
sample ER graph $\rightarrow$ direct it $\rightarrow$ pick start and end $\rightarrow$ diffuse

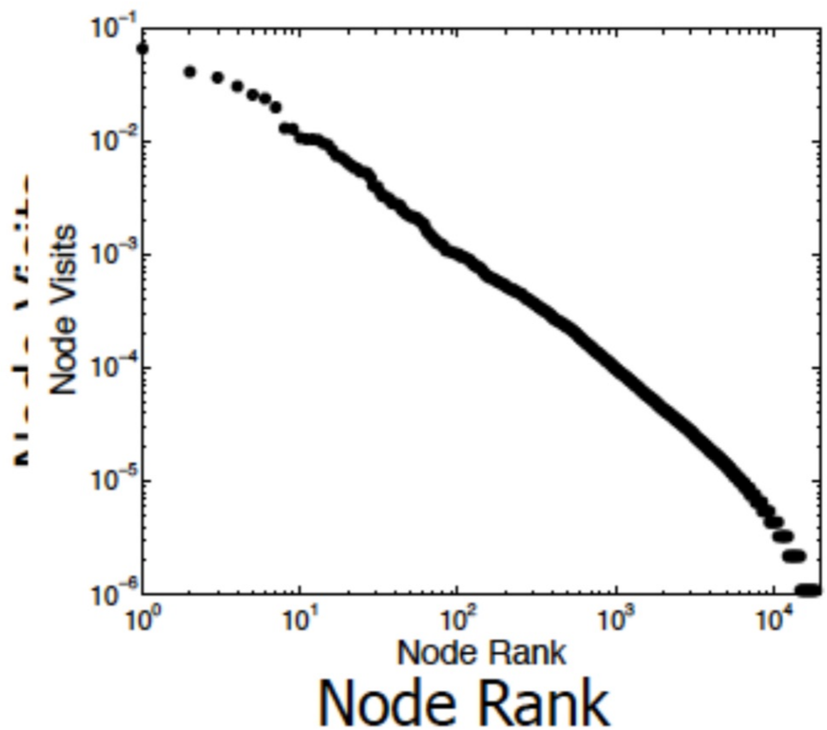
Zipf holds for any link probability



Exponential NW

HEP Co-authors

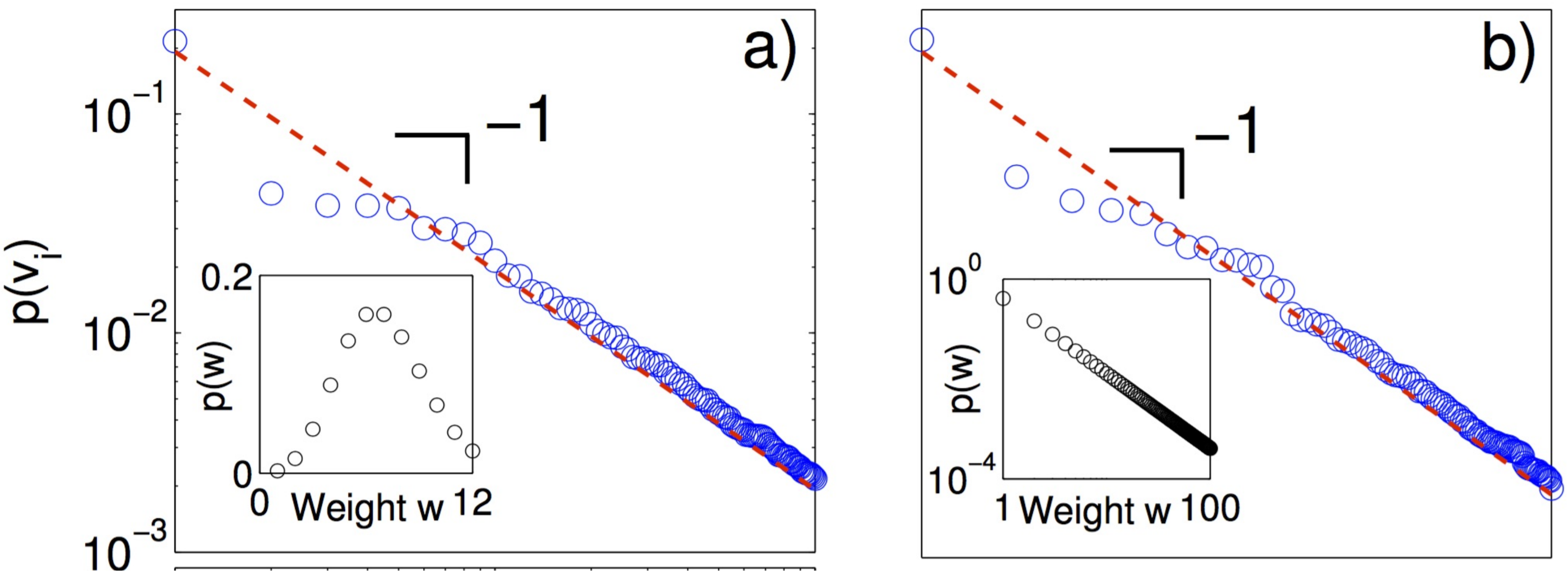




prior probabilities are practically irrelevant!

What happens if introduce weights on links?

ER Graph

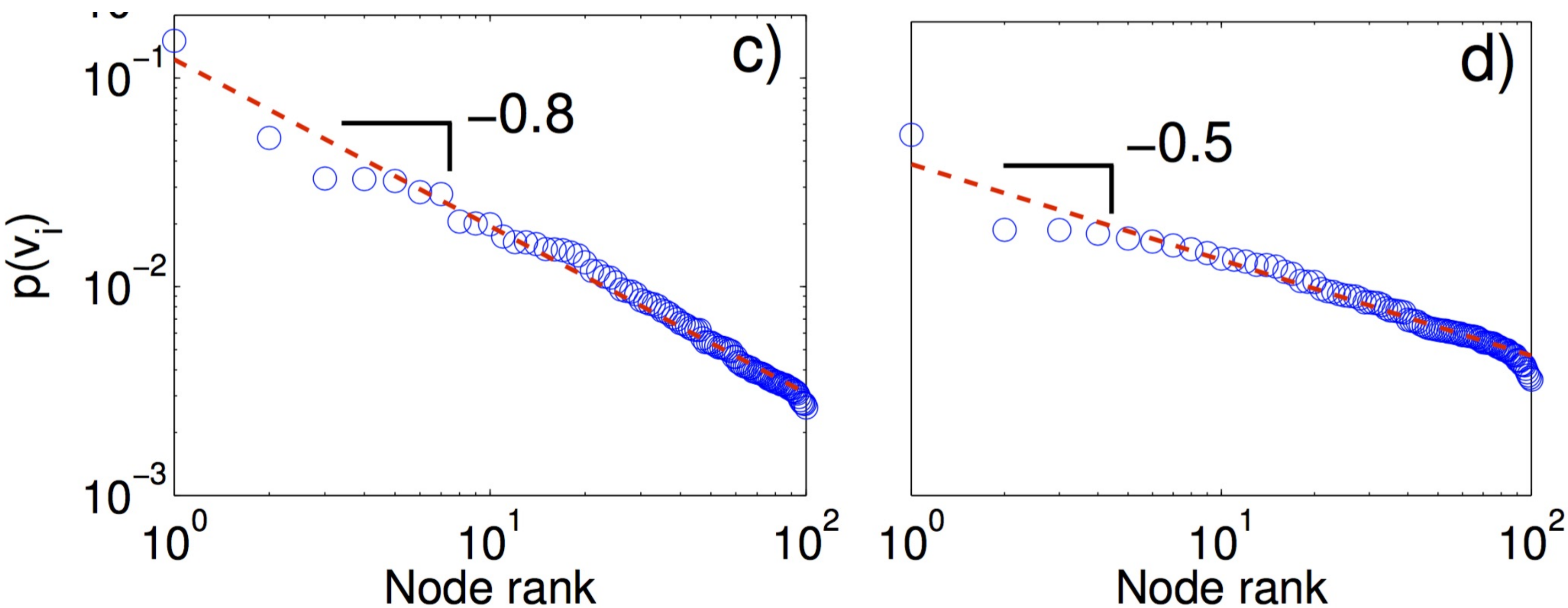


poisson weights   power weights

prior probabilities are practically irrelevant!

What happens if introduce cycles?

ER $\rightarrow$ direct it $\rightarrow$ change link to random direction with $1 - \lambda$



noise level $\lambda = 0.8$ $\lambda = 0.5$

Zipf's law is an immense attractor!

Zipf's law is an attractor

- no matter what the network topology is $\rightarrow$ Zipf
- no matter what the link weights are $\rightarrow$ Zipf
- if have cycles $\rightarrow$ exponent is less than one

Every good search process is SSR!

What is good search?

Search is a SSR process. Good search is ...

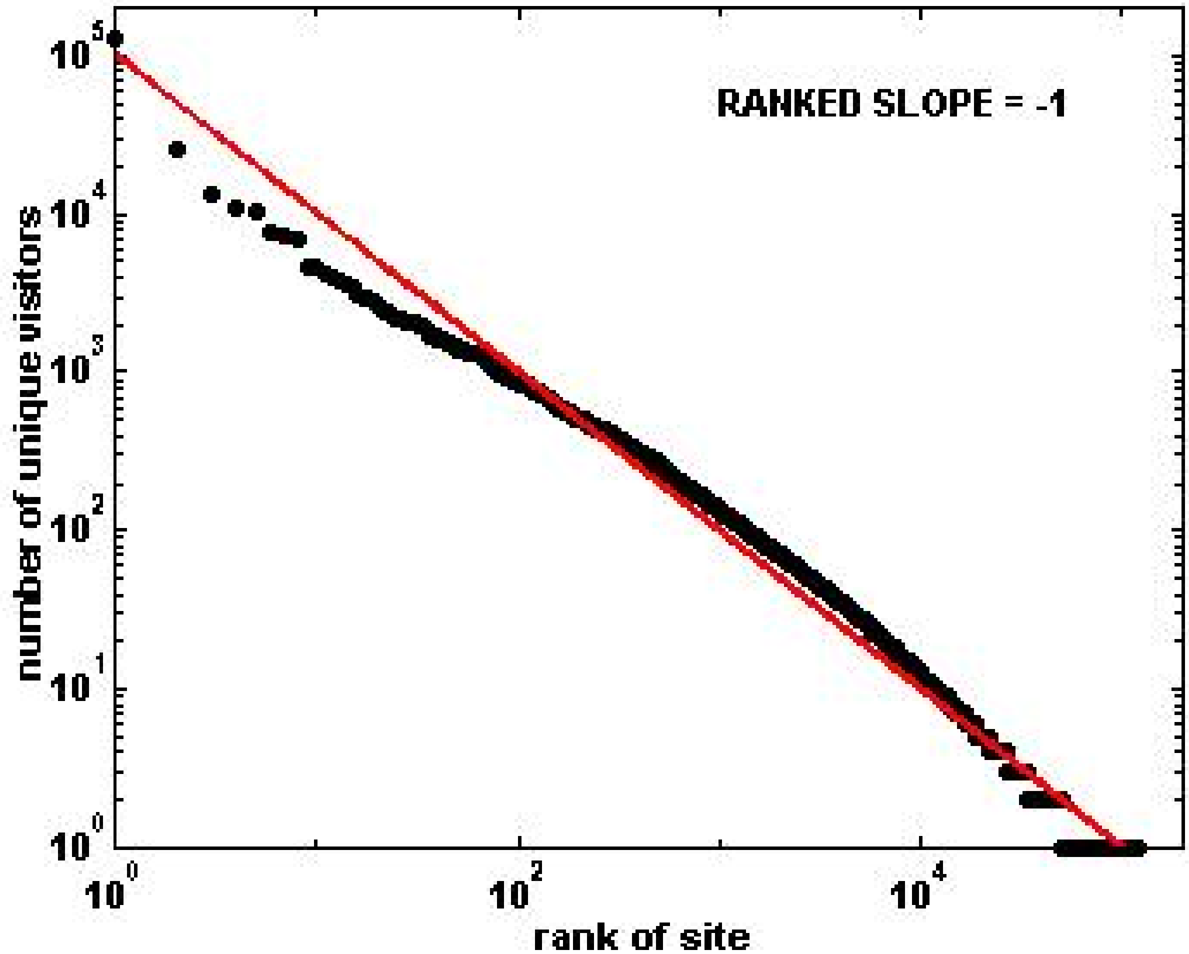
- ... if at every step you eliminate more possibilities than you actually sample

- ... every step you take eliminates branches of possibilities

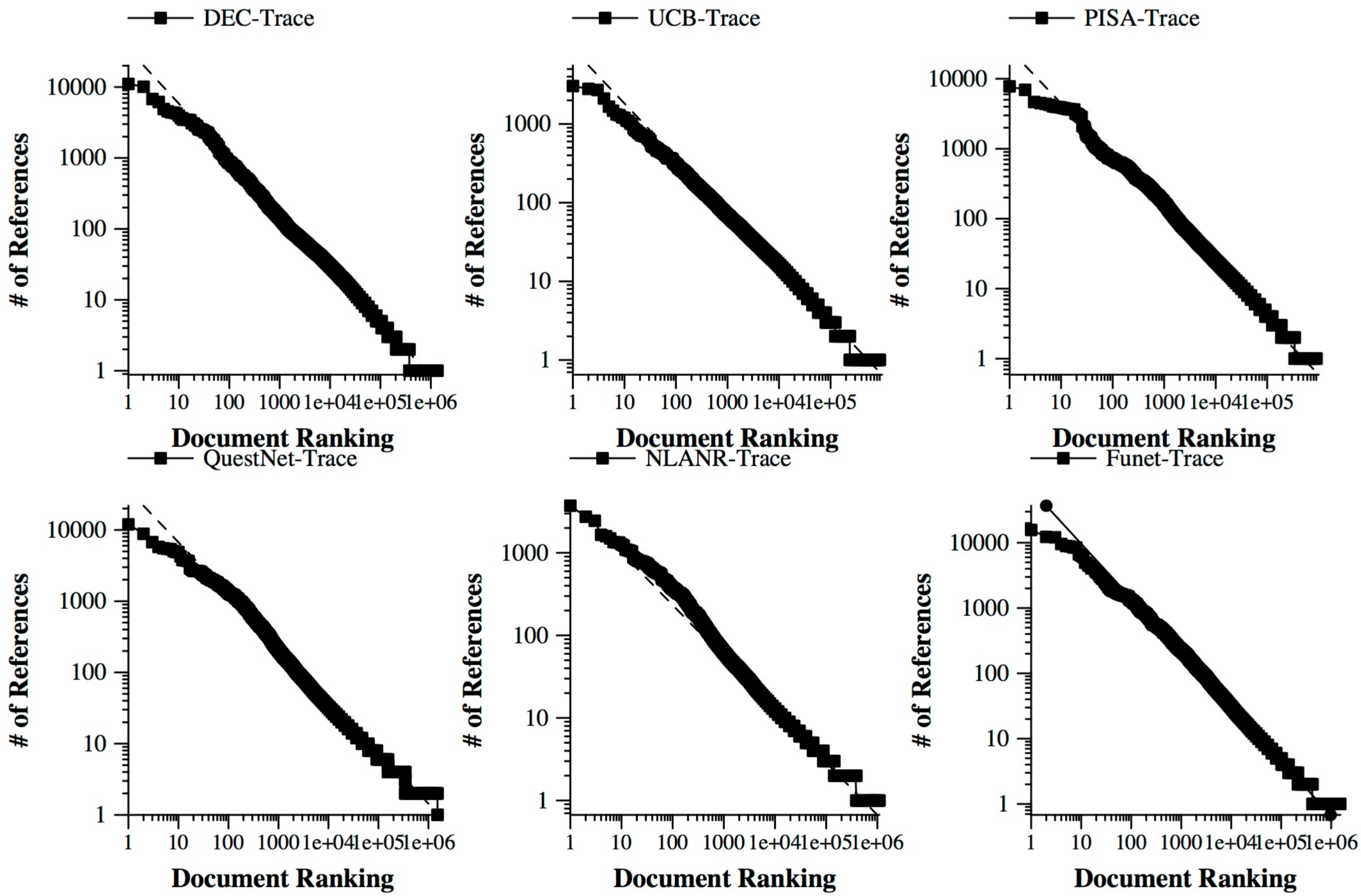
if eliminate fast enough $\rightarrow$ power law in visiting times

if eliminate too little $\rightarrow$ sample entire space (exhaustive search)

Clicking on web page is often result of search process

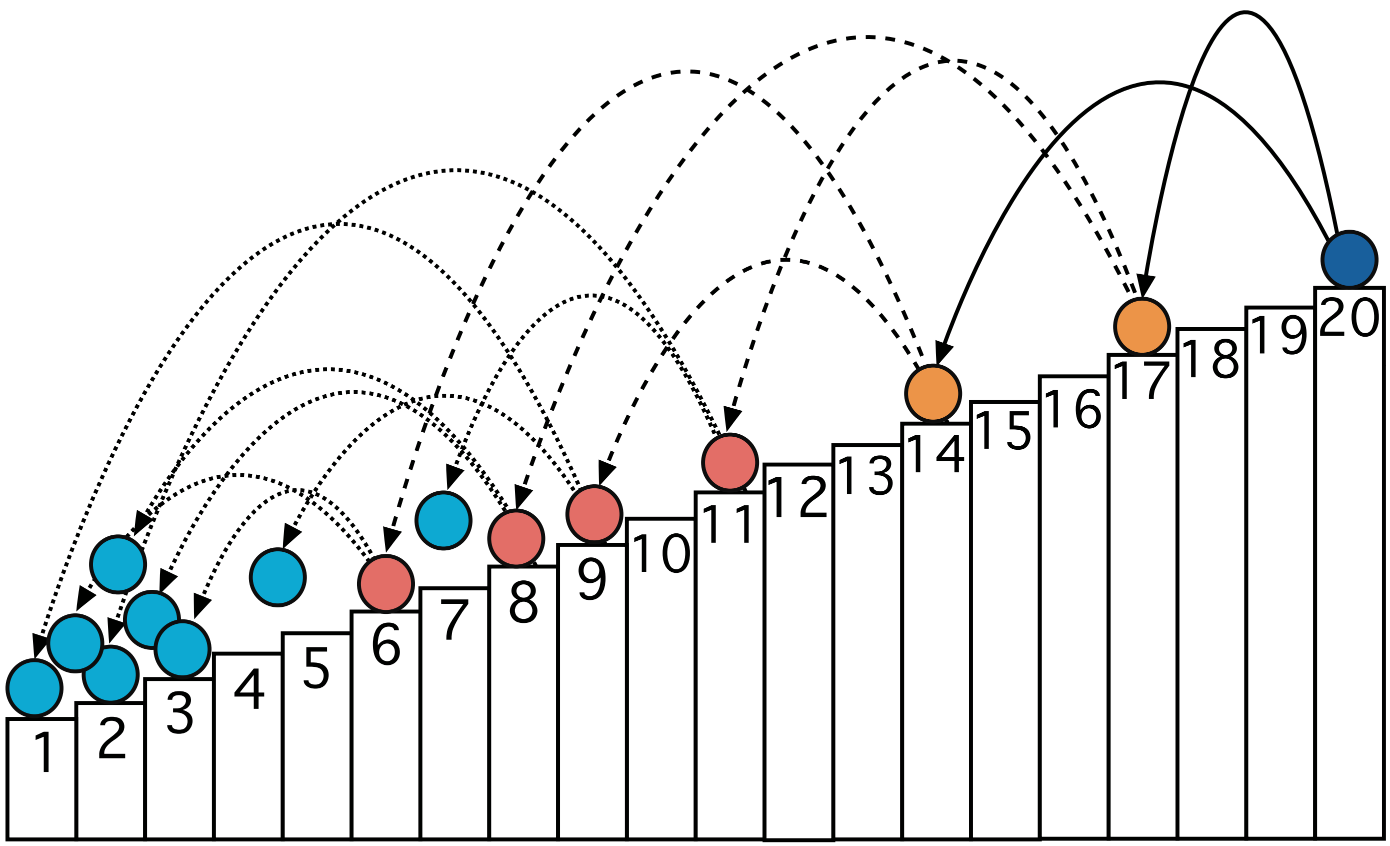


adamic & hubermann 2002



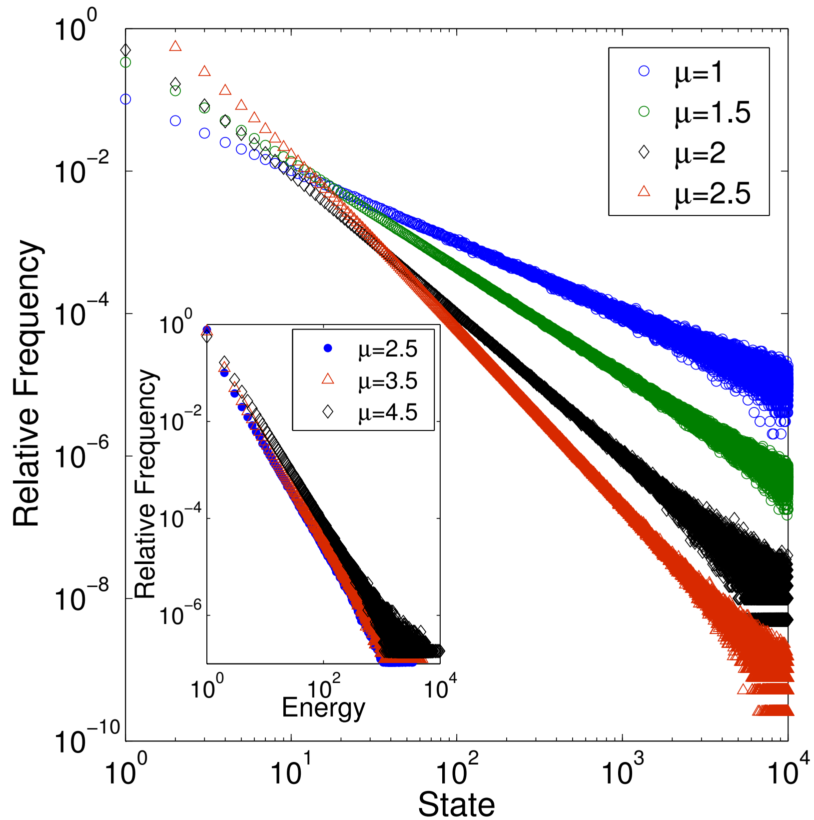
breslau et al 99

What about exponents > 1?



Multiplication factor μ

$$\rightarrow p(i) = i^{-\mu}$$


μ=1
μ=1.5
μ=2
μ=2.5
Relative Frequency
State
μ=2.5
μ=3.5
μ=4.5
Relative Frequency
Energy

What if we introduce conservation laws?

Conservation laws in SSR processes

Assume that you have duplication at every jump $\mu = 2$

If you are at $i \rightarrow$ duplicate $\rightarrow$ one jumps to j, the other to k

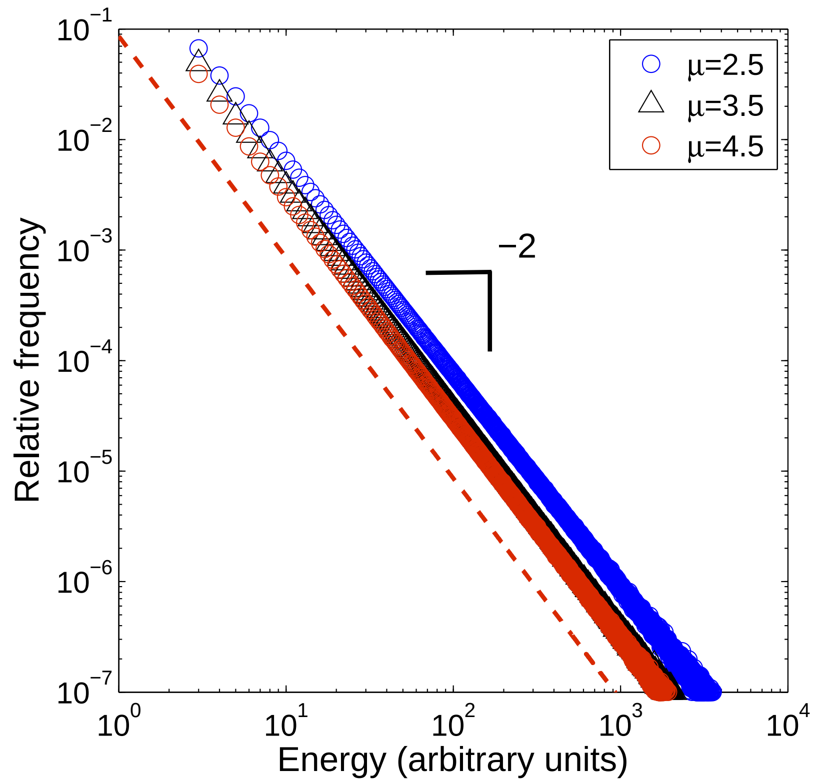
conservation means: $i = j + k$.

For any μ, conservation means:

$i = \mathrm{state}_1 + \mathrm{state}_2 \; + \cdots + \mathrm{state}_\mu$

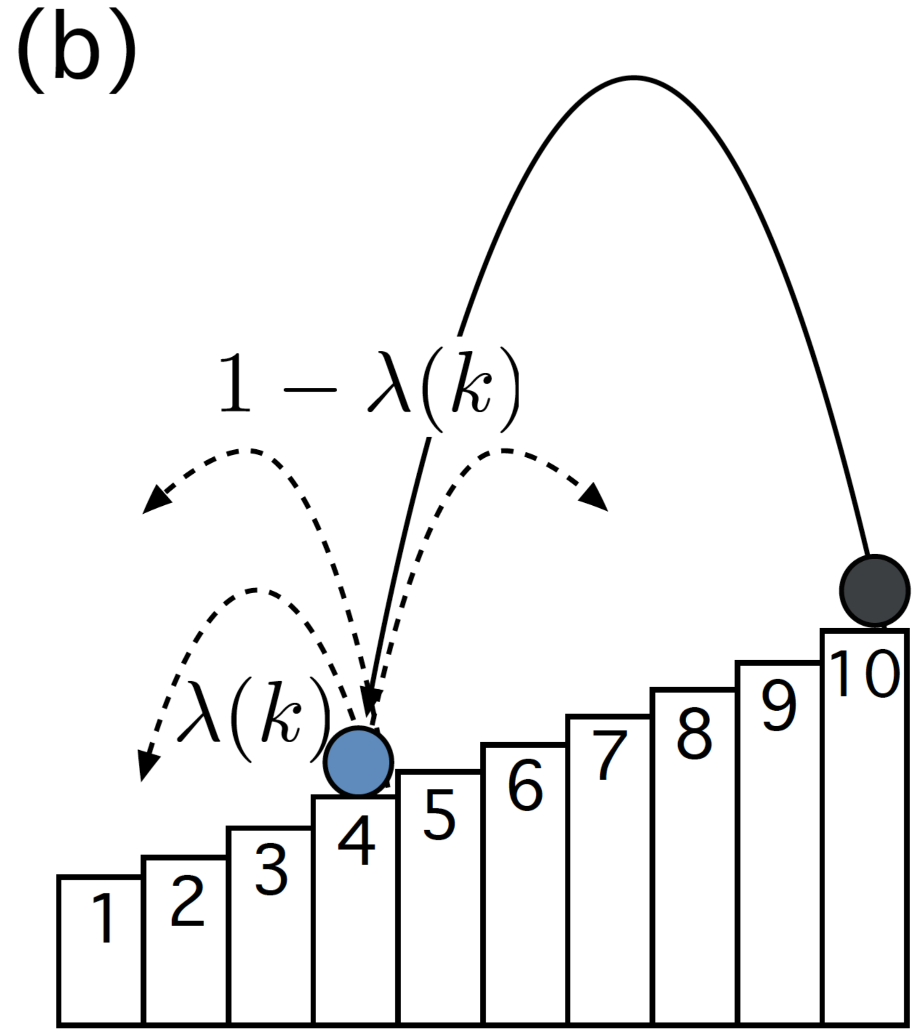
$$\rightarrow p(i) = i^{-2} \quad \text{for all} \quad \mu$$

This was found by E. Fermi for particle cascades



μ=2.5
μ=3.5
μ=4.5
−2
Relative frequency
Energy (arbitrary units)

Assume that driving rate depends on the state $\lambda(i)$



Can derive the relation

$$\rightarrow \lambda(x) = -x\frac{d}{dx}\log p_\lambda(x)$$

That can be proved as a theorem.

Proof

The transition probabilities from state k to i are

$$p_{\mathrm{SSR}}(i|k) = \begin{cases} \lambda(k)\frac{q_i}{g(k-1)} + (1-\lambda(k))q_i & \text{if} \quad i<k \\ (1-\lambda(k))q_k \; \text{otherwise} & , \end{cases}$$

$g(k)$ is the cdf of q_i, $g(k) = \sum_{i\leq k} q_i$. Observing that

$$\frac{p_{\lambda,q}(i+1)}{q_{i+1}}\left(1+\lambda(i+1)\frac{q_{i+1}}{g(i)}\right) = \frac{p_{\lambda,q}(i)}{q_i}$$

we get

$$p_{\lambda,q}(i) = \frac{q_i}{Z_{\lambda,q}}\prod_{1<j\leq i}\left(1+\lambda(j)\frac{q_j}{g(j-1)}\right)^{-1} \sim \frac{q(i)}{Z_{\lambda,q}}e^{-\sum_{j\leq i}\lambda(j)\frac{q(j)}{g(j-1)}}$$

$Z_{\lambda,q}$ is the normalisation constant. For uniform priors, taking logs and going to continuous variables gives the result $\lambda(x) = -x\frac{d}{dx}\log p_\lambda(x)$.

Special cases $\lambda(x) = -x\frac{d}{dx}\log p_\lambda(x)$

- Zipf: no noise $\rightarrow$ $p(x) = x^{-1}$
- Power-law: $\lambda(x) = \alpha$ $\rightarrow$ $p(x) = x^{-\alpha}$
- Exponential: $\lambda(x) = \beta x$ $\rightarrow$ $p(x) = e^{-\beta(x-1)}$
- Power-law + cut-off: $\lambda(x) = \alpha + \beta x$ $\rightarrow$ $p(x) = x^{-\alpha}e^{-\beta x}$
- Gamma: $\lambda(x) = 1 - \alpha + \beta x$ $\rightarrow$ $p(x) = x^{\alpha-1}e^{-\beta x}$

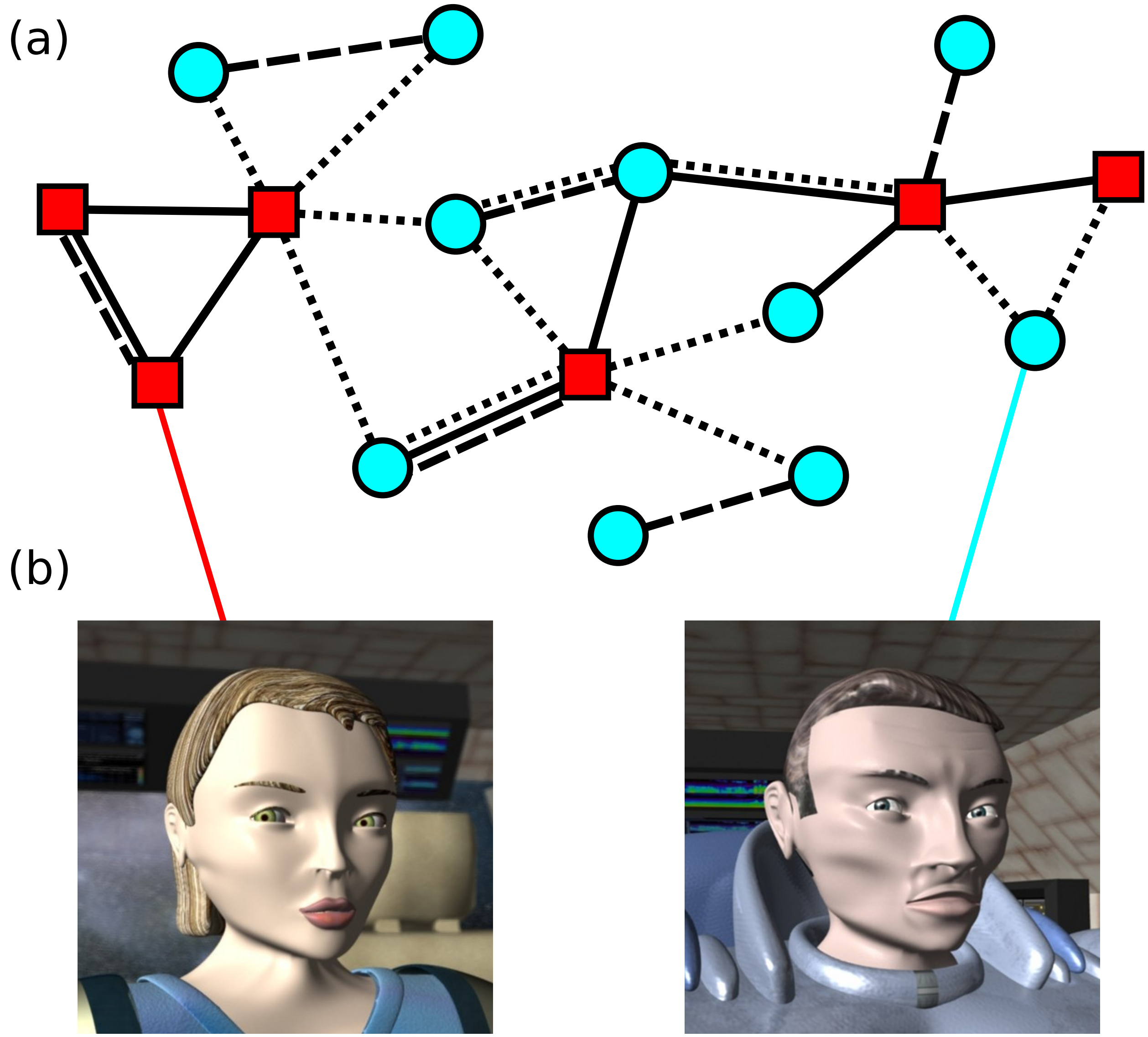
Special cases $\lambda(x) = -x\frac{d}{dx}\log p_\lambda(x)$

- Normal: $\lambda(x) = 2\beta x^2 \quad \rightarrow \quad p(x) = e^{-\frac{\beta}{2}(x-1)^2}$
- Stretched exp: $\lambda(x) = \alpha\beta|x|^\alpha \quad \rightarrow \quad p(x) = e^{-\frac{\beta}{\alpha}(x-1)^\alpha}$
- Gompertz: $\lambda(x) = (\beta e^{\alpha x} - 1)\beta x \quad \rightarrow \quad p(x) = e^{\beta x - \alpha e^{\beta x}}$
- Weibull: $\lambda(x) = \beta^{-\alpha}\alpha x^\alpha + \alpha - 1 \rightarrow p(x) = \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha}$
- Tsallis: $\lambda(x) = \frac{\beta x}{1-\beta x(1-Q)} \quad \rightarrow \quad p(x) = (1-(1-Q)\beta x)^{\frac{1}{1-Q}}$

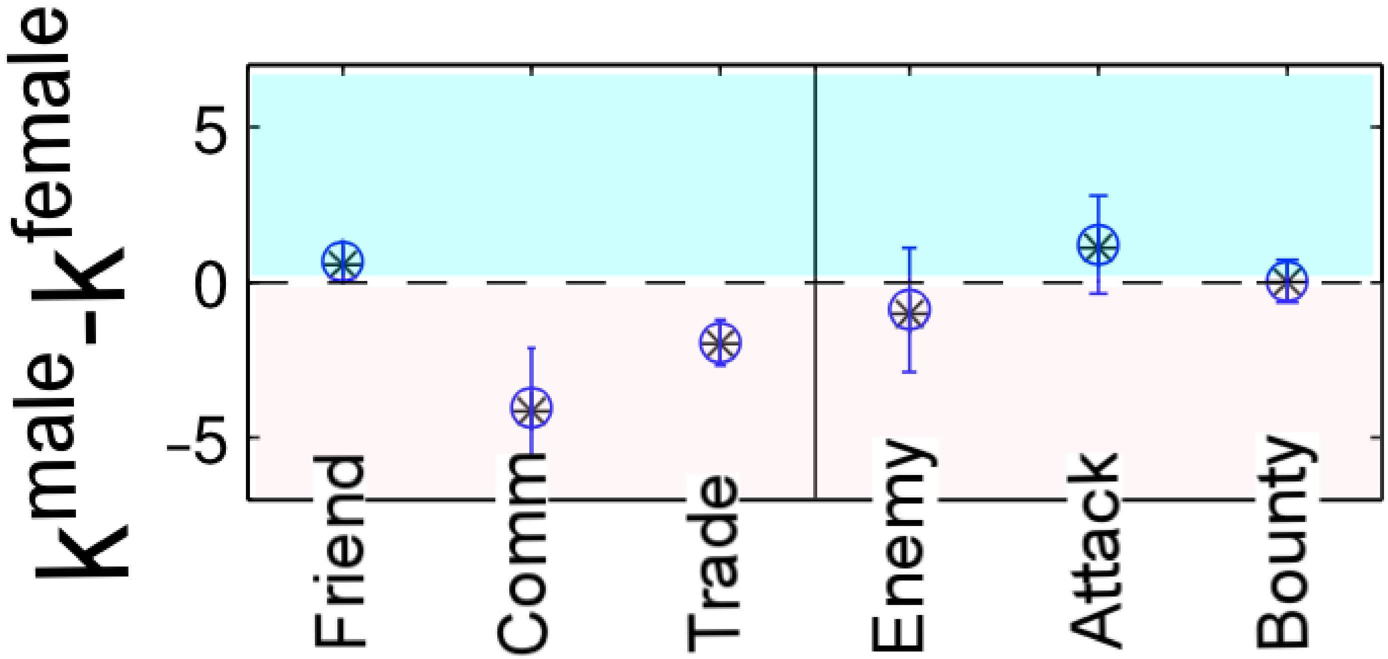
Problems that are of SSR nature

- search, e.g. targeted diffusion
- language: sentence formation
- fragmentation: break spaghetti
- scattering processes
- sequences of human behavior
- games: go
- internet communication

IV Gender differences

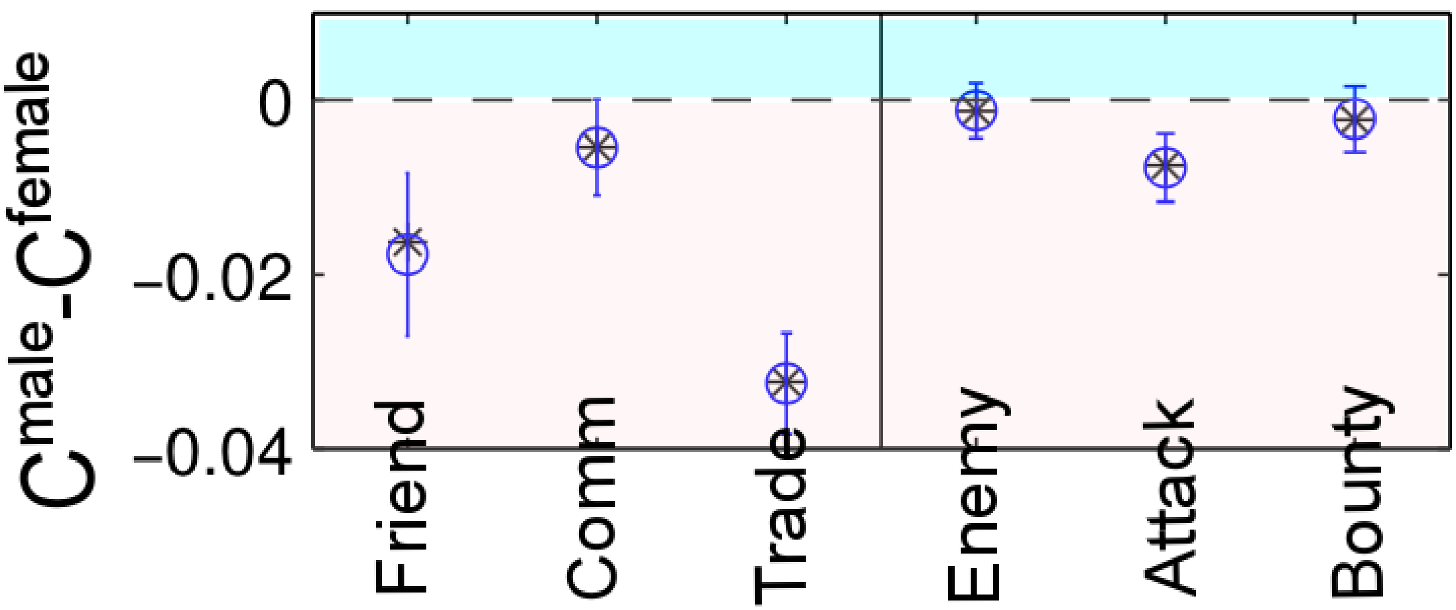

(a)
(b)

Gender differences: degree



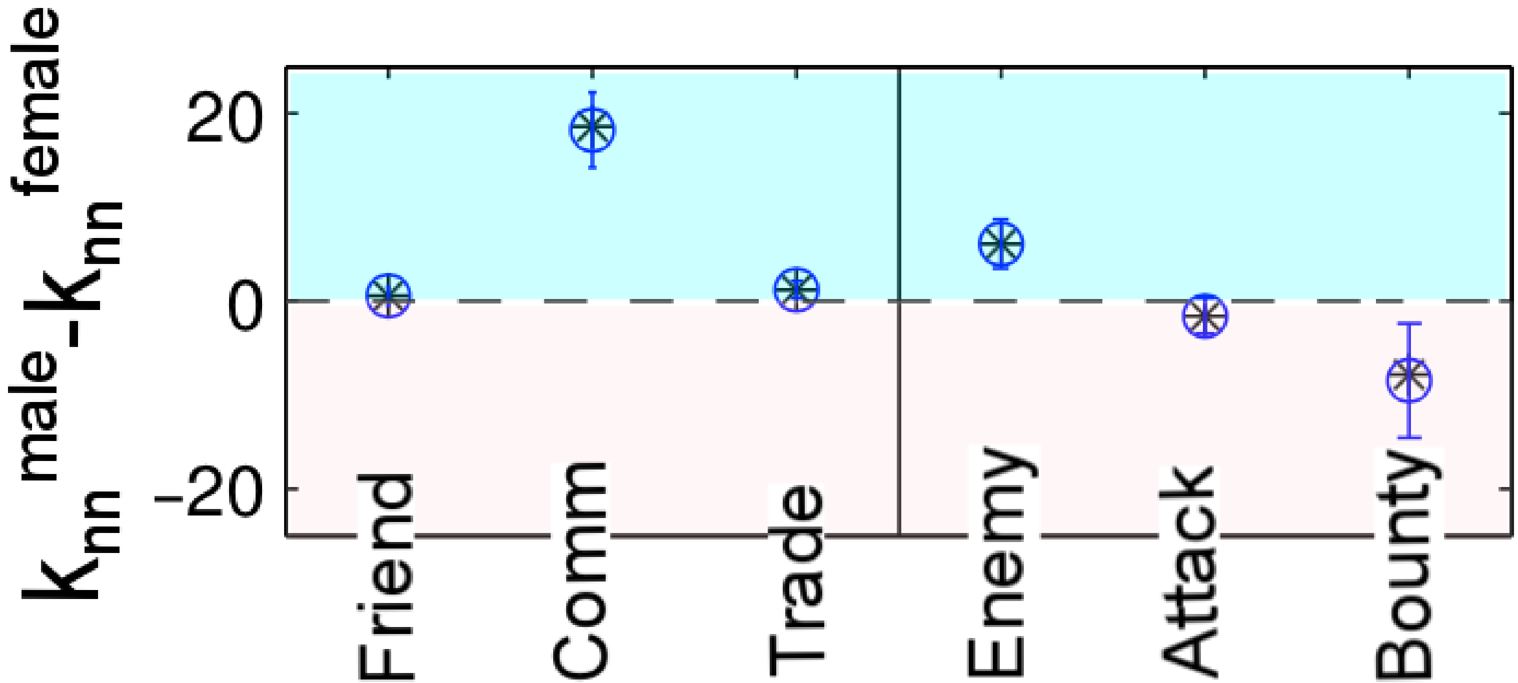
→ **Females have 5 more communication partners, on average**

Gender differences: clustering



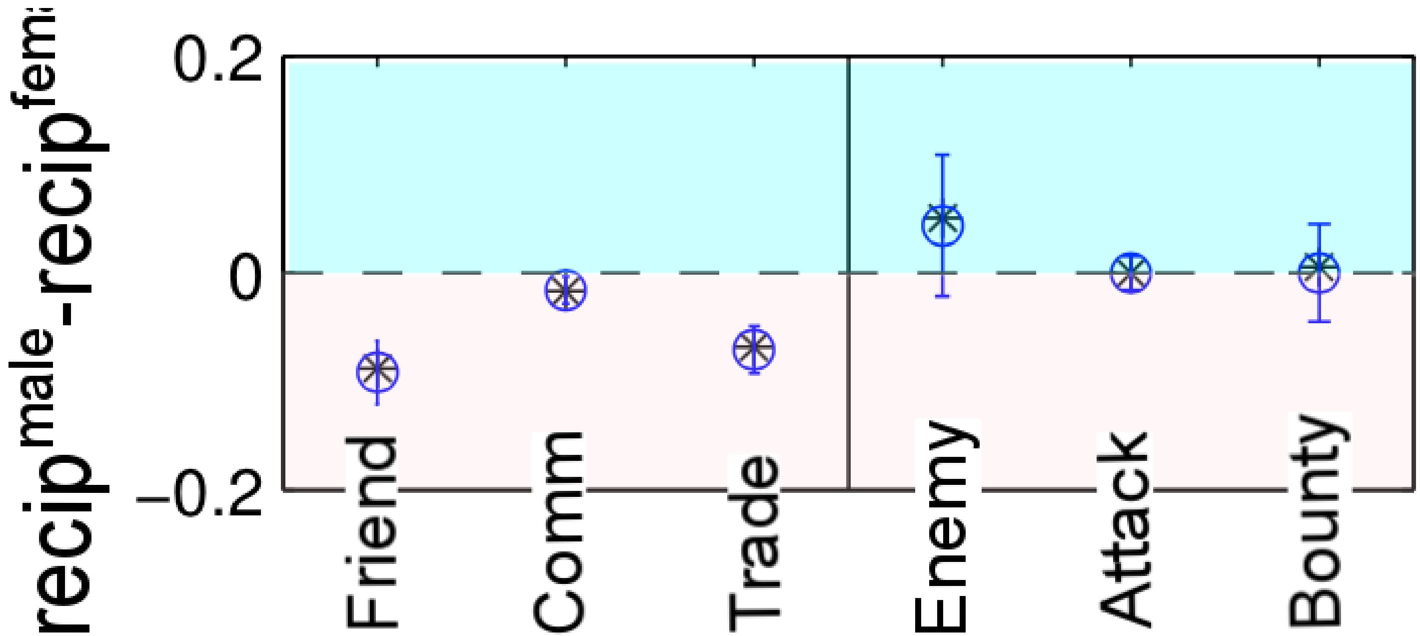
→ **Females build much more triangles than males**

Gender differences: nearest-neighbor degree



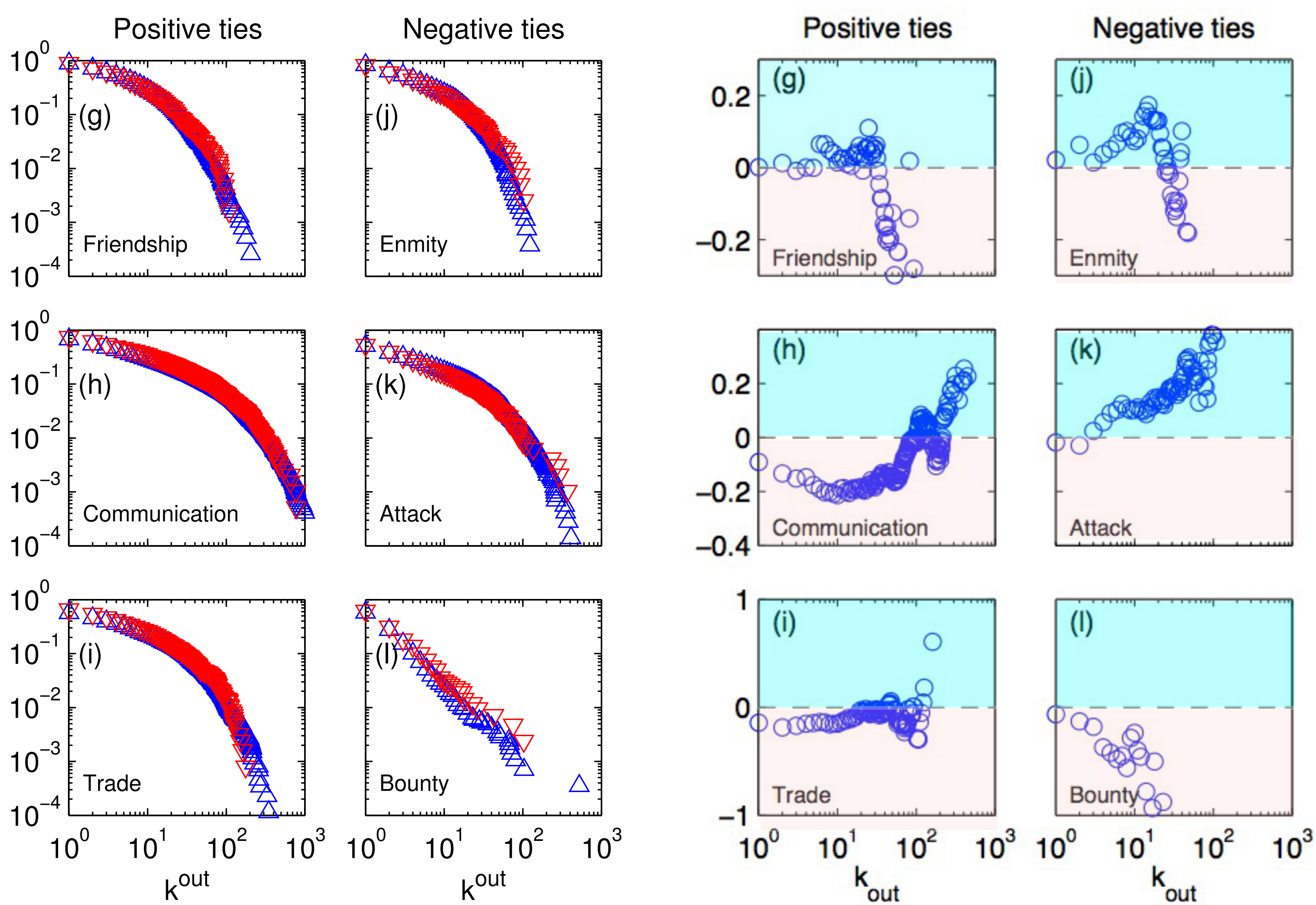
→ **Males link to well-connected communicators**

Gender differences: reciprocation

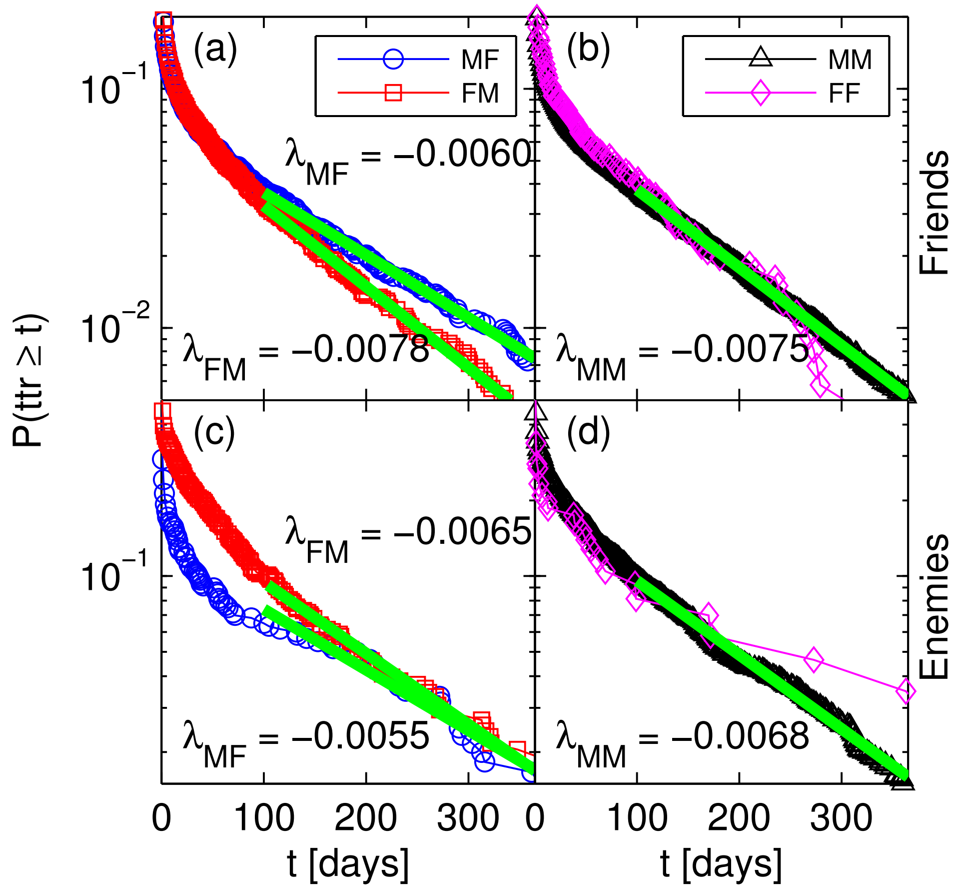


→ **Females reciprocate**

Gender differences in networking



Temporal behavioral gender differences



Another physical social (behavioral) law?

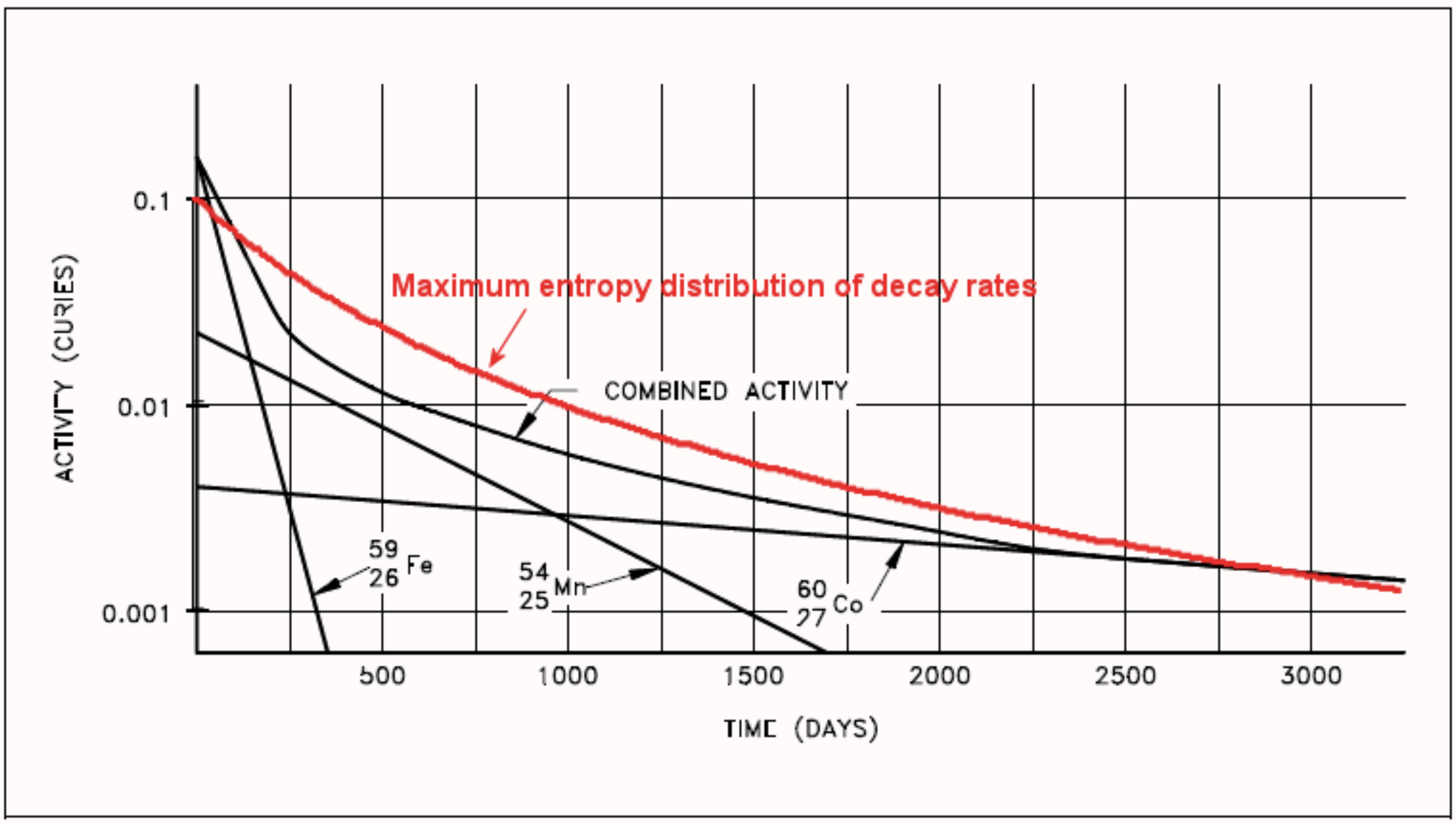
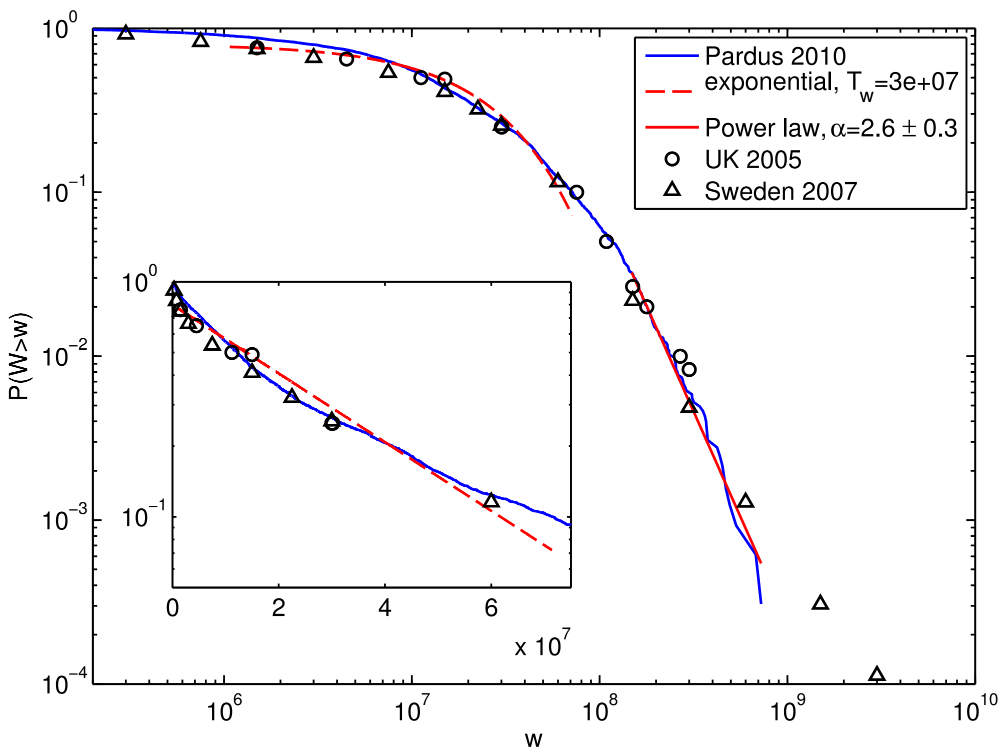


Figure 12 Combined Decay of Iron-56, Manganese-54, and Cobalt-60

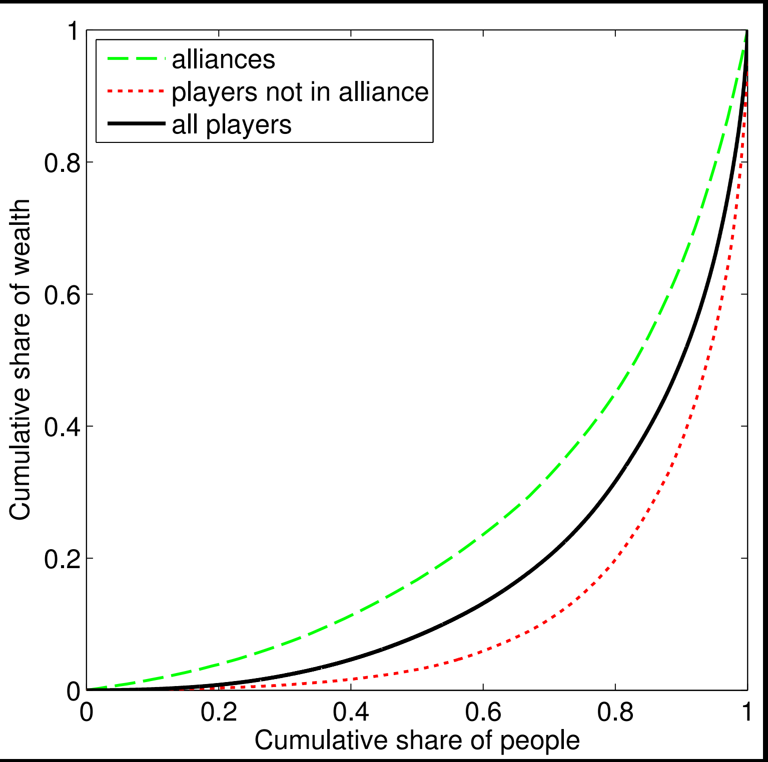
V Wealth of virtual nations

Can we understand the origin of wealth distributions?

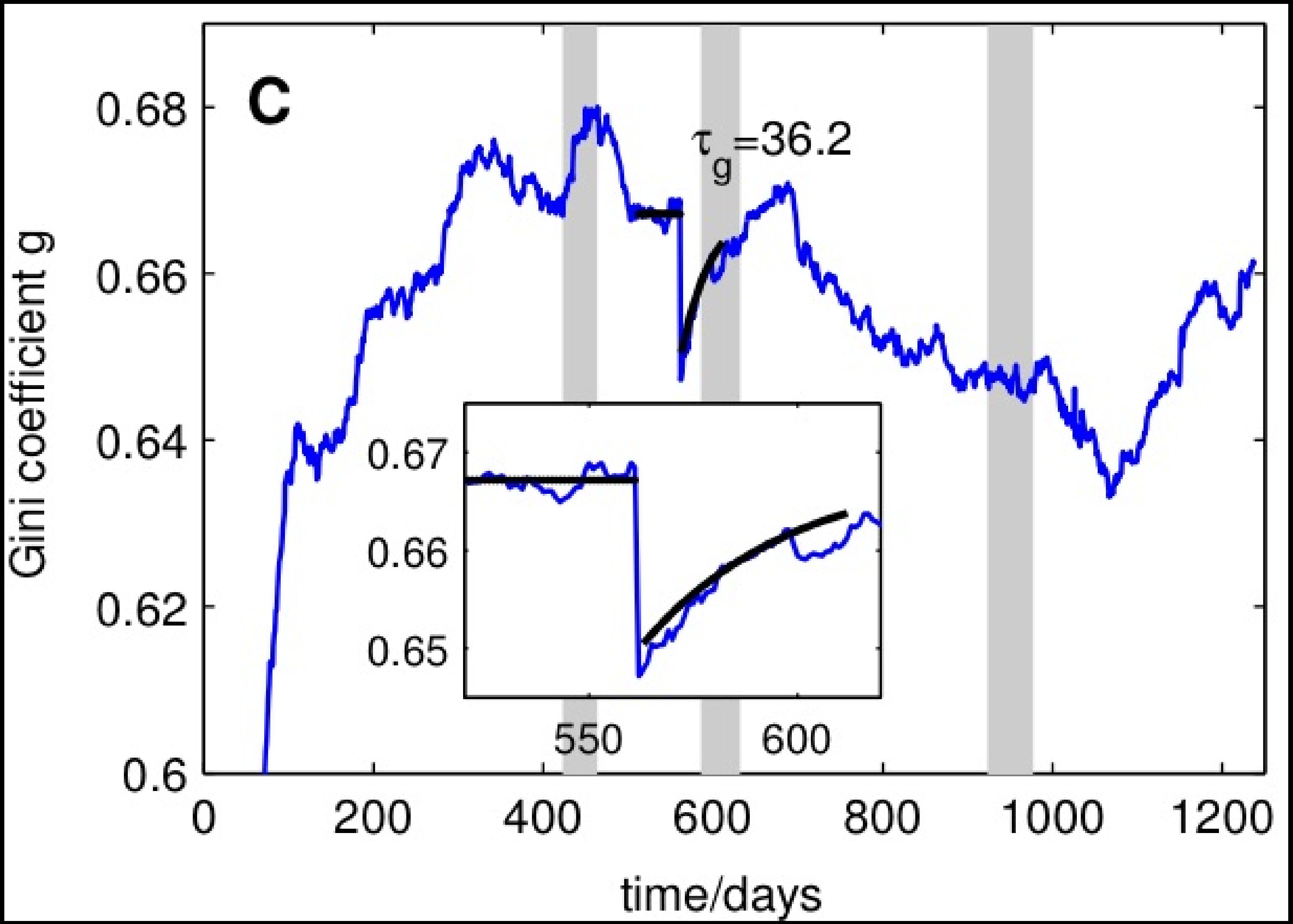
wealth distribution



Gini-curve



Wealth is not luck – it is structural



yanisvaroufakis.eu

DIGITAL ECONOMIES: Markets, Money and Democratic Politics Revisited | Yanis Varoufakis

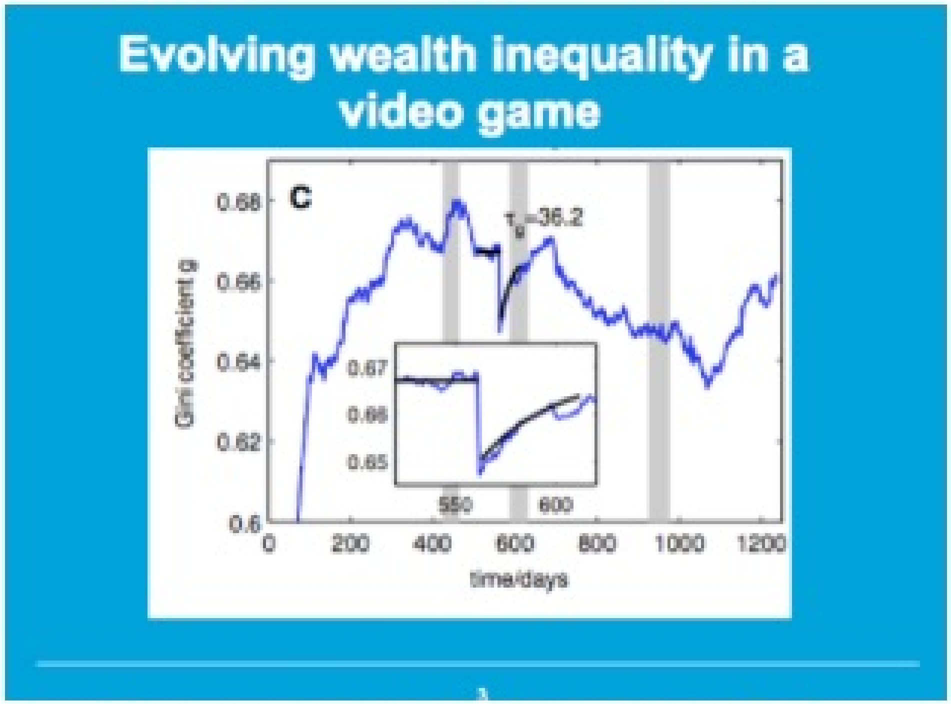


Diagram courtesy of Fuchs and Stefan Thurner (2014).

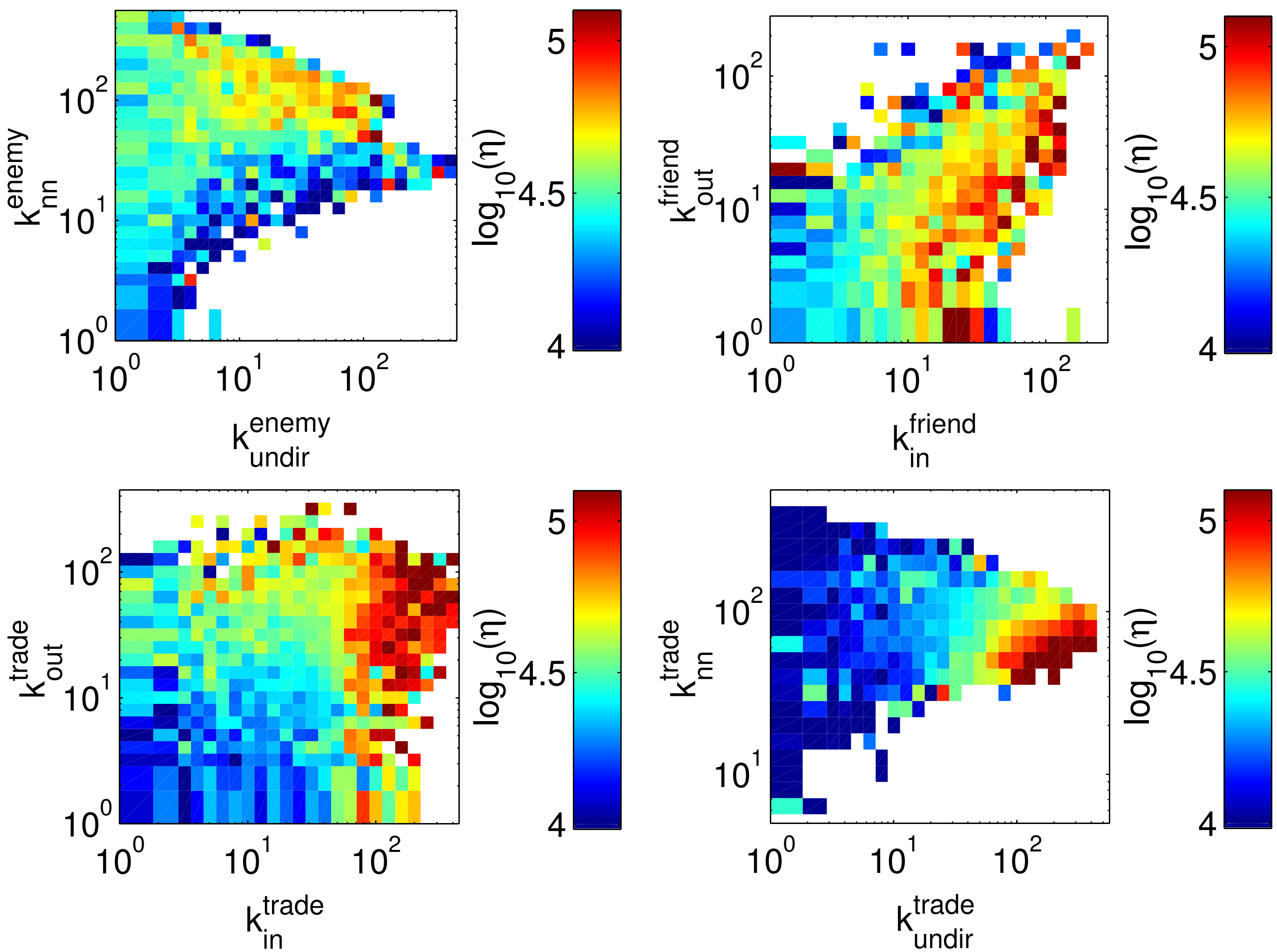
It is also of interest to juxtapose this video game community's wealth inequality data against data derived from Sweden and the UK, which as

Follow

MEDICAL UNIVERSITY OF VIENNA
SANTA FE INSTITUTE
IIASA
COMPLEXITY SCIENCE HUB VIENNA

MEDICAL UNIVERSITY OF VIENNA SANTA FE INSTITUTE IIASA COMPLEXITY SCIENCE HUB VIENNA

Wealth depends on position within multiplex network



What have we learned about the homo sapiens?

- Strength of human interactions measurable
- Humans are triangle-closers
- Humans organize in stable signed triangles
- Humans organize in groups that are (roughly) multiples of four
- Males and females organize their local social networks very different
- Females and males handle aggression differently
- Humans boost aggression levels when treated badly
- Good and bad actions lead to different organization of networks
- Wealth is function of your position in the social multiplex