

Life Data Epidemiology

Lect 20: Complex contagion

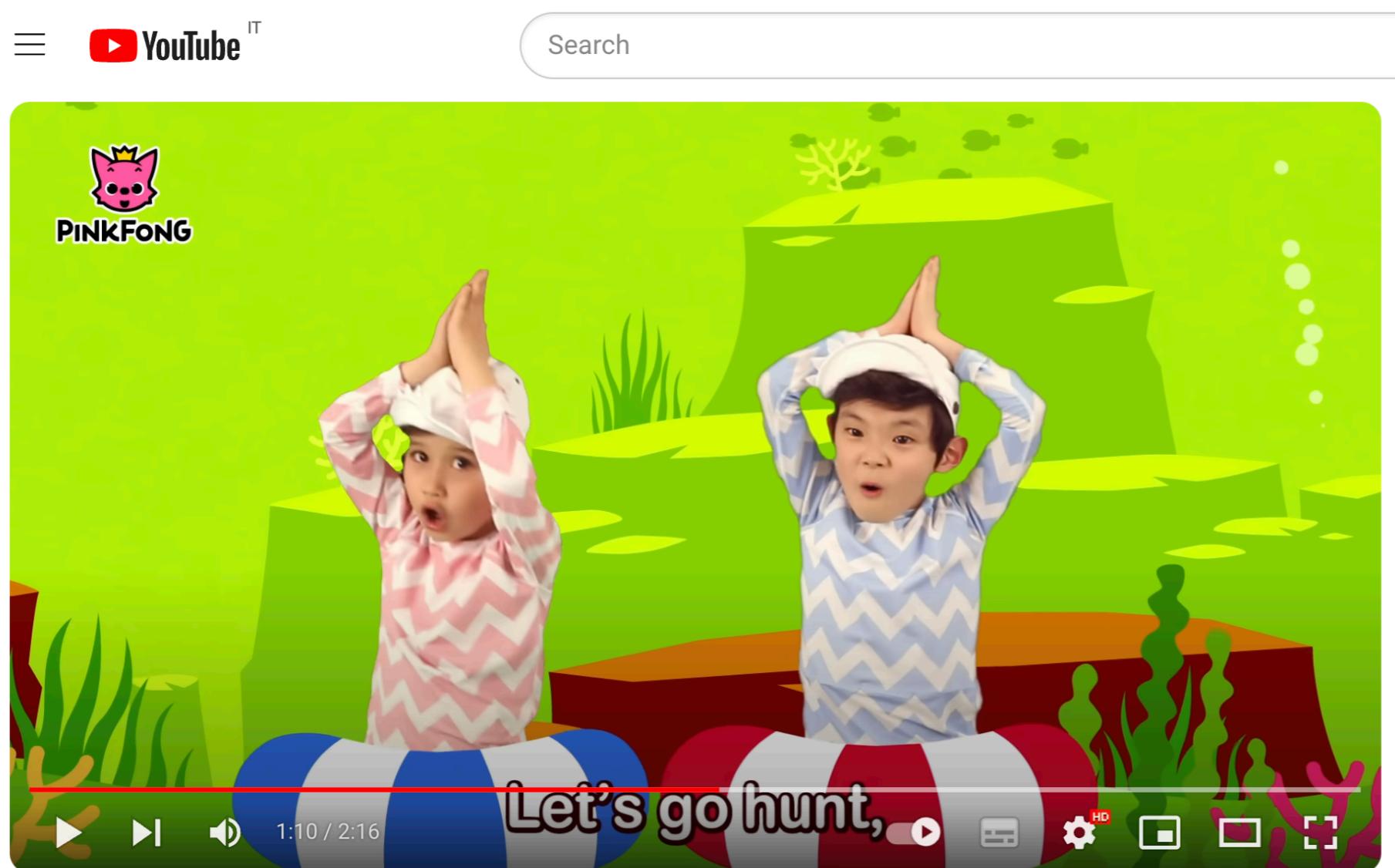
Chiara Poletto

mail: chiara.poletto@unipd.it

web: chiara-poletto.github.io

bsky: @chpoletto.bsky.social

going viral ...



Baby Shark Dance | #babyshark Most Viewed Video | Animal Songs | PINKFONG Songs for Children



13,903,196,466 views 7 years ago

★ Pinkfong Sing-Along Movie 3: Catch the Gingerbread Man - In Cinemark September 9 (USA Only) ★

information spread as an epidemic ...

or better as an explosive epidemic!

Other contagion processes

- Social Norms
- Economic Crisis
- Viral Information
- Social movements



(Grace HeeJung Kim)

Spreading of information

Epidemic models (and adaptations) are often used to study information diffusion in social systems

Important for

- Fake news
- Echo chambers
- Infodemics



(Courtesy: Claus Lunau/Science Photo Library)

Spreading of information: differences with epidemic spreading

However, in social contexts things are a bit more complex:

- Information transmission is an intentional act for both sender and receiver
- Often beneficial for senders and receivers (e.g. reinforcement)
- Influenced by cognitive and psychological factors
- Content of information matters (e.g. homophily)

Complex contagion

The spreading of information/innovation is caused by a different kind of contagion:

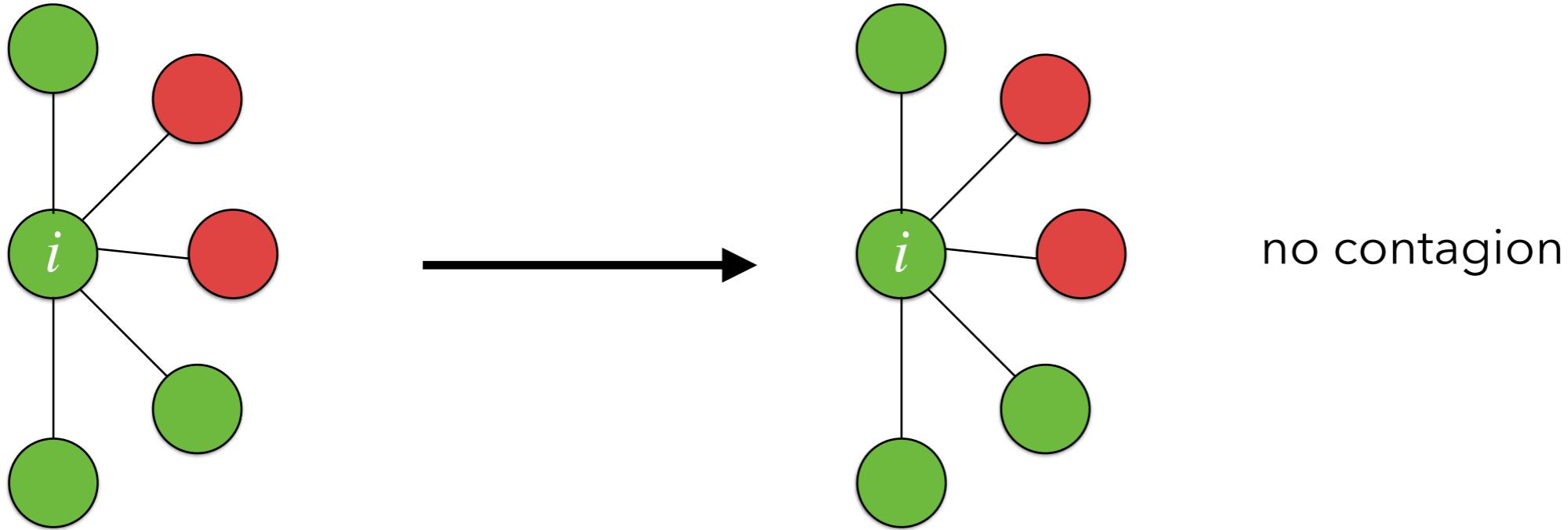
It involves multiple exposures and reinforcement, i.e. **memory of past interactions**

e.g. threshold models: "If half of my friends buy a given product, I will likely buy it too"

This is named **complex contagion** in contraposition with **simple contagion** where interactions are independent

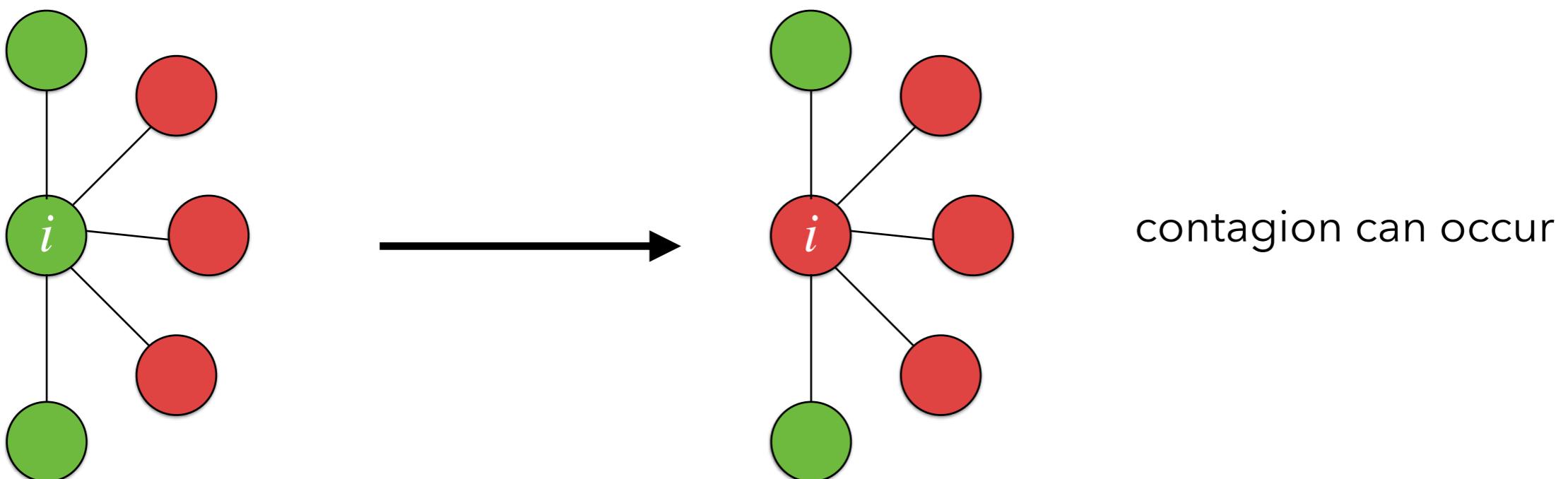
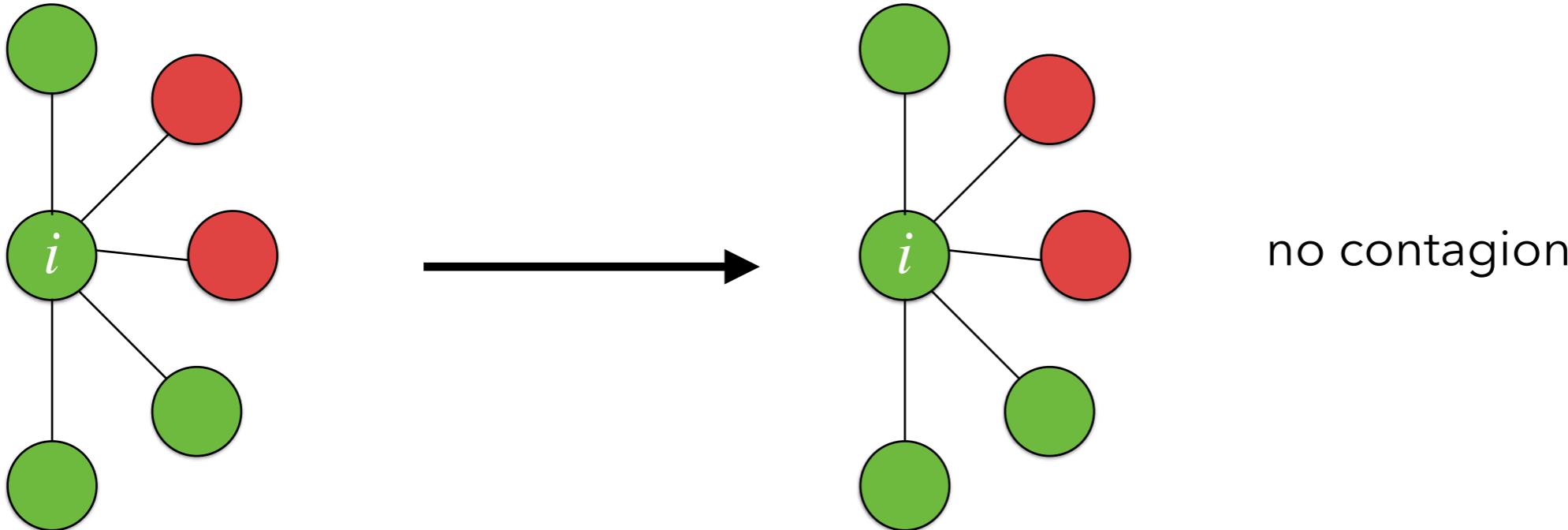
Threshold model

"If half of my friends buy a product, I will likely buy it too."



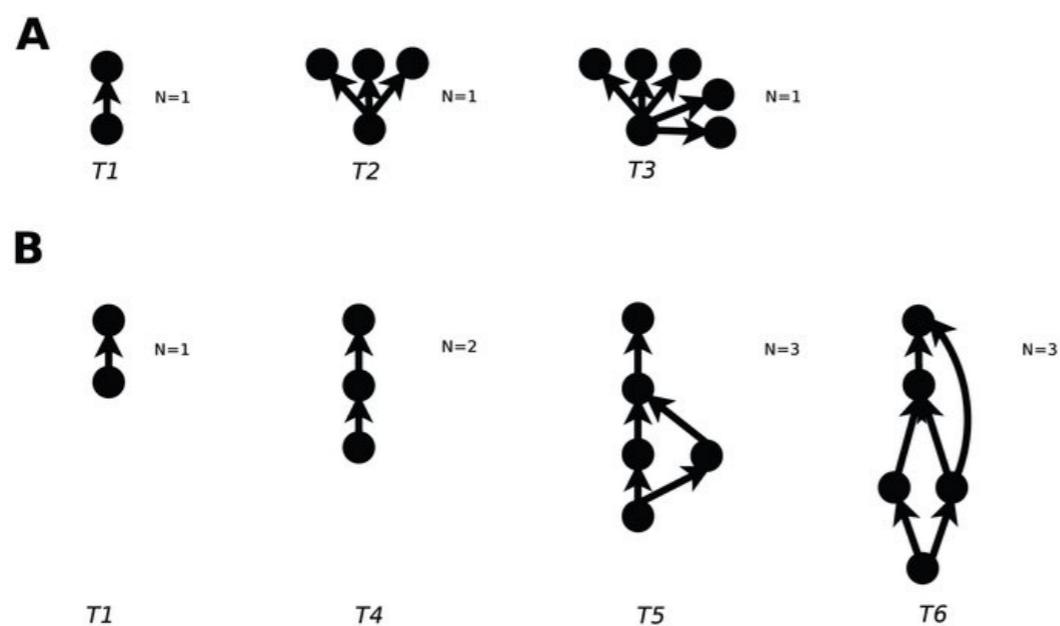
Threshold model

"If half of my friends buy an iPhone, I will likely buy it too."



Threshold model

Lead to **information cascades**:



[Vaquero, Luis & Cebrian, Manuel. The Rich Club Phenomenon in the Classroom. Scientific reports. 3. 1174. 10.1038/srep01174. (2013)]

generalised complex contagion model

a generalised contagion model able to reproduce both simple and complex contagion:

- Well-mixed populations
- S, I, R
- Memory of past interactions up to time T
- From each successful interaction with an infected j , a susceptible i gets a “dose” of infection $d_i(t)$
- If the accumulated dose $D_i(t) = \sum_{t'=t-T+1}^t d_i(t')$ exceeds a threshold d_i^* , then the susceptible becomes infected

generalised complex contagion model

contagion process:

- At each time step t : each individual i contacts a random individual j
 - If $i = S$ and $j = I$, with probability p individual i gets a "dose" of infection $d_i(t)$ from a dose size distribution $f(d)$
 - Otherwise $d_i(t) = 0$
- Each individual keeps a "cumulative" dose of the T previous time steps:
$$D_i(t) = \sum_{t'=t-T+1}^t d_i(t')$$
- If $D_i(t)$ is larger than the individual threshold d_i^* , individual i gets infected

generalised complex contagion model

recovery process:

If $D_i(t)$ gets below d_i^* , i recovers with probability r

It also possible to add an $R \rightarrow S$ transition with probability r' to simulate an SIRS model

- With $r = 1$ and $r' = 1$ it recovers an SIS-like dynamics

generalised complex contagion model

parameters

- p and r infection and recovery probability (same role as β and μ)
- $d_i(t)$ “dose” per infection distributed following $f(d)$
- d_i^* threshold, distributed following $g(d^*)$
- $f(d)$ and $g(d^*)$ can be general distributions
- With specific choices of p , $f(d)$ and $g(d^*)$, it is possible to recover both pure epidemic and threshold dynamics.

generalised complex contagion model

Let's define $P_{inf}(K)$ the probability that an individual with $K < T$ contact gets infected as:

$$P_{inf}(K) = \sum_{k=1}^K \binom{K}{k} p^k (1-p)^{K-k} P_k$$

With P_k the average fraction of infected after receiving k doses in T time steps

$$P_k = \int_0^\infty dd^* g(d^*) P\left(\sum_{i=1}^k d_i \geq d^*\right)$$

And

$$P\left(\sum_{i=1}^k d_i \geq d^*\right) \text{ the probability that } k \text{ doses exceed } d^*$$

generalised complex contagion model

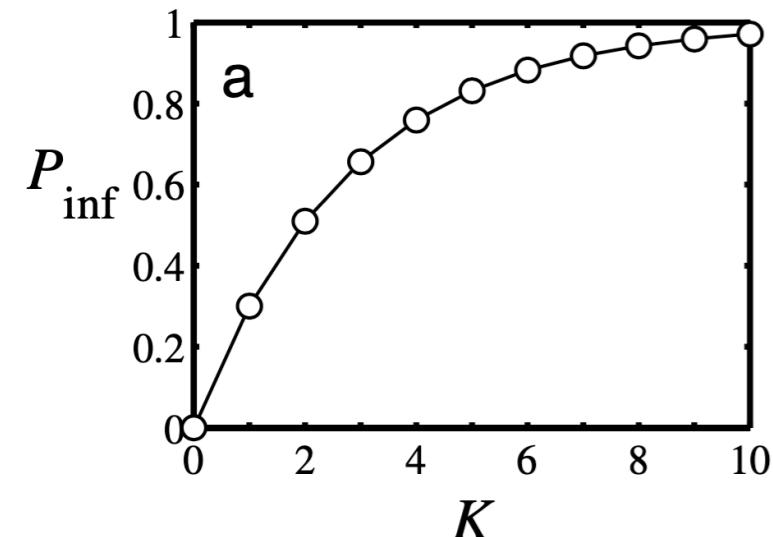
The model can be solved numerically for any distribution of $f(d)$ and $g(d^*)$

Specific cases:

If $p < 1$ and fixed dose size $f(d) = \delta(d - 1)$ and fixed threshold $g(d^*) = \delta(d^* - 1)$

Epidemic spreading (independent interactions)

- All contacts have the same infection probability
- The threshold is $d^* = 1$ (one successful contact)



generalised complex contagion model

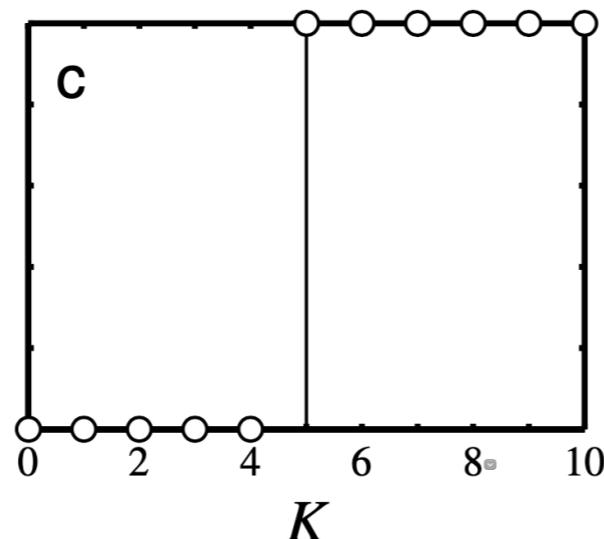
The model can be solved numerically for any distribution of $f(d)$ and $g(d^*)$

Specific cases:

If $p = 1$ and fixed dose size $f(d) = \delta(d - 1)$ and fixed threshold $g(d^*) = \delta(d^* - 5)$

Deterministic threshold model

- Threshold is fixed $d^* = 5$
- Needs at least 5 encounters to get infected



generalised complex contagion model

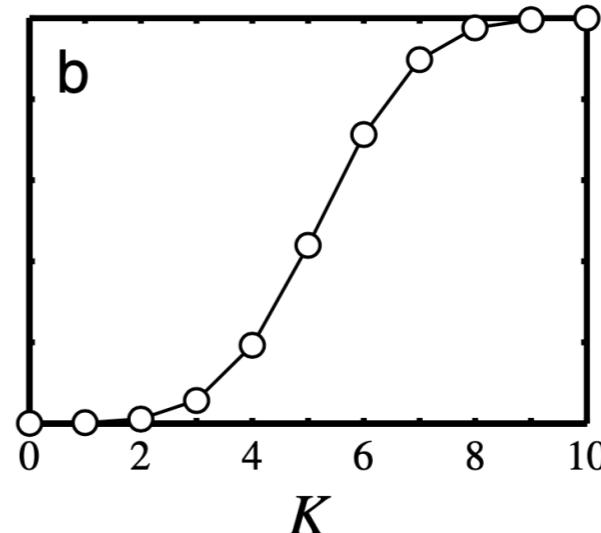
The model can be solved numerically for any distribution of $f(d)$ and $g(d^*)$

Specific cases:

If $p = 1$ and dose size $f(d)$ is distributed log-normally and fixed threshold
 $g(d^*) = \delta(d^* - 5)$

Stochastic threshold model

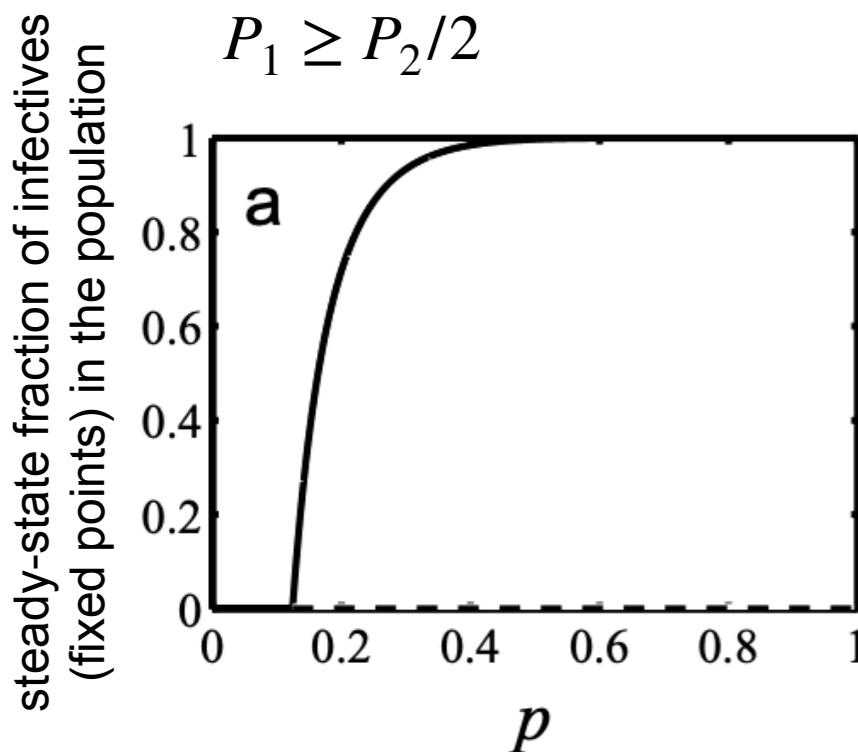
- Threshold is fixed $d^* = 5$
- But the “dose” for each contact varies.



generalised complex contagion model

Possible outcomes depend on P_1 and P_2 , the probability that an individual will become infected as a result of one and two exposures respectively

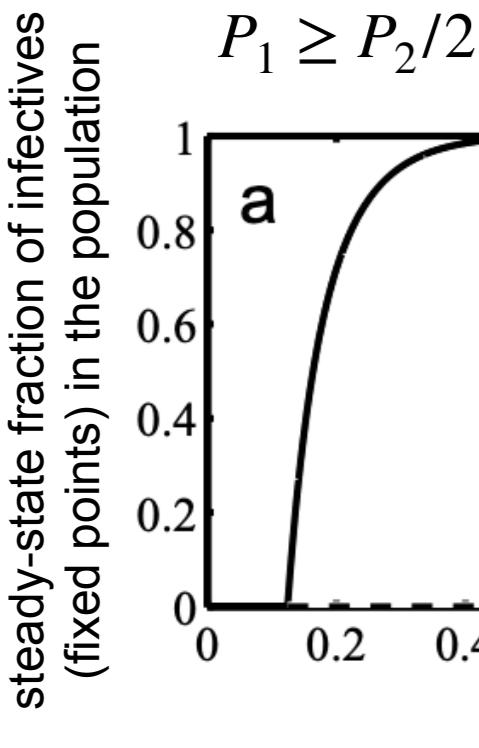
epidemic threshold model



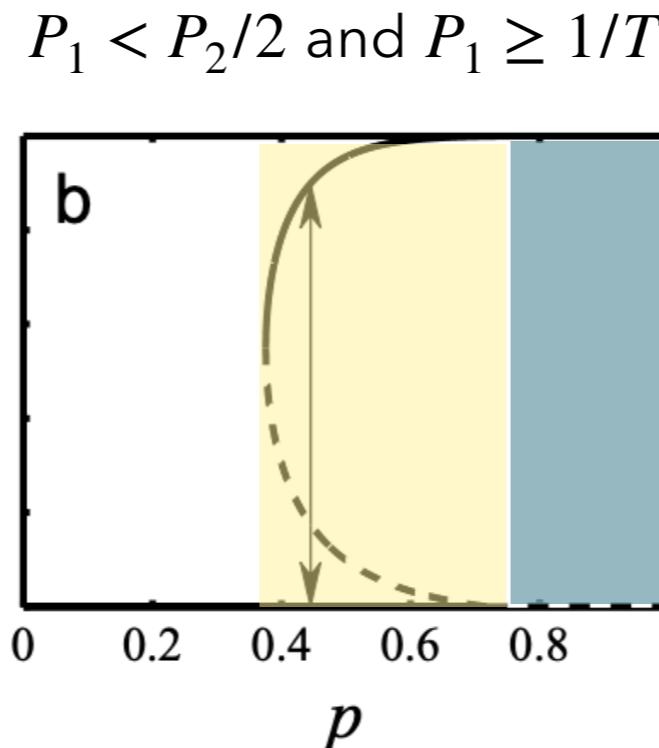
generalised complex contagion model

Possible outcomes depend on P_1 and P_2 , the probability that an individual will become infected as a result of one and two exposures respectively

epidemic threshold model

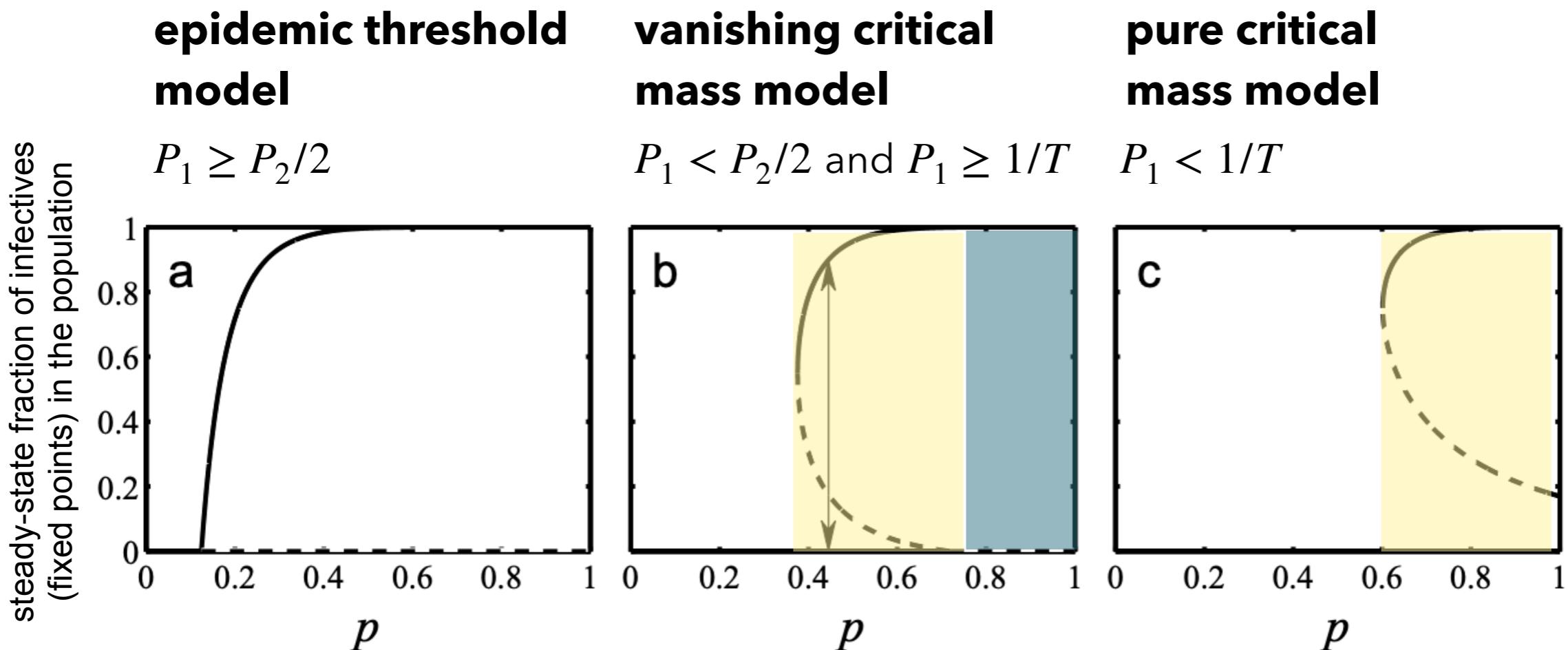


vanishing critical mass model

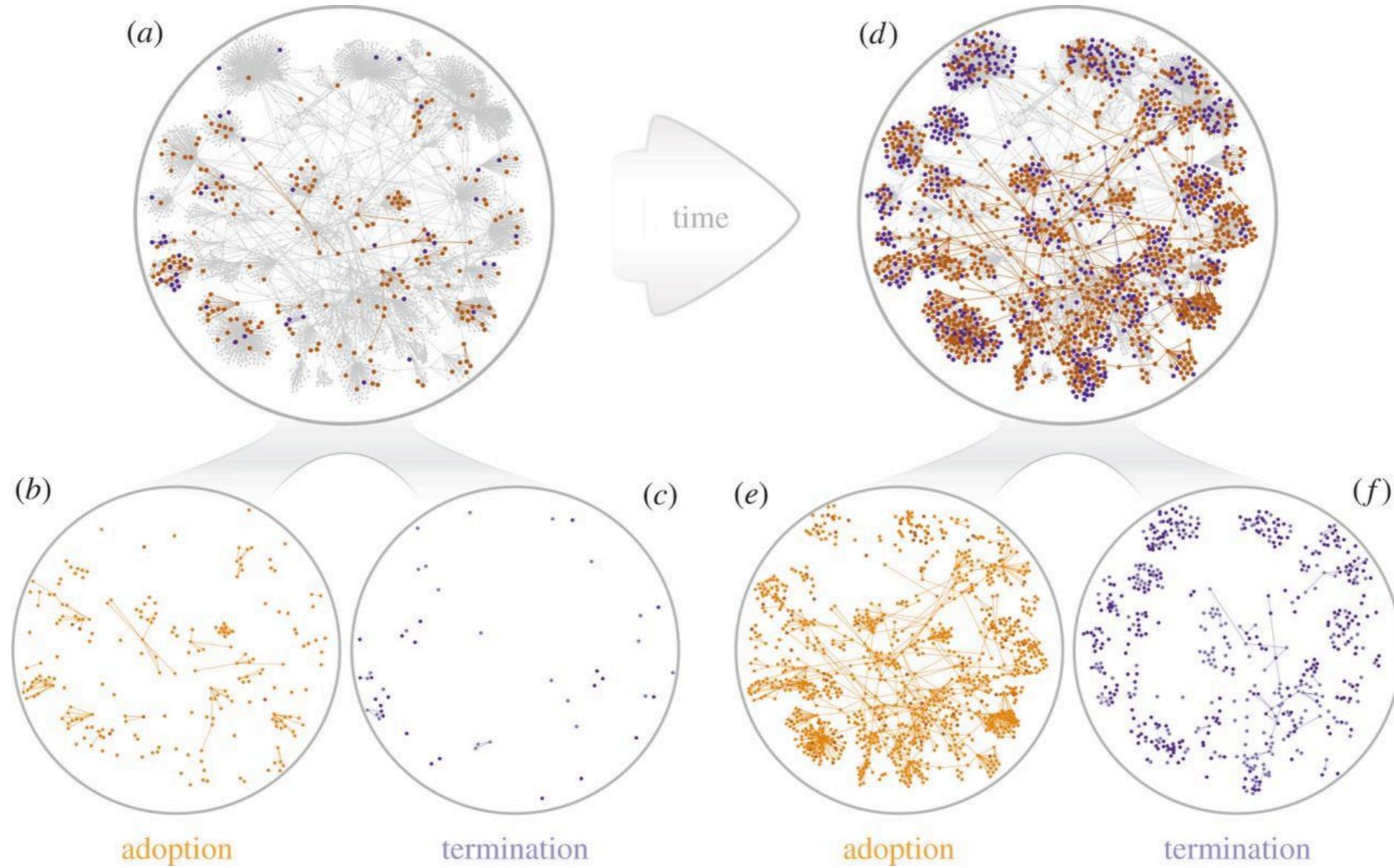


generalised complex contagion model

possible outcomes depend on P_1 and P_2 , the probability that an individual will become infected as a result of one and two exposures respectively



Empirical evidences

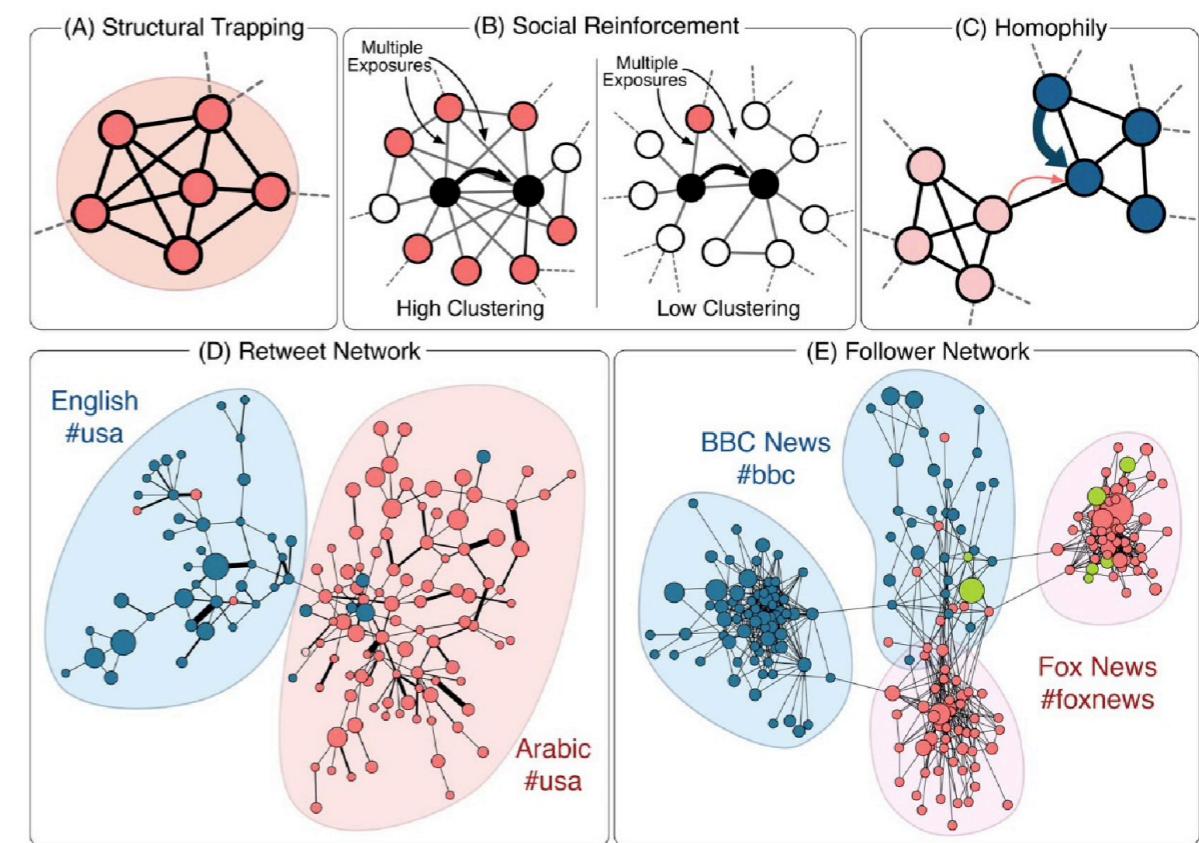


analysis of a dataset recording the spreading of Skype service to empirically support the assumptions behind models of social contagion. The probability of adoption via social influence depends on the fraction of adopting neighbours.

Empirical evidences

Question: how hashtags (memes) spread in online social networks?

- Simple vs. Complex Contagion
- Example: Twitter
- Studying how retweets (RT) and mentions (@) diffuse hashtags in different communities
- Effect of communities: reinforcement, homophily, etc.
- Differences between “Popular” and “Normal” hashtags



[Weng, L., Menczer, F. & Ahn, Y. Virality Prediction and Community Structure in Social Networks. Sci. Rep. 3, 2522; (2013)]

Information spreading in online social networks

Weights to quantify the fraction of information that flows inside a community and outside:

$\langle w_{\circlearrowleft} \rangle_c$ average weight (number of tweets) per link inside the community

$\langle w_{\nwarrow} \rangle_c$ average weight (number of tweets) per link outside the community

Users activity:

f_{\circlearrowleft} : focus of an individual, i.e. fraction of activity inside the community

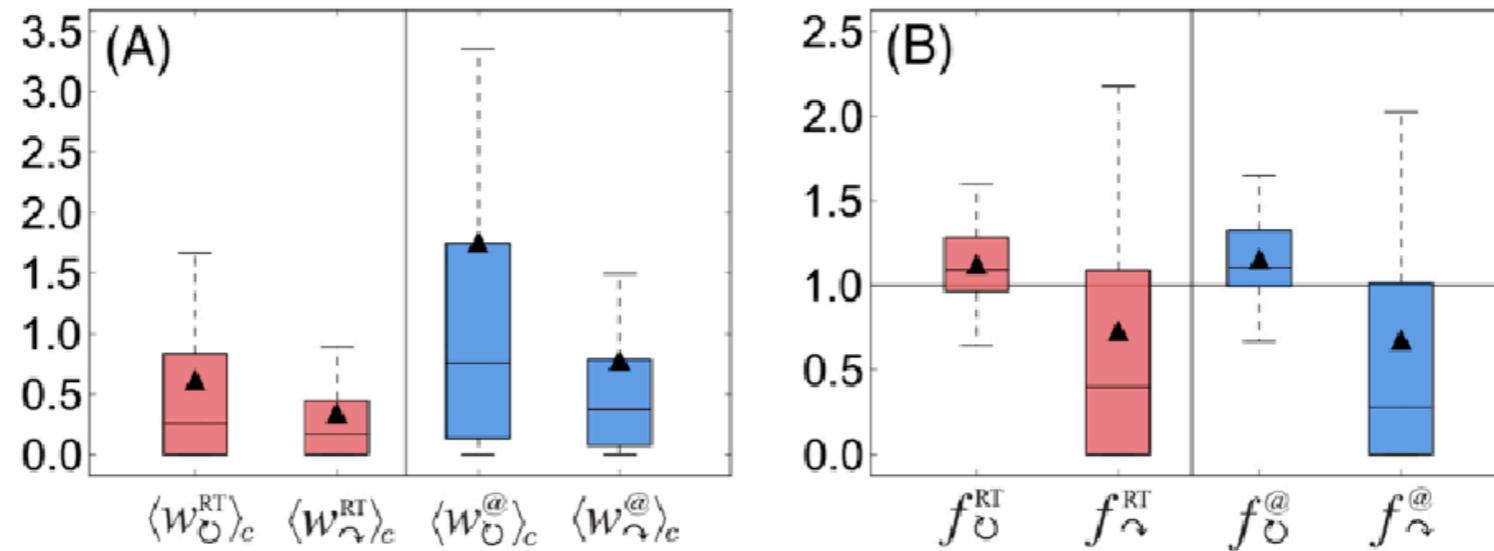
f_{\nwarrow} : focus of an individual, i.e. fraction of activity outside the community

If simple contagion: there should be no differences in spreading inside and outside a community (e.g. no reinforcement)

Information spreading in online social networks

$\langle w_{\circlearrowleft} \rangle_c$, $\langle w_{\curvearrowright} \rangle_c$ and f_{\circlearrowleft} , f_{\curvearrowright}

It seems that community spreading is favored (effects of homophily and reinforcement)



[Weng, L., Menczer, F. & Ahn, Y. Virality Prediction and Community Structure in Social Networks. Sci. Rep. 3, 2522; (2013)]

Information spreading in online social networks

Hashtag (h) level measures:

Usage dominance $r(h)$: proportion of tweets produced inside the “main” community of h out of the total number of tweets containing h : $T(h)$

Expected low for viral spreading and high for complex contagion

Usage entropy $H(h)$: how h is distributed across communities

High for viral spreading and low for complex contagion

Average exposure $N(h)$: average number of exposures needed to adopt hashtag h

Low for viral spreading and high for complex contagion

Information spreading in online social networks

- Reference models
 - 4 models $M_{1,\dots,4}$ to represent different baseline behavior

Table 1 | Baseline models for information diffusion

Community effects			Simulation implementation
Network	Reinforcement	Homophily	
M_1			For a given hashtag h , M_1 randomly samples the same number of tweets or users as in the real data.
M_2	✓		M_2 takes the network structure into account while neglecting social reinforcement and homophily. M_2 starts with a random seed user. At each step, with probability p , an infected node is randomly selected and one of its neighbors adopts the meme, or with probability $1 - p$, the process restarts from a new seed user ($p = 0.85$).
M_3	✓	✓	The cascade in M_3 is generated similarly to M_2 but at each step the user with the maximum number of infected neighbors adopts the meme.
M_4	✓		In M_4 , the simple cascading process is simulated in the same way as in M_2 but subject to the constraint that at each step, only neighbors in the same community have a chance to adopt the meme.

Average behavior

Simple contagion

Complex contagion

Information spreading in online social networks

$r(h)$, $H(h)$ and $N(h)$ as function of:

Number of tweets T

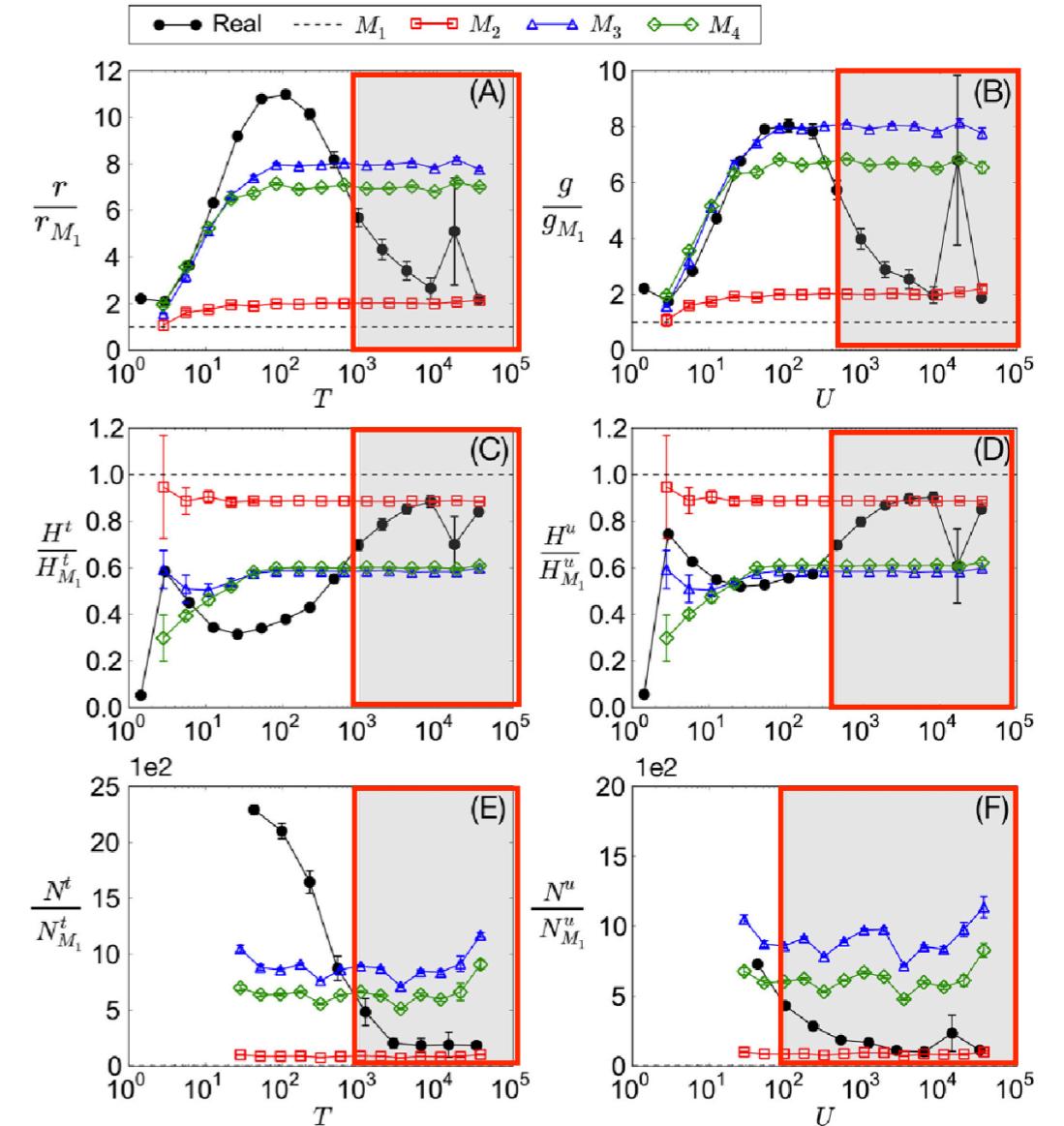
Number of users U

Baseline behavior: model M_1

Two behaviors:

Popular hashtags (large T and U) spread like epidemics (Viral)

Less popular ones follow a complex contagion



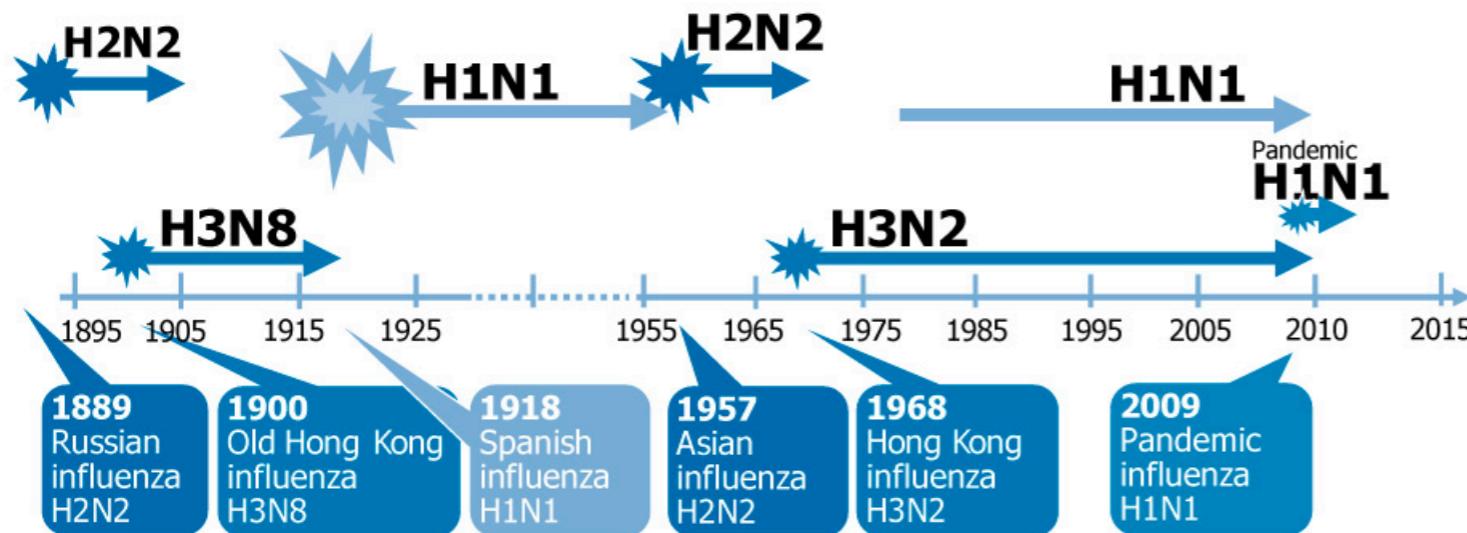
Refs.

- Section X of:
R. Pastor-Satorras, C. Castellano, P. Van Mieghem and A. Vespignani. **Epidemic processes in complex networks**. Rev. Mod. Phys. 87, 925 (2015) (<https://arxiv.org/abs/1408.2701>).
Offers an overview of spreading models for social contagion
- The Generalized Complex Contagion model is presented in:
P. S. Dodds and D.J. Watts. **Universal Behavior in a Generalized Model of Contagion**. Physical Review Letters, 92 218701 (2004) (<http://www.uvm.edu/pdodds/research/papers//years/2004/dodds2004a.pdf>).
- The twitter data analysis is presented in:
L. Weng, F. Menczer and Y. Ahn. **Virality Prediction and Community Structure in Social Networks**. Scientific Reports 3, 2522 (2013) (<https://doi.org/10.1038/srep02522>).

multistrain & influenza

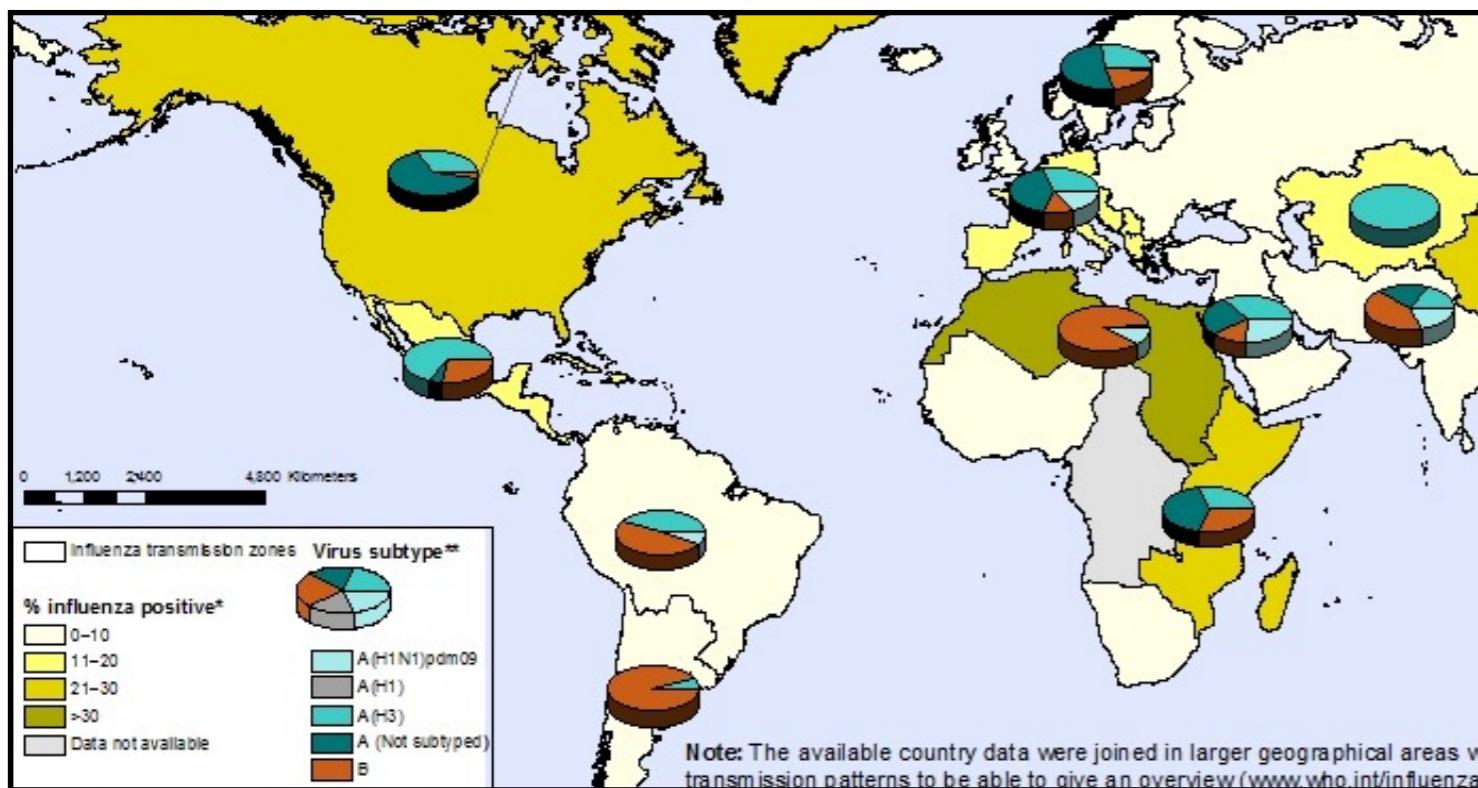
FIGURE

Recorded human pandemic influenzas since 1885 (early sub-types inferred)



Source: European Centre for Disease Prevention and Control (ECDC) 2009

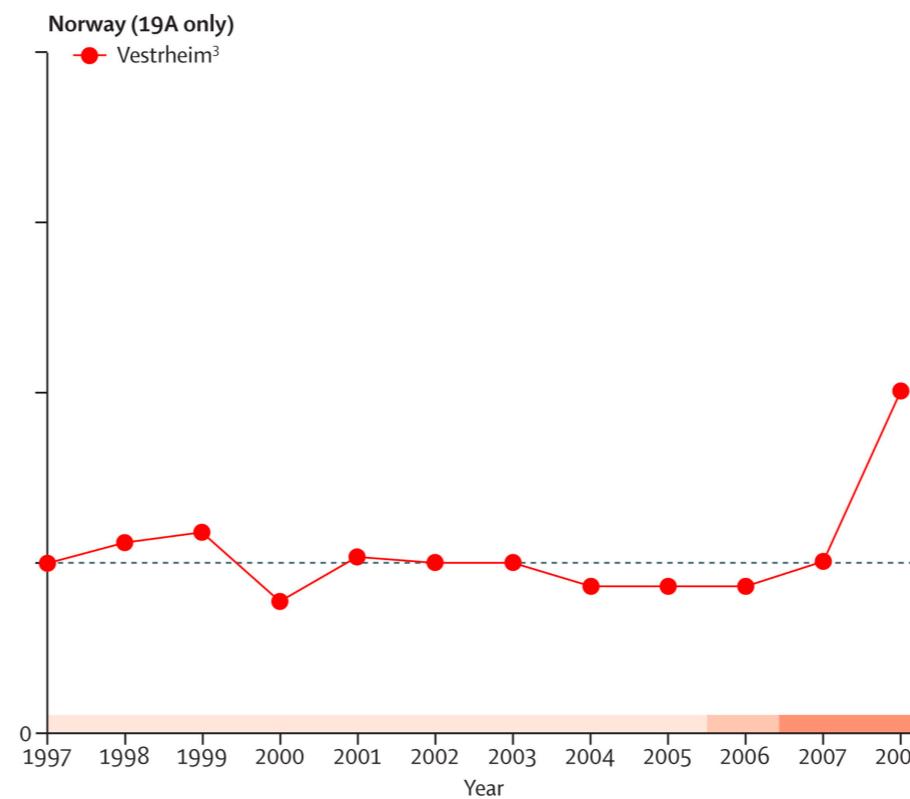
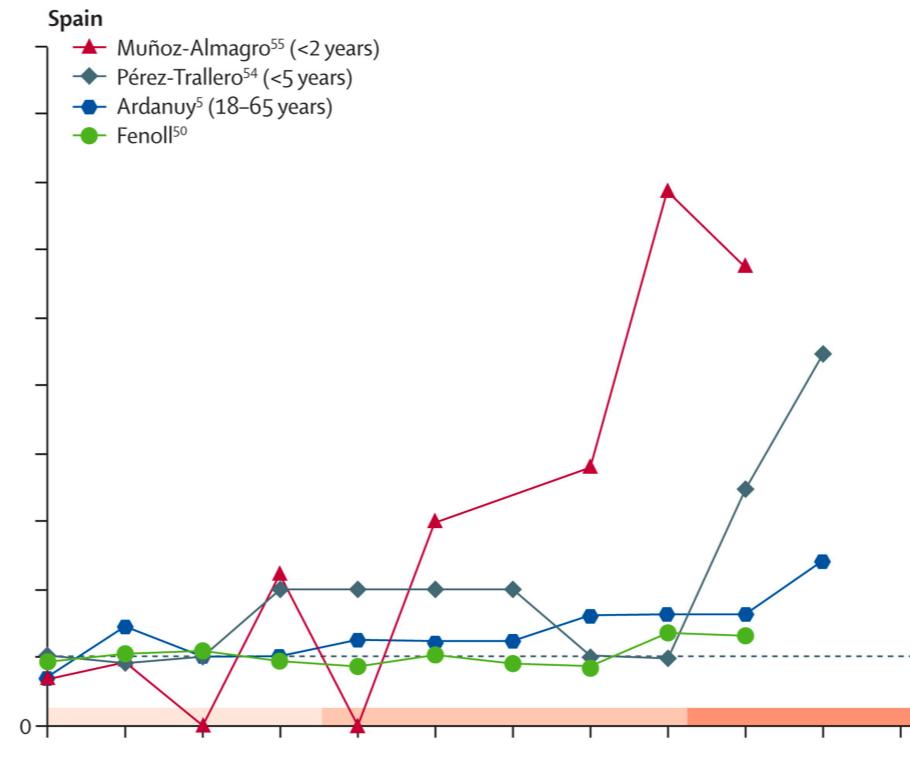
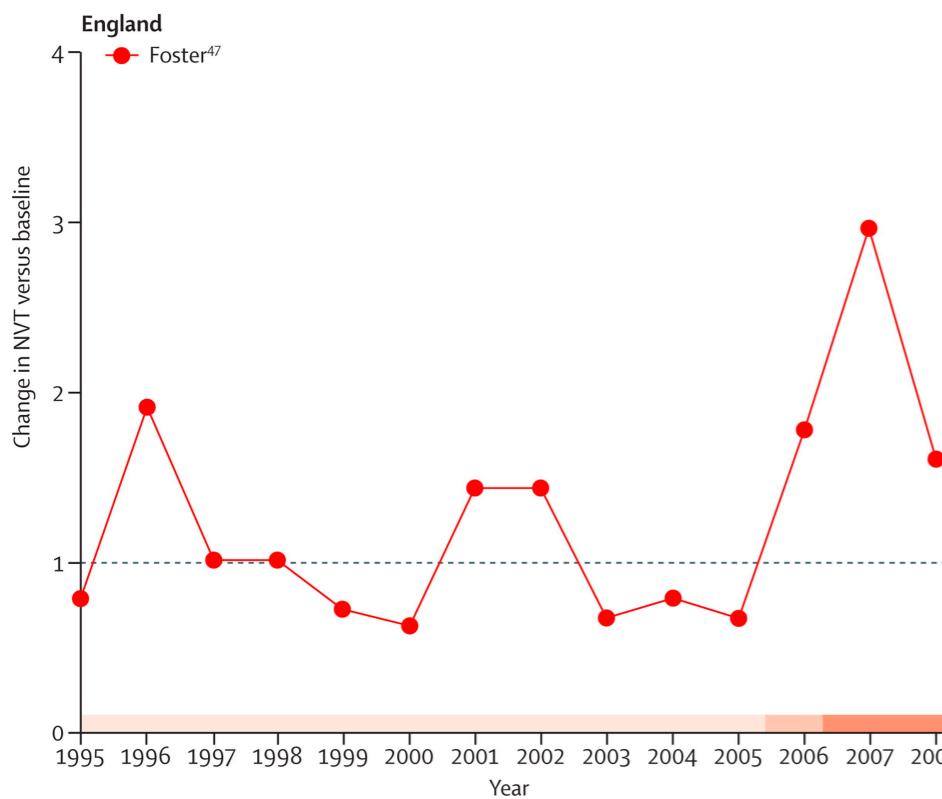
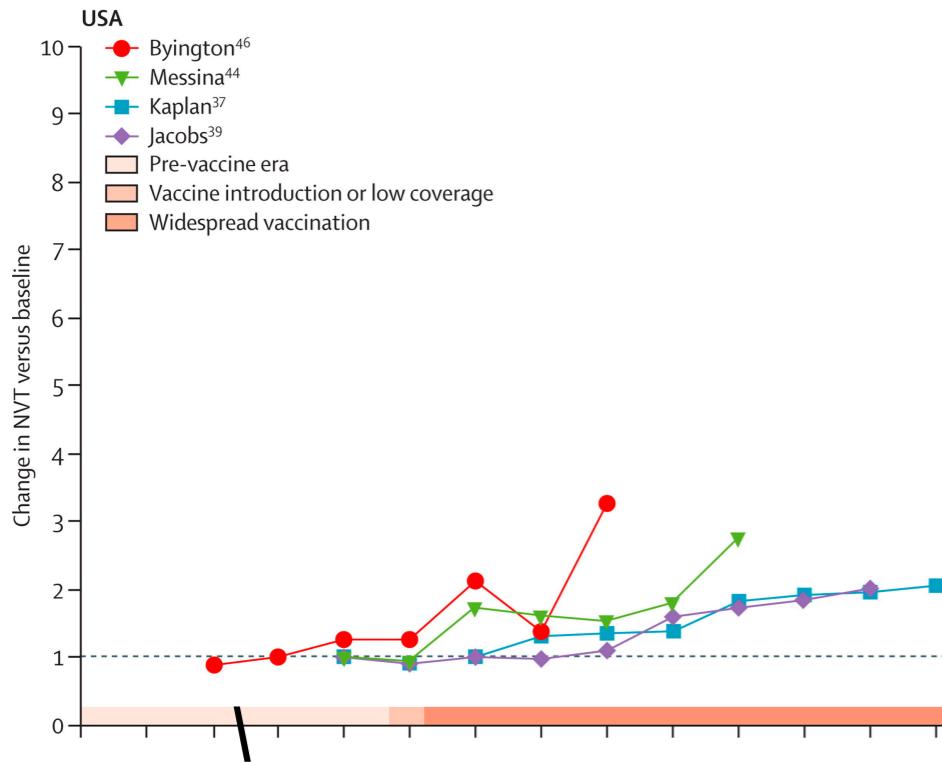
Reproduced and adapted (2009) with permission of Dr Masato Tashiro, Director, Center for Influenza Virus Research, National Institute of Infectious Diseases (NIID), Japan.



- 3 circulating subtypes + continuous mutation
- ~ billion cases annually (~3-5 million severe illness)
- ~300,000-500,000 deaths
- vaccine effectiveness:
23% (in 2014-15)
47% (in 2015-16)

can we increase the impact of vaccination?

multistrain & pneumococcal disease

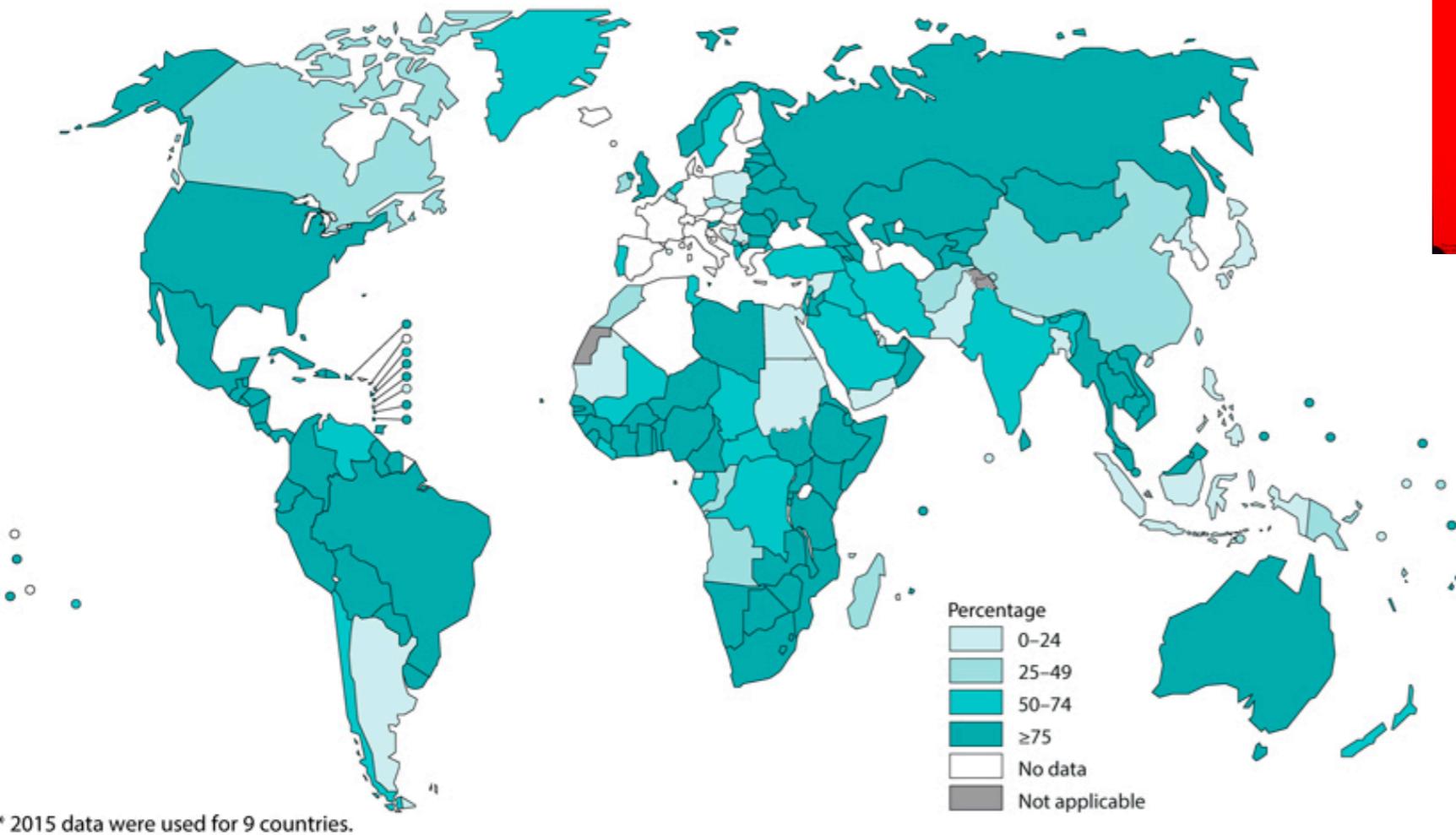


- *S. pneumoniae* major cause of meningitis, sepsis, pneumonia
- ~90 serotypes
- vaccine target few of them

strain replacement?

HIV & TB

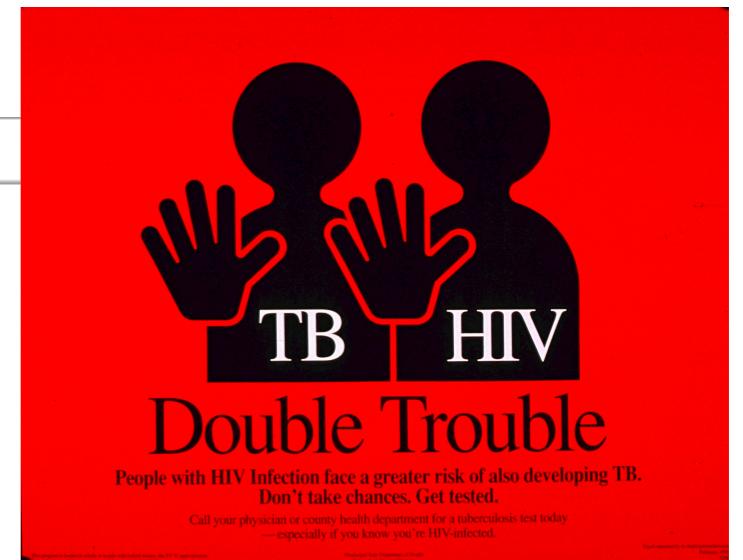
Percentage of new and relapse TB cases with documented HIV status, 2016



any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.
Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

Data Source: Global Tuberculosis Report 2017. WHO, 2017.

© WHO 2017. All rights reserved.



- risk of TB ~16-27 times greater in people living with HIV
- 10.4 million cases of TB disease globally, 1.2 million [11%] among HIV patient

multi strain: fundamental questions

cooperation

EPIDEMIOLOGICAL IMPACT

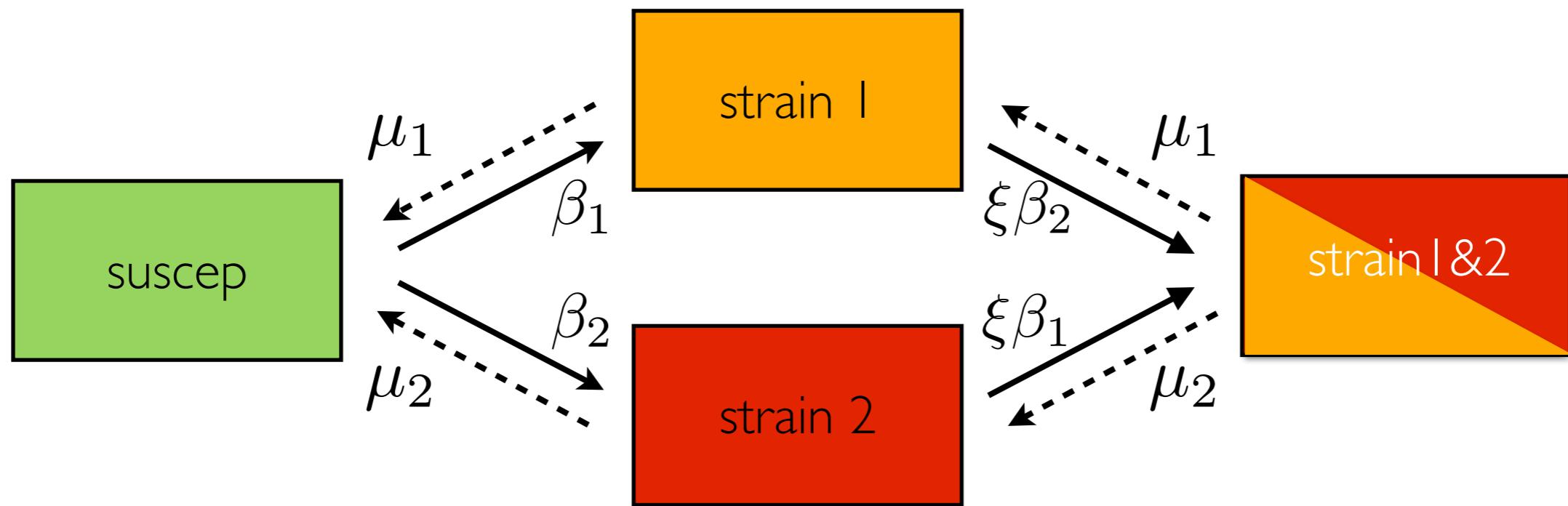


competition

dominance/co-dominance

COEXISTENCE

a simple model with 2 strains

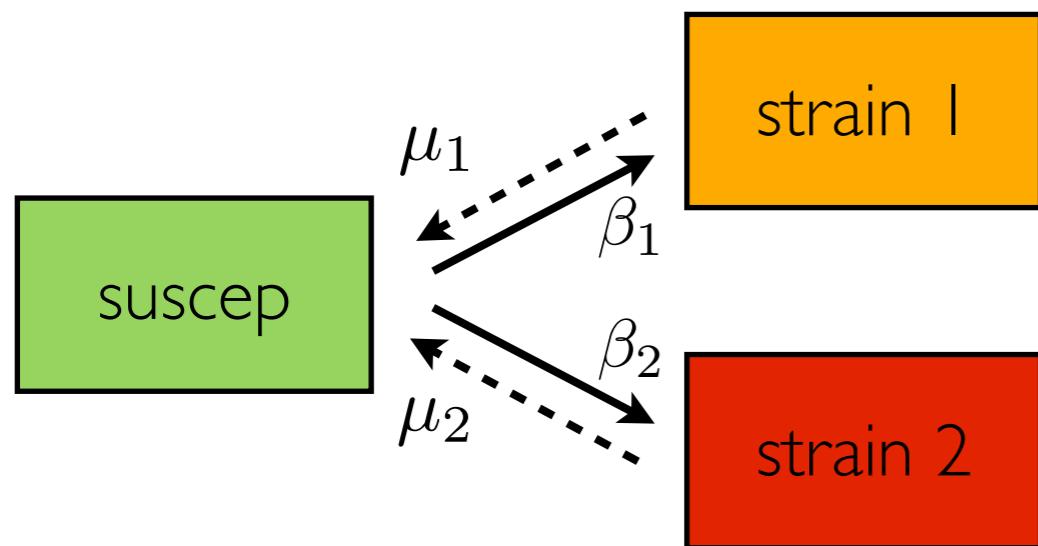


$\xi > 1$ cooperative

$\xi < 1$ competitive

competitive diseases: winner takes it all!

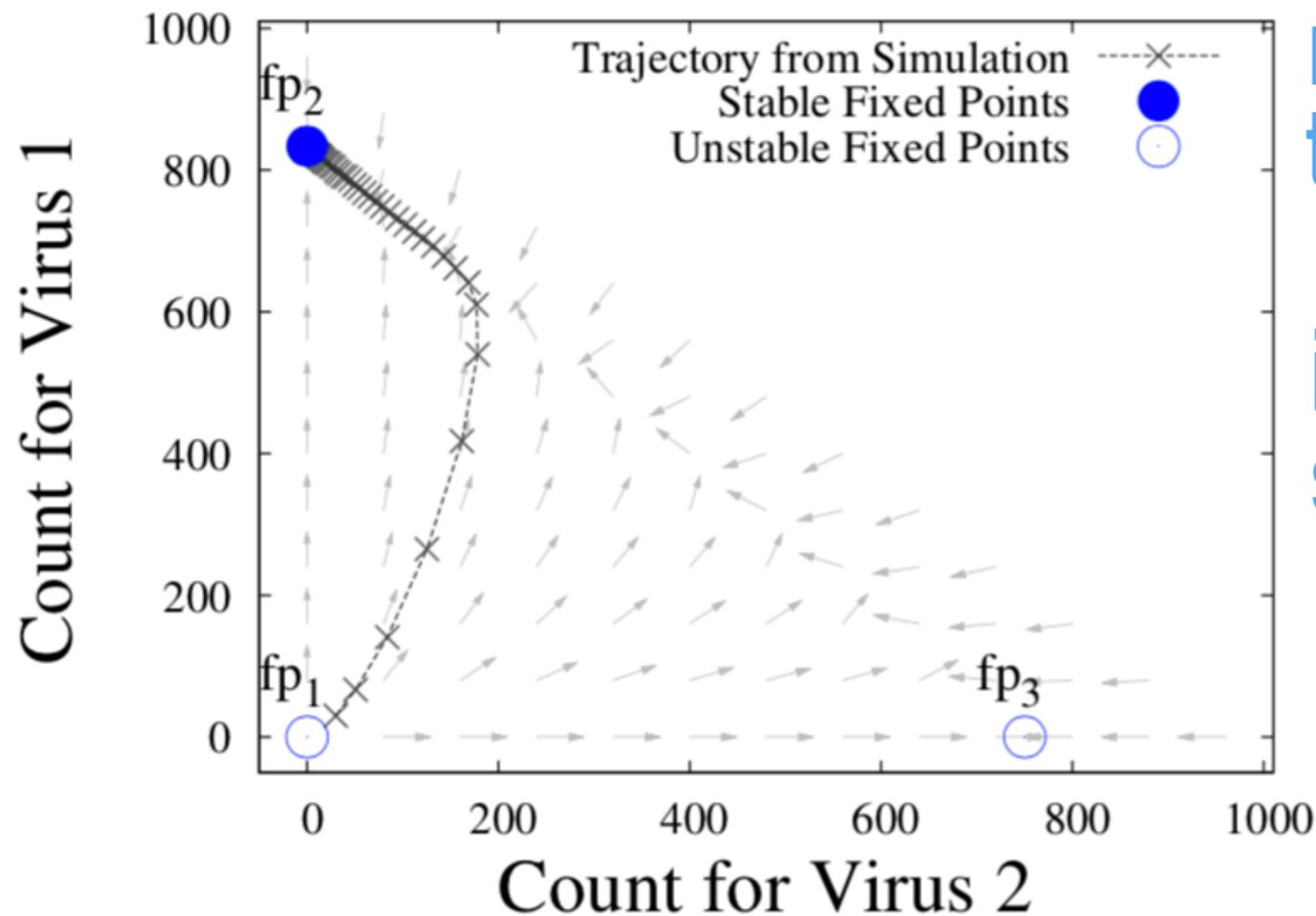
$\xi = 0$ mutual exclusion



A_{ij} adjacency matrix
 $\rho [A]$ spectral radius

$\rho [A] \beta / \mu$ strength of the strain

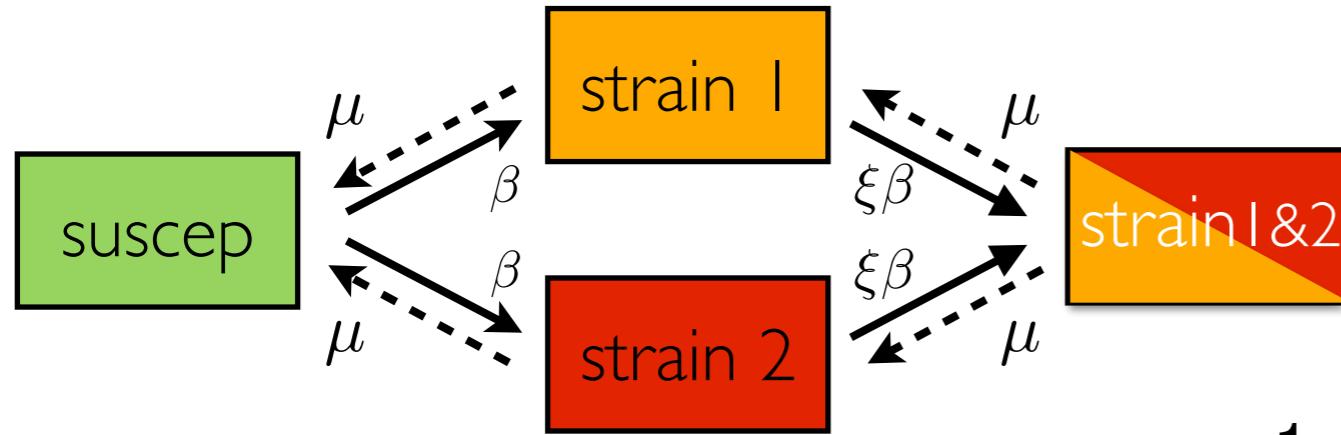
competitive diseases: winner takes it all!



**no co-existence
the stronger strain always wins**

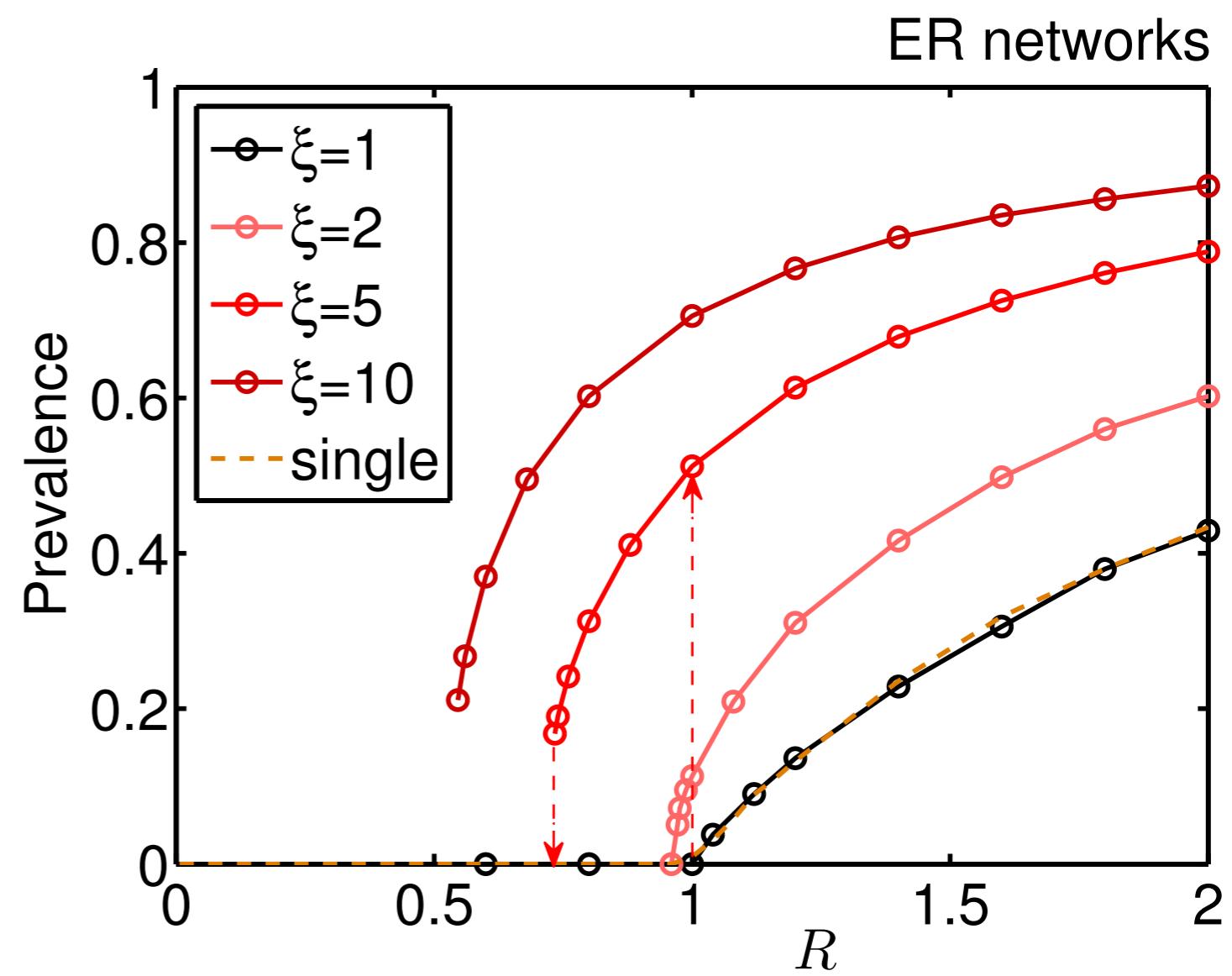
**if the 2 are equal
still only 1 strain survives**

cooperative diseases

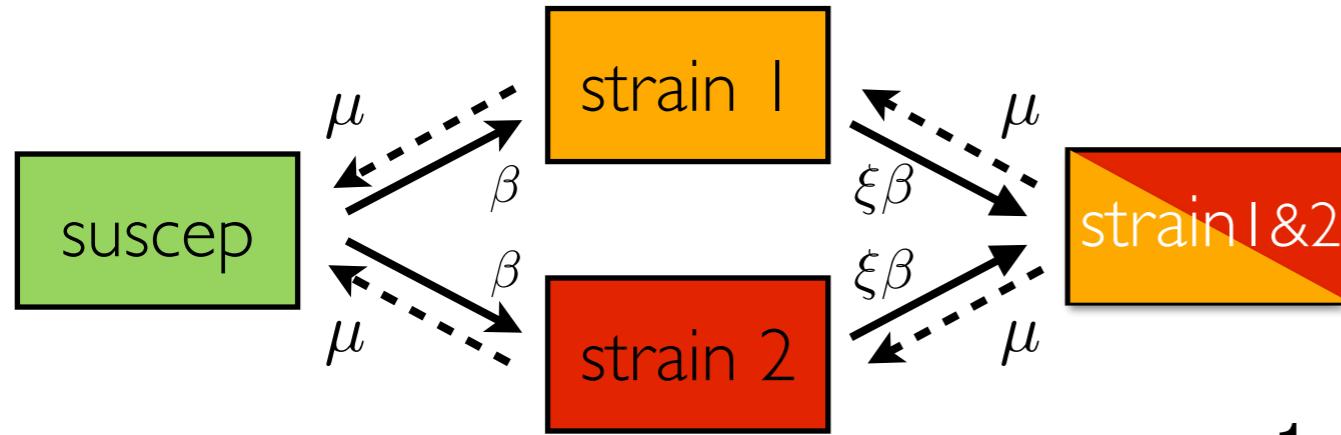


$\xi > 1$

- abrupt phase transitions

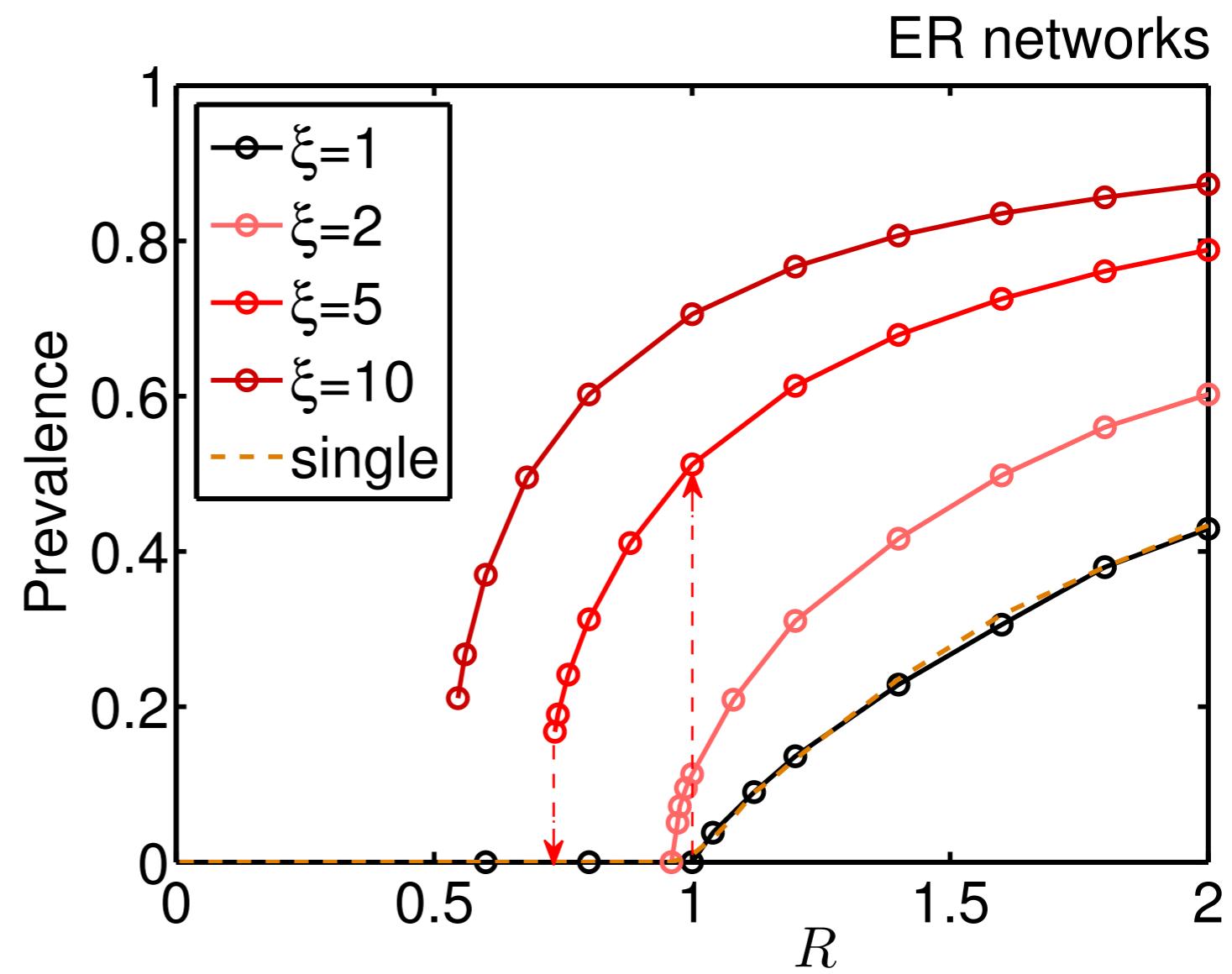


cooperative diseases

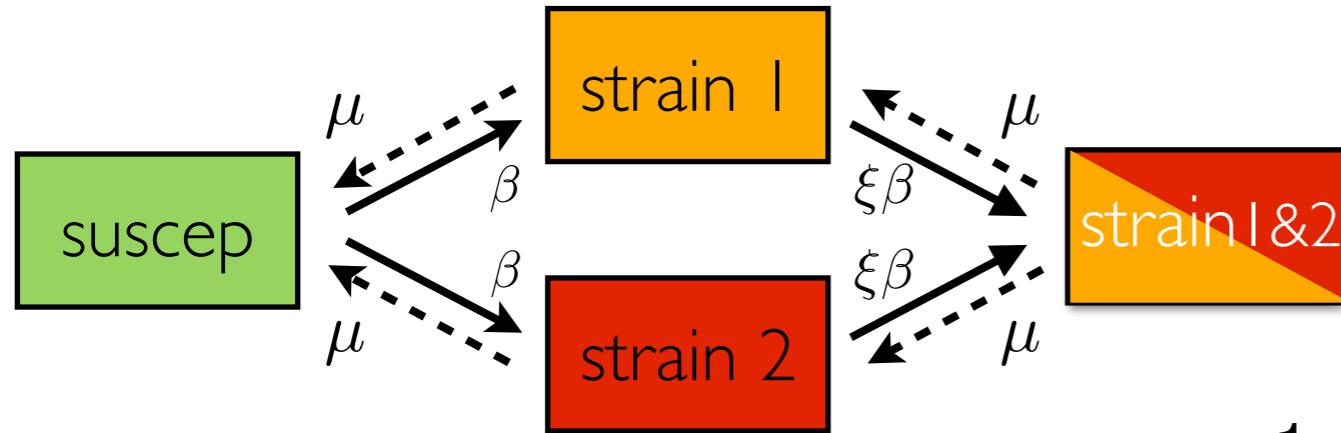


$\xi > 1$

- abrupt phase transitions
- hysteresis: epidemic threshold is different from the eradication threshold

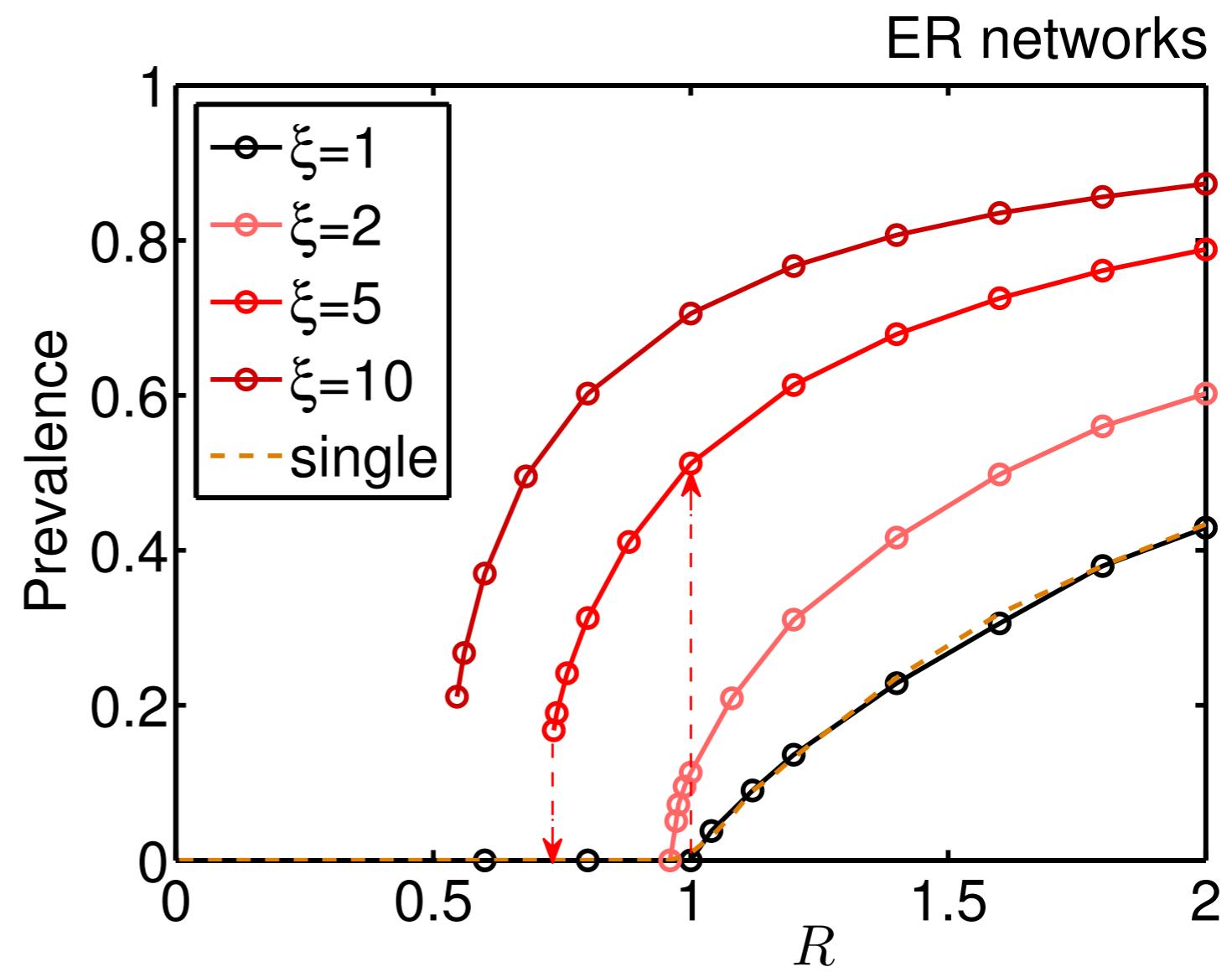


cooperative diseases

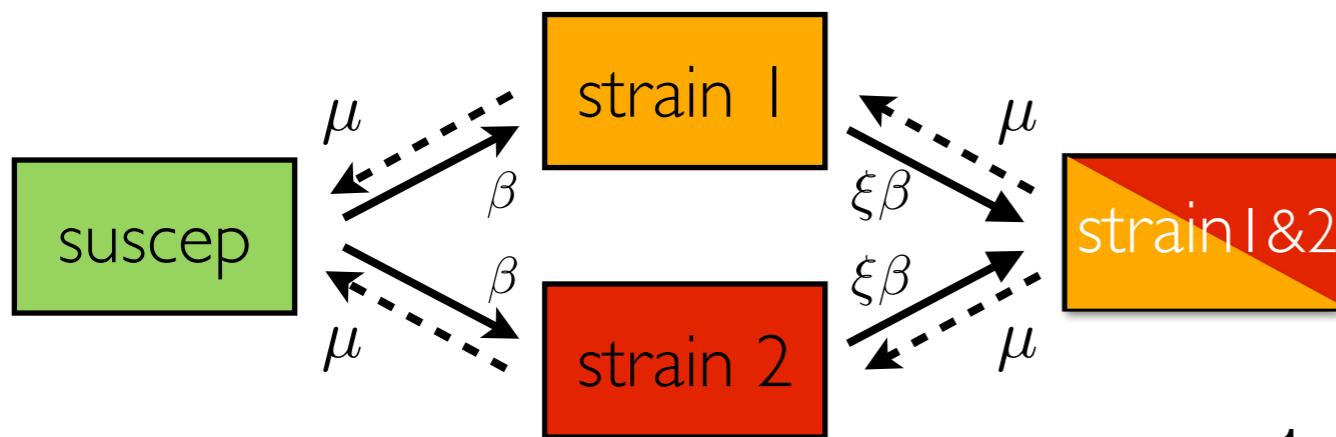


$\xi > 1$

- abrupt phase transitions
- hysteresis: epidemic threshold is different from the eradication threshold
- the behaviour depends on the network topology



cooperative diseases



$\xi > 1$

- abrupt phase transitions
- hysteresis: epidemic threshold is different from the eradication threshold
- the behaviour depends on the network topology

macroscopic patterns of interacting contagions are indistinguishable from social reinforcement

