Individual identity for disease metapopulations in scale-free networks

Franco Aquistapace

Physics of Life, Data and Epidemiology
Master's Degree in Physics of Data, DFA

Università
degli Studi
di Padova

Motivation

Disease metapopulations:

- A powerful approach for introducing spatial heterogeneities into epidemic dynamics. Allows to account for the mobility of individuals among different subpopulations.

Individual identity:

- Mobility can be introduced into the epidemic dynamics with different levels of detail. For real—world networks, the choice of mobility model has been shown to have a noticeable impact on the epidemic [1].

Scale—free population networks:

- The degree distribution of human population networks has been shown to follow a power—law behaviour $\longrightarrow$ scale—free networks.

[1] Keeling, M. J., Danon, L., Vernon, M. C., & House, T. A. (2010). Individual identity and movement networks for disease metapopulations. *Proceedings of the national academy of sciences*, *107*(19), 8866-8870.

SIR dynamics — multiple infectious stages

The epidemic dynamics are described in this work by the SIR model with m infectious stages. If x is a given subgroup of the entire population, then the epidemic process is given by the following reactions:

$$(S_x,\, I_x^1) \xrightarrow{\lambda_x} (S_x - 1,\, I_x^1 + 1)$$

Infection

$$(I_x^i,\, I_x^{i+1}) \xrightarrow{\gamma_i} (I_x^i - 1,\, I_x^{i+1} + 1)\ ,\ i = 1,\, \cdots,\, m-1$$

Disease progression

$$(I_x^m,\, R_x) \xrightarrow{\gamma_m} (I_x^m - 1,\, R_x + 1)$$

Recovery

- λ_x : force of infection (FOI) on subgroup x
- γ_i : recovery rate of infectious stage i

Here, the recovery rate is the same for all infectious stages, γ.

$\Rightarrow$ The infectious period is Gamma distributed.

SIR dynamics — human commuters

The specific details describing the effect of mobility on the epidemic dynamics can now be introduced in the model through the FOI.

Three alternative descriptions of mobility are considered here:

- **Kernel transmission (KT)**: The mobility network defines an effective spatial transmission kernel between subpopulations.

- **Random movers (RM)**: Each day, randomly chosen individuals move between subpopulations according to the mobility network.

- **Identity based (IB)**: The commuting individuals from each subpopulation are fixed and identified from the start of the epidemic.

At night the FOI is purely determined by the amount of infectious individuals within the home location.

SIR dynamics — human commuters

The FOI during daytime for the different mobility models is then given by:

- **KT model**: the identifier $x = i$ simply denotes the home location of individuals.

$$\lambda_i^{day} = \sum_j \sum_n \beta^{(n)} I_j^{(n)} \sum_k M_{ik} M_{jk} \left(\sum_l N_l M_{lk} \right)^{-1} \qquad \lambda_i^{night} = \sum_n \beta^{(n)} I_i^{(n)} N_i^{-1}$$

- N_l : Resident population in l.
- $\beta^{(n)}$: Transmission rate for infectious stage n.
- M_{ij} : Mobility matrix, proportion of individuals that live in population i and commute to j every day.

SIR dynamics — human commuters

The FOI during daytime for the different mobility models is then given by:

- **RM and IB models**: the identifier $x = (h, d)$ denotes both the home and daytime location.

$$\lambda^{day}_{(h,d)} = \sum_j \sum_n \beta^{(n)} I^{(n)}_{(j,d)} \left(\sum_k N_{(k,d)} \right)^{-1} \qquad \lambda^{night}_{(h,d)} = \sum_c \sum_n \beta^{(n)} I^{(n)}_{(h,c)} \left(\sum_k N_{(h,k)} \right)^{-1}$$

- $N_{(h,d)}$: Number of individuals with home location h that commute to d.
- **RM**: The commute locations d are randomly reassigned to each individual every day according to the mobility matrix.
- **IB**: The commute locations are fixed from the start.

Gillespie algorithm

For a system under Markovian dynamics, described as a set of reactions with rates $\{\omega_i\}$, the time to the next event of type i is distributed as $p(t_i) \sim \exp(-w_i t)$. Then, the time to the next event of any type is distributed as:

$$p(t) \sim \exp\left(-\sum_i \omega_i t\right)$$

Where an event represents the transition of the system from one sate to another. The probability that the event is of type i is given by:

$$p(i \mid t) = \frac{\omega_i}{\sum_j \omega_j}$$

Then, the system is evolved in time iteratively by:

1. Sampling the time t of the next transition from $p(t)$.
2. Sampling the transition type from $p(i \mid t)$.
3. Recalculating $\{\omega_i\}$ if necessary.

Linear population networks

Epidemic dynamics are simulated on a linear chain of $L = 50$ cities, with $N = 2 \times 10^3$ individuals per city.

Mobility matrix M constructed so that a fixed fraction σ of individuals in a given city i commutes only to the adjacent cities $i - 1$ and $i + 1$ (closed boundary conditions).

Three main focus points:

- Different mobility fractions σ.
- Different recovery rates γ and basic reproductive ratios R_0.
- Stochastic variability of the system.

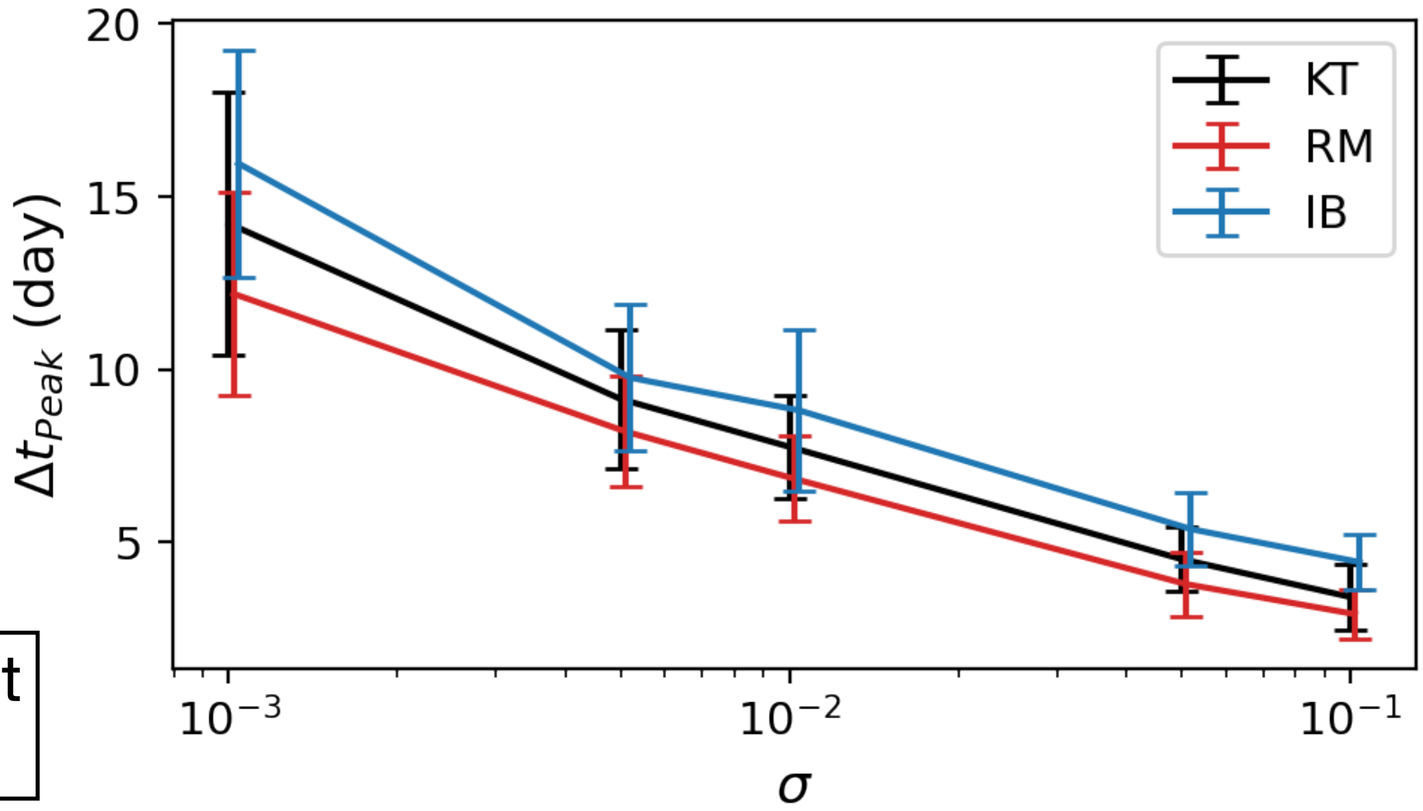
Test of the methods $\longrightarrow$ Expected results are known [1].

[1] Keeling, M. J., Danon, L., Vernon, M. C., & House, T. A. (2010). Individual identity and movement networks for disease metapopulations. *Proceedings of the national academy of sciences*, *107*(19), 8866-8870.

Linear population networks - Mobility fraction

- Mobility fractions $\sigma \in [10^{-3}, 10^{-1}]$.
- Remaining parameters fixed as $\gamma = 0.24/\text{day} = 0.01/\text{hour}$, $R_0 = 2$ and $m = 5$ infectious stages (~influenza).
- Each simulation starts with one infected individual in the first node of the chain.

Time difference between prevalence peaks of adjacent cities.

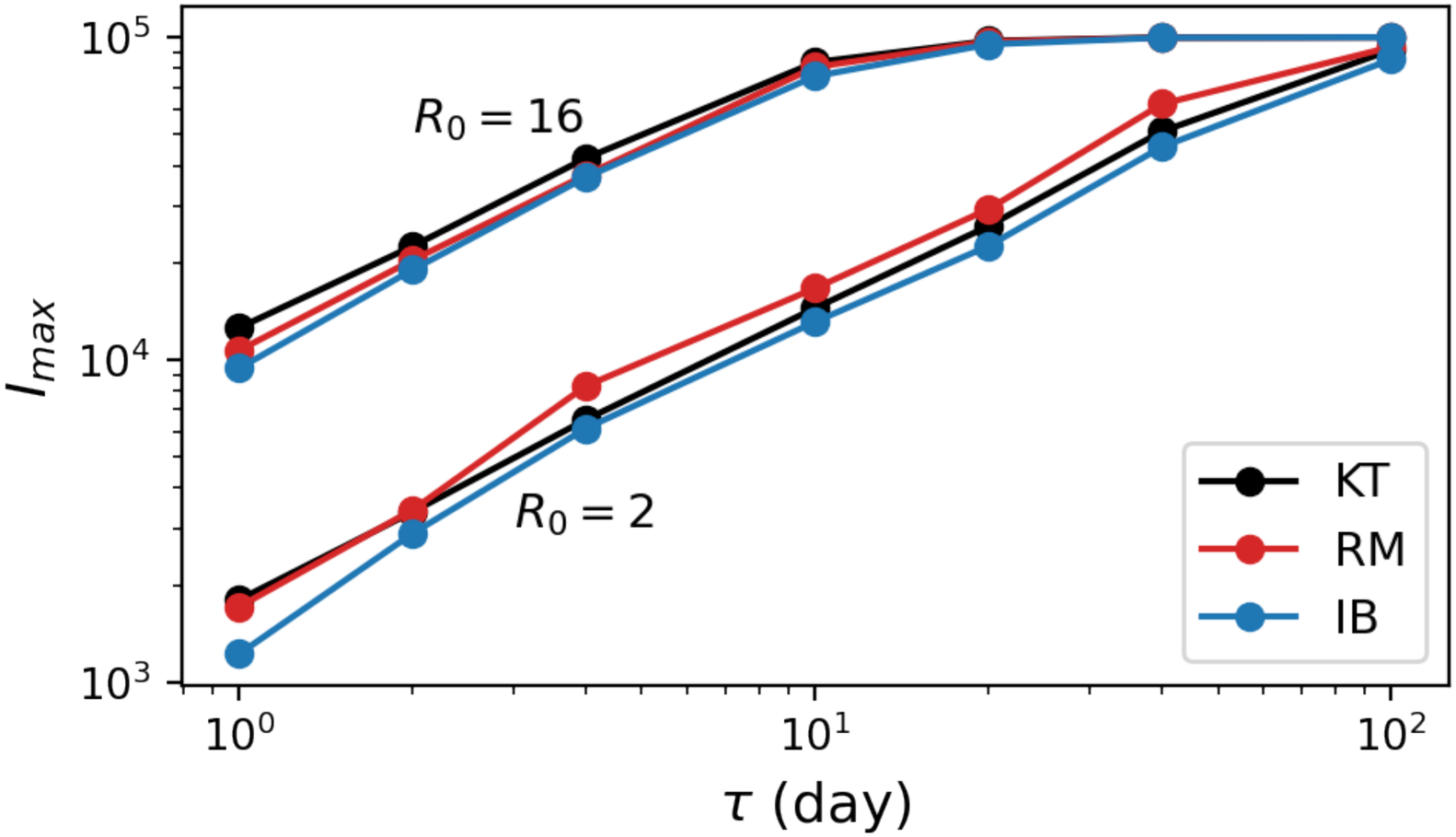


Difference between KT and RM might be caused by small population size

Linear population networks - Epidemic parameters

- Mobility fraction $\sigma = 10^{-2}$ and $m = 5$ infectious stages.
- Recovery rate $\gamma \in [10^{-2}/\text{day}, 1/\text{day}]$, infectious period $\tau = \gamma^{-1}$.
- Basic reproductive ratio $R_0 = 2,\ 16.$
- Each simulation starts with one infected individual in the first node of the chain.

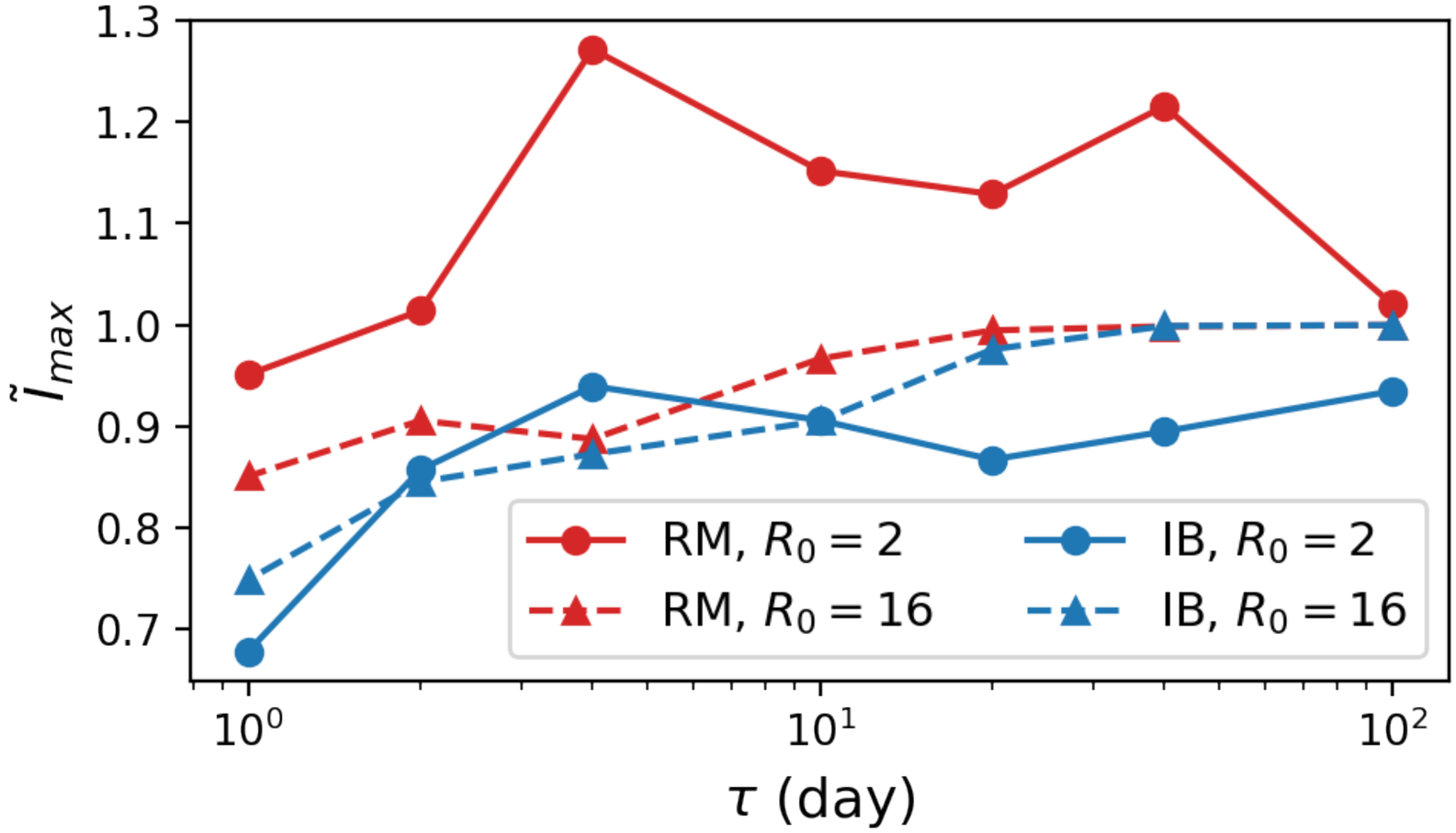
Global prevalence peak height



Linear population networks - Epidemic parameters

- Mobility fraction $\sigma = 10^{-2}$ and $m = 5$ infectious stages.
- Recovery rate $\gamma \in [10^{-2}/\text{day}, 1/\text{day}]$, infectious period $\tau = \gamma^{-1}$.
- Basic reproductive ratio $R_0 = 2,\ 16.$
- Each simulation starts with one infected individual in the first node of the chain.

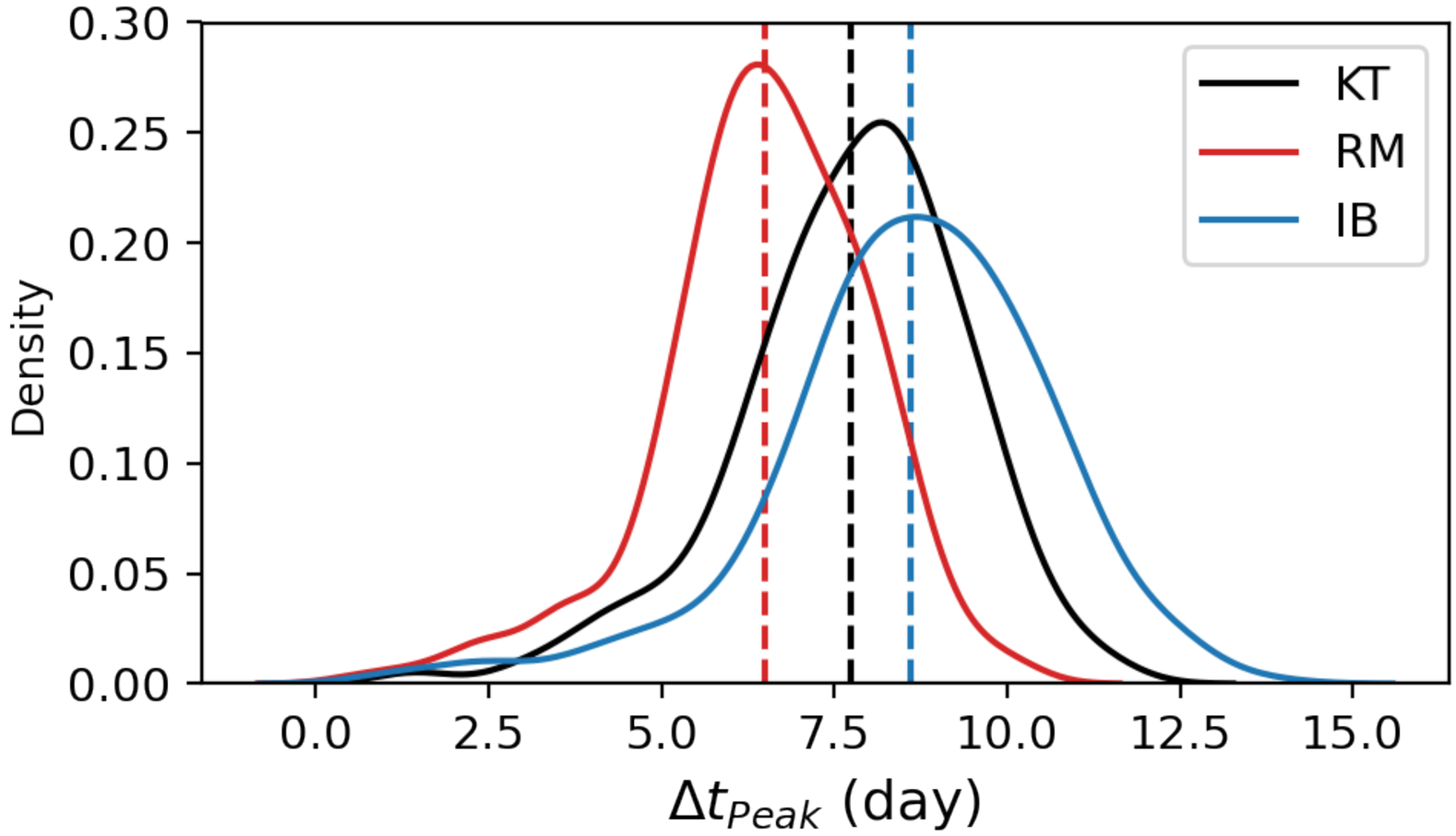
Relative peak $\tilde{I}_{max} = \dfrac{I_{max}}{I_{max}^{(KT)}}$



Linear population networks - Stochasticity

- Mobility fraction $\sigma = 10^{-2}$.
- Influenza—like parameters: $\gamma = 0.24/\text{day}$, $R_0 = 2$ and $m = 5$.
- Each simulation starts with one infected individual in the first node of the chain.
- 10 simulations per model.

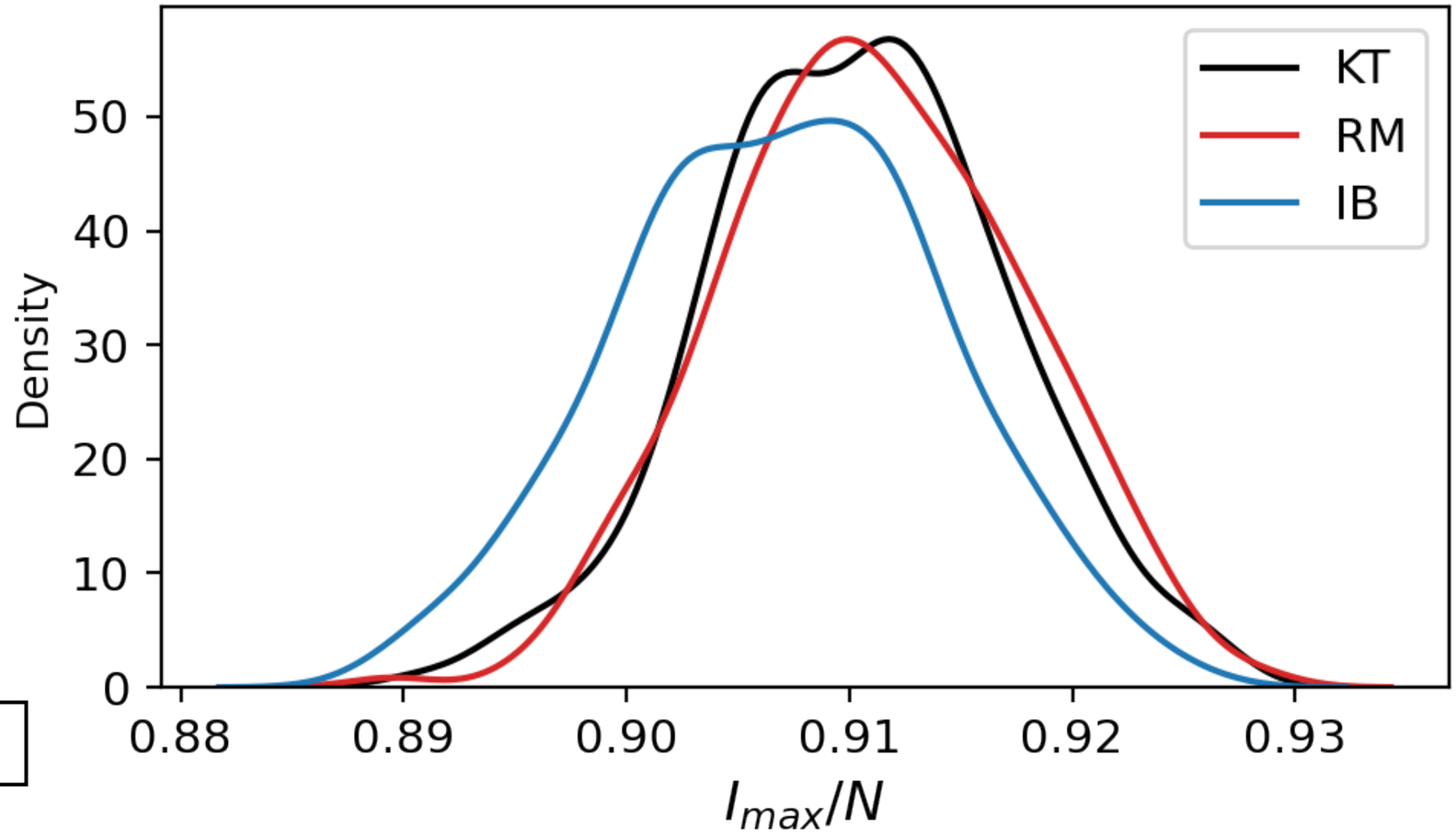
Distribution of time difference between prevalence peaks



Linear population networks - Stochasticity

- Mobility fraction $\sigma = 10^{-2}$.
- Influenza—like parameters: $\gamma = 0.24/\text{day}$, $R_0 = 2$ and $m = 5$.
- Each simulation starts with one infected individual in the first node of the chain.
- 10 simulations per model.

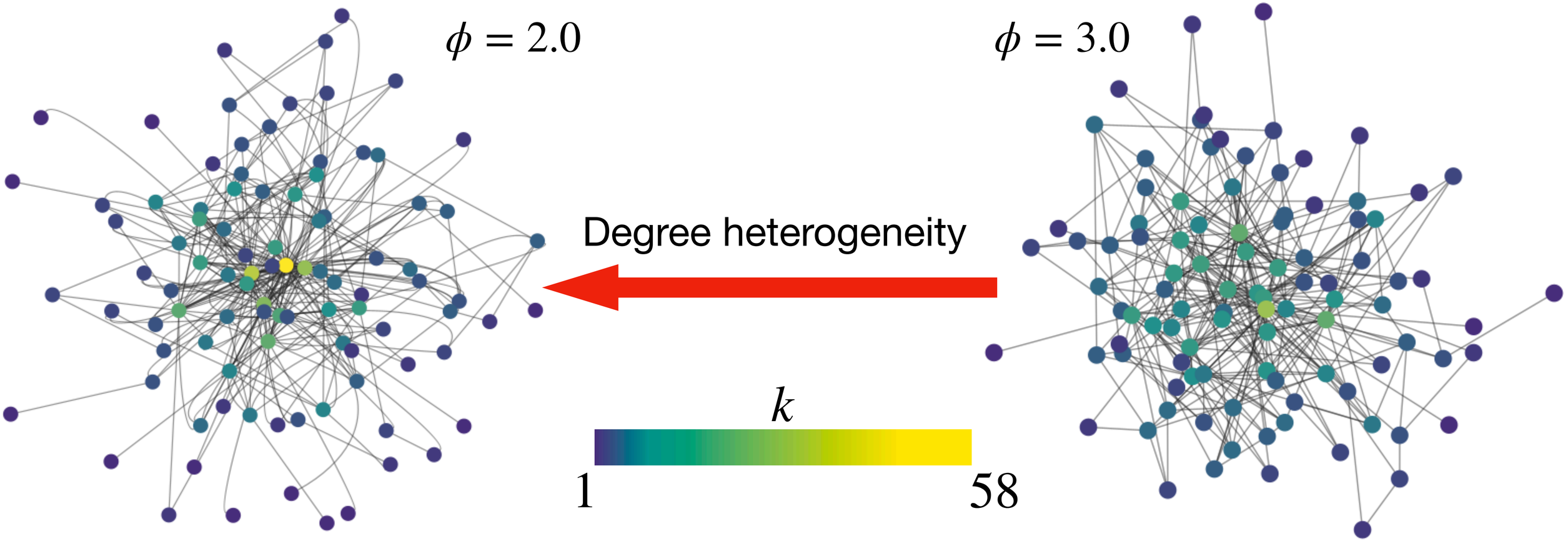
Distribution of prevalence peak relative heights



This is expected!

Scale—free population networks

- Networks with power—law degree distributions: $P(k) \sim k^{-\phi}$, usually $\phi \in [2,3]$.
- Scale—free networks are characterised by the presence of few highly connected **hubs**, and many **peripheral** nodes with few connections.



Scale—free population networks

The static method [2] is used here to generate scale—free networks with desired properties:

1. A set of N nodes, indexed as $i = 1,\cdots,N$, are assigned a weight $w_i = i^{-\alpha}$, where $\alpha \in [0,1]$ controls the degree distribution of the network.
2. In an iterative manner, two nodes i and j are selected with relative probabilities w_i and w_j, and an edge is created between them if not already present.
3. This process is repeated until the average degree of the network reaches a desired value $\langle k \rangle$.
4. For large N, the degrees of the network are distributed as $P(k) \sim k^{-\phi}$, with $\phi = 1 + 1/\alpha$.

[2] Goh, K. I., Kahng, B., & Kim, D. (2001). Universal behavior of load distribution in scale-free networks. *Physical review letters*, *87*(27), 278701.

Scale—free population networks

After the population network is generated, the amount of individuals for population with degree k is [3]:

$$N_k = \bar{N}\frac{k}{\langle k \rangle}$$

The fraction of individuals commuting from a population with degree k to a neighbour population with degree k', $\sigma_{kk'}$, is defined as:

$$\sigma_{kk'} = \sigma\frac{N_{k'}}{N_k + N_k^{nn}}$$

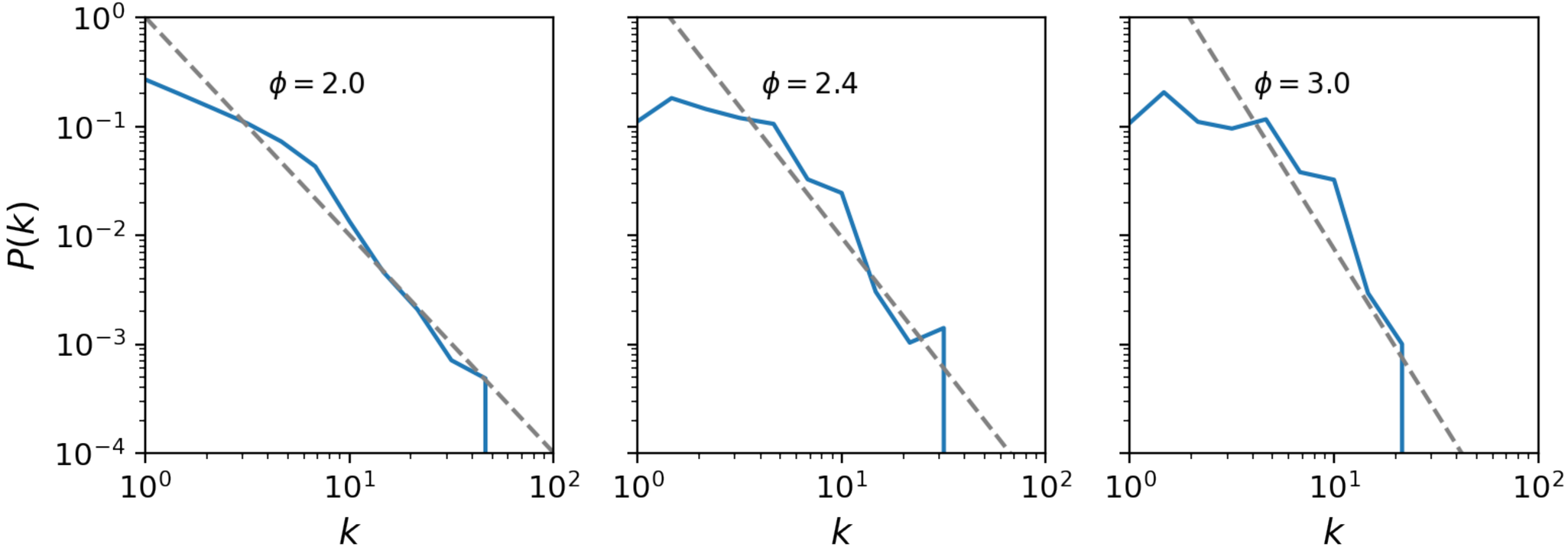
- $\bar{N}$: Average number of individuals per population
- σ : Mobility fraction parameter.
- N_k^{nn} : Average number of individuals among the neighbouring populations

[3] Balcan, D., & Vespignani, A. (2011). Phase transitions in contagion processes mediated by recurrent mobility patterns. *Nature physics*, *7*(7), 581-586.

Scale—free population networks

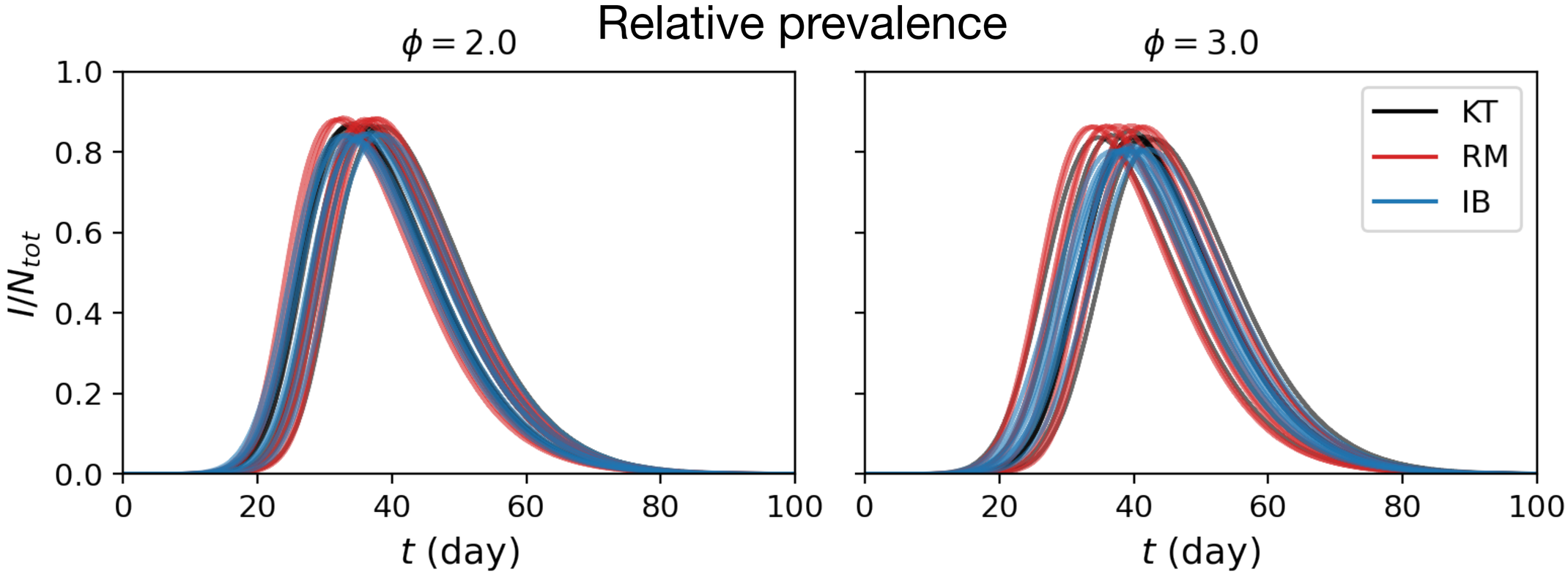
Scale—free networks are produced with $\phi \in [2.0,3.0]$, $\langle k \rangle \approx 6$ and $N \approx 100$. The nodes with degree zero are removed from each network.

The population of each node is generated with $\bar{N} = 2 \times 10^3$, and the mobility network is generated with $\sigma = 10^{-2}$.



Scale—free population networks — Global behaviour

- Simulations are performed with parameters $\gamma = 0.24/\text{day}$, $R_0 = 2$ and $m = 5$.
- Each simulation starts with one infected individual in the most populated node of the network.
- $N_{sim} = 8$ simulations per model.

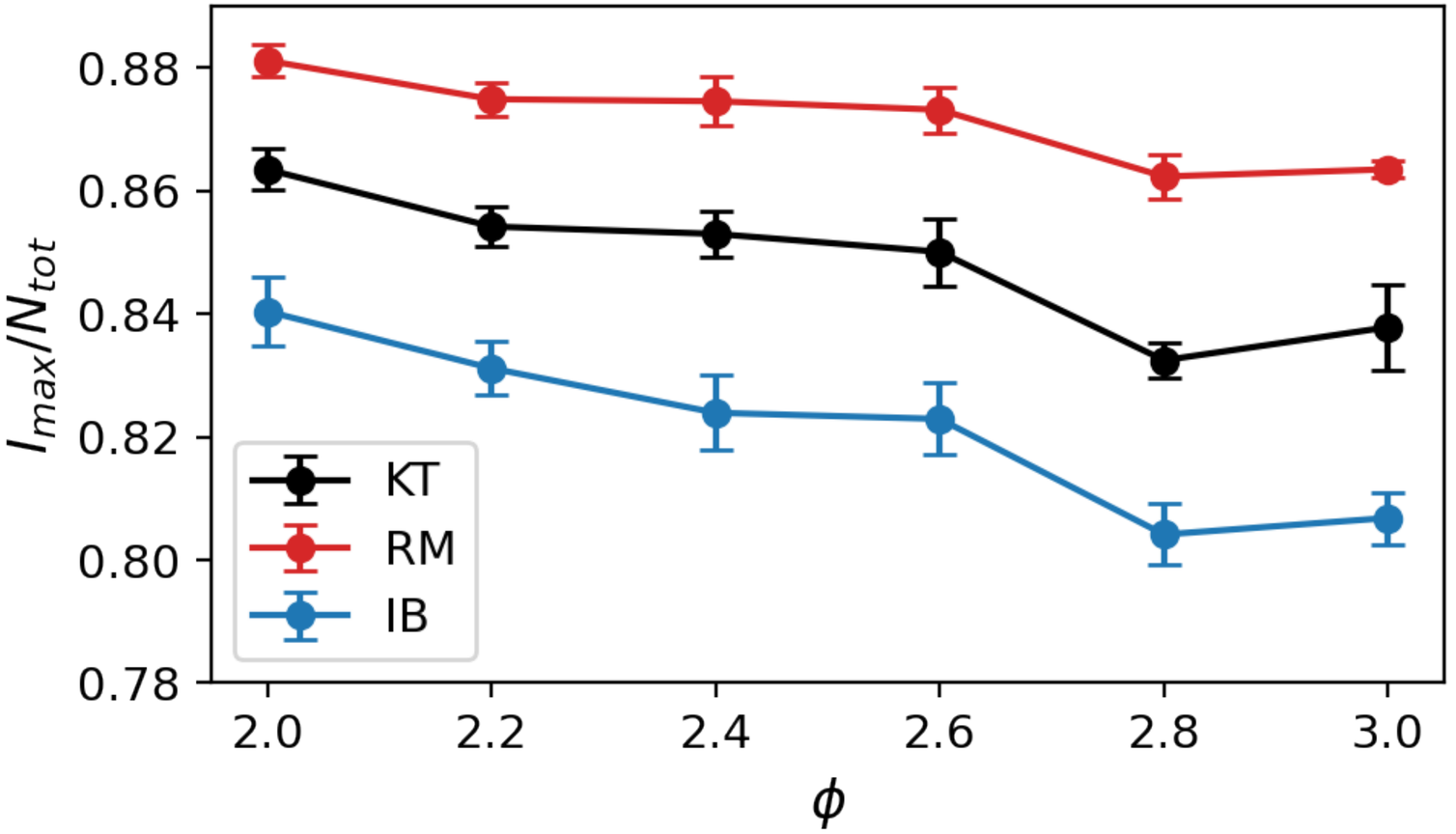


Scale—free population networks — Global behaviour

- I_{max} : Maximum prevalence (height of the prevalence peak).
- Accounting only for discrete individuals leads to a higher peak.
- Accounting for individual identity leads to a lower peak.
- No significant difference for the time of the peak.

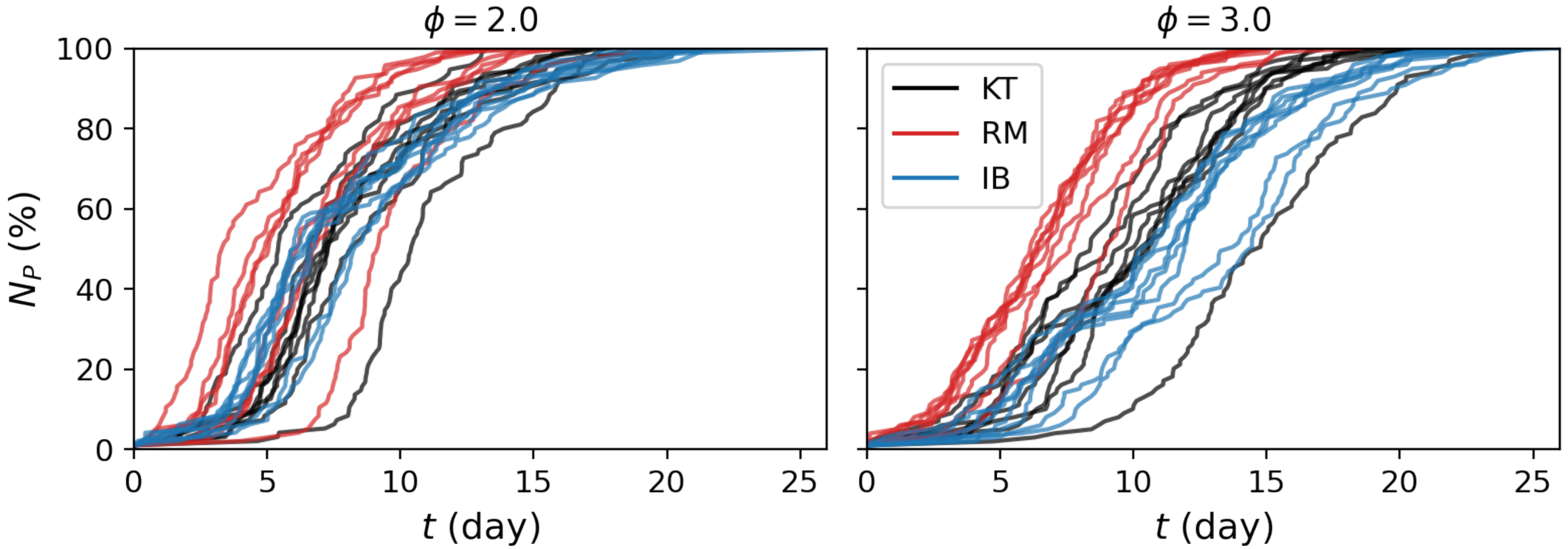
In line with simulations on real—world networks.

Relative prevalence peak height



Scale—free population networks — Percolation

- $N_P(t)$: Amount of populations with $t_{Peak} \leq t$. Time centred at t_{Peak} of seed node.
- Degree homogeneity appears to hinder the spread of the disease
  ⇒ Percolation generally occurs later
- Insight about the spatial spread of the disease.



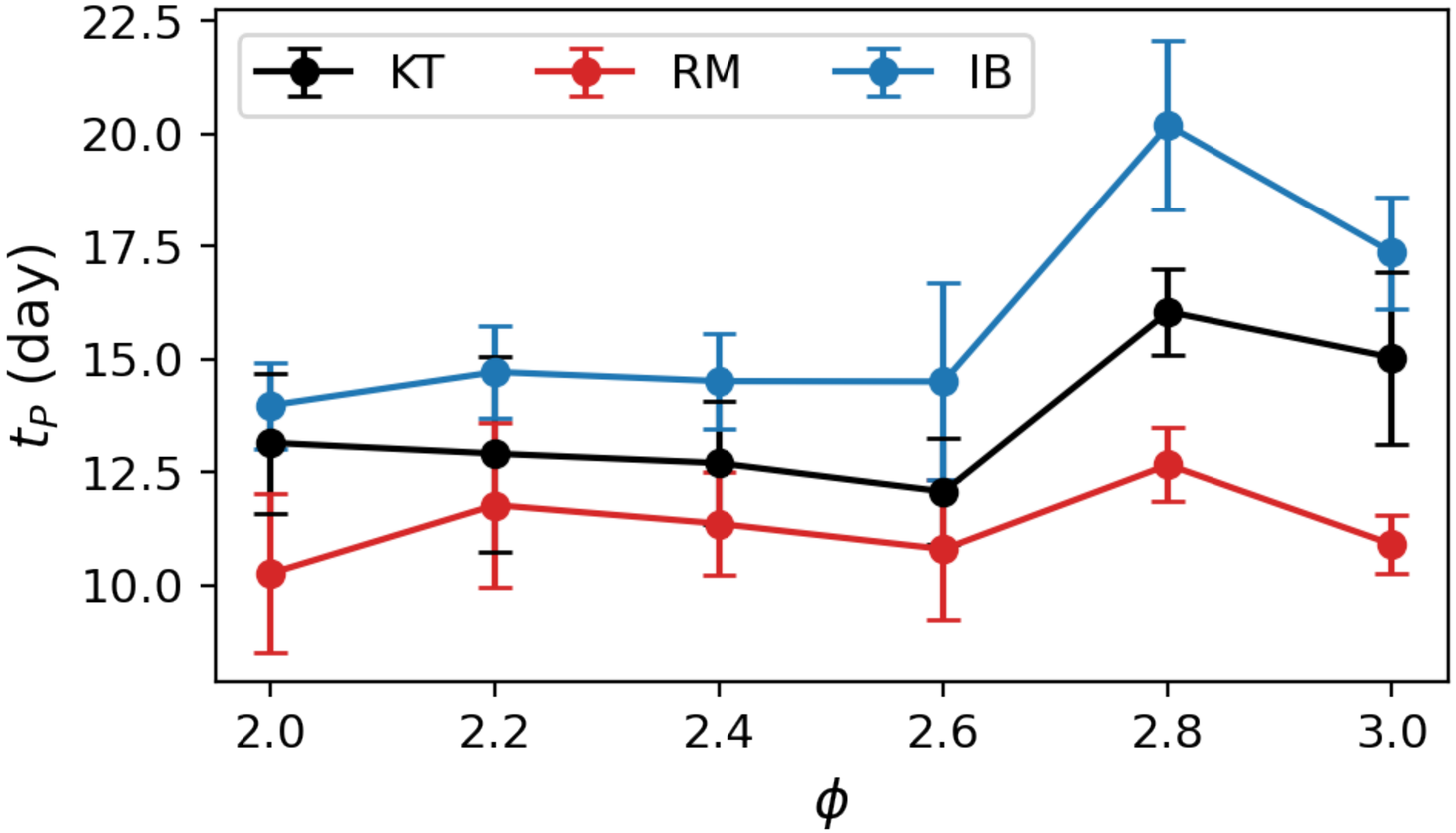
Scale—free population networks — Percolation

- Percolation time:

$$t_P = t_{90\%} - t_{FP}$$

- t_{FP} : Time of prevalence peak in the seed population.
- $t_{90\%}$: Time at which $N_P = 90\,\%$

Percolation time



Overall slower spread when identity is accounted for!

Concluding remarks

- Degree distribution in scale—free networks has an effect over the epidemic dynamics:

  ➡Higher heterogeneity in the network leads to a higher prevalence peak and generally faster spatial spreading.

- The level of detail in the mobility model has a noticeable impact on scale—free networks, regardless of the degree distribution:

  ➡Accounting for individual identity leads to a lower prevalence peak and slower spatial spreading.

Follow up

- Explore effect of epidemic parameters in more detail.
- Sampling more synthetic networks → One network does not describe the ensemble
- Different population sizes → Do our results change with significantly larger N?
- Different initial conditions → Does it matter if the seed node is a hub?
- More simulations!

Data availability

The code used to perform the simulations and visualise the results, as well as the files with simulation data, are available at a public GitHub repository:

https://github.com/FrancoAquistapace/ComplexNetworksEpidemiology