Individual based infection models on (not so) dense large random networks

Aurélien Velleret

Université Paris-Saclay, INRAE, MaIAGE, Jouy-en-Josas, France

aurelien.velleret@nsup.org

24. July 2023, SPA 2023

Individual-based models on prescribed graphs

- SIS in continuous time and on a G oriented graph (fixed throughout the epidemic)
- Weighted edges : intensity of infectious transmission
- Individuals can be either susceptible S or infectious I.
- n individuals are considered, n being large.
- Remission at a rate γ_i for individual $i \in [\![1, n]\!]$.
- Provided that i is susceptible and j infected, j infects i at rate $w_{i,j}^G$.

A characterization of individuals

- Individual $i \in [\![1, n]\!]$ is described by some type X_i on a general state space $\mathcal{X}$
- Age, Comorbidity, Spatial distribution or Social Belongings...
- μ is seen as the generating distribution for the X_i's (i.i.d. representation or simply limiting empirical distribution)
- Rate of remission : $\gamma_i = \gamma^{(n)}(X_i)$
- These characteristics shall also be responsible for the structure of the graph and the weights of the edges.

The generation of a large random graph

As a basic assumption, pairwise independent connections :

- Given X_i and X_j,
  the weight $w^G_{i,j}$ can only be 0 (no connection)
  or $w_I^{(n)}(X_i, X_j)$.
- Value $w_I^{(n)}(X_i, X_j)$ with probability $w_E^{(n)}(X_i, X_j)$, $w_E^{(n)}$ symmetrical.
- The presence of the contact is undirected (hence the symmetry),
  yet the intensity (weight) can be directed
  (differences in susceptibility or infectivity).
- Fixed traits : Level of vaccination or immunity not described.

The description with a kernel w

Model of infection at the characteristic level :

$$\partial_t u(t,x) = (1-u(t,x)) \int_{\mathcal{X}} \mu(dy) u(t,y) w(x,y) - \gamma(x) u(t,x).$$

- $u(t,x)$: probability for an individuals of type x to be in state I at time t, $(1-u(t,x))$ to be in state S
- $\mu(dy)u(t,y)$: weighted type distribution of infected individuals.
- formula of connection between the models :

$$w \sim n \cdot w_I^{(n)} \cdot w_E^{(n)}$$

- $w_I^{(n)}$ constant $\Rightarrow$ w as a graphon, for it encodes the graph structure.

At which level of sparsity is something different ?

- Kernel description valid provided the level of interaction per link tends to zero (average criterion on $w_I^{(n)}$).
- No issue from individuals having degree of order n^α, for any $\alpha \in (0, 1]$.
- May be seen for small α at the fluctuation level (like CLT rather than LLN)

Individual based SIS models on (not so) dense large random networks

postdoctoral research in collaboration
with Jean-François Delmas, Viet Chi Tran, Pierre-André Zitt (on Marne)
together with Federica Garin and Paolo Frasca (on Grenoble)

Summary

The main results

- The limiting dynamics on the trait space
- General interaction kernel

Numerics

- Discrepancy with the limit
- Very sparse interactions

Summary

Convergence of the empirical measure

For each individual $i \in [\![1, n]\!]$:

- $X_i = X_i^{(n)} \in \mathcal{X}$: fixed type,
- $E_t^i = E_t^{i,(n)} \in \{I, S\}$: infectious status at time t.

Empirical measure of interest :

$$\eta_t^{(n)} := \frac{1}{n} \sum_{i \leq n} \delta_{(X_i, E_t^i)} \in \mathcal{M}_1(\mathcal{X} \times \{I, S\})$$

Expected limit :

$$\eta_t(dx, de) := \mu(dx) \cdot ([1 - u(t, x)] \cdot \delta_S(de) + u(t, x) \cdot \delta_I(de)).$$

Existence and uniqueness conditions

$$\partial_t u(t,x) = (1-u(t,x)) \int_{\mathcal{X}} \mu(dy) u(t,y) w(x,y) - \gamma(x) u(t,x). \tag{1}$$

(A1) : Boundedness and regularity conditions

- The non-negative function γ is bounded μ-a.e. continuous .
- The non-negative function w is bounded $(\mu \otimes \mu)$-a.e. continuous function.
- The probability measure η_0 is absolutely continuous with respect to μ.

Delmas, J.F., Dronnier, D., Zitt, P.A. ; An Infinite-Dimensional SIS Model (2022)

The regularity condition on the kernel is mild

These choices for w are covered :

- Stochastic Block Model : finite $\mathcal{X}$.
- w piecewise Lipschitz on $[0,1]\times[0,1]$
  (classically assumed for graphon convergence).
- w continuous on a subspace of $\mathbb{R}^d$.
- Geometric random graphs : $w_E^{(n)}(x,y) = 1_{\{|x-y|<r\}}$ for a constant $r>0$, with uniformly distributed traits on $\mathcal{X} = [0,1]^d$.
- $w(x,y) = \hat{w}[d(x,y)]$,
  where d is the distance on a variety $\mathcal{X}$
  and $\hat{w}$ is piecewise continuous.

A slighly more general problem

For any f on $\mathcal{X} \times \{S, I\}$, bounded measurable and integrable in x.

$$\begin{aligned}\langle \eta_t \mid f \rangle = \langle \eta_0 \mid f \rangle + \int_0^t \int_{\mathcal{X}} [\eta_r(dx, S) \cdot \mathcal{A}^S f(x; \eta_r) \\ + \eta_r(dx, I) \cdot \mathcal{A}^I f(x; \eta_r)] dr,\end{aligned} \tag{2}$$

where $\eta_0(dx, de)$ is the initial condition,
$\mathcal{A}^S$ and $\mathcal{A}^I$ are related to the transition rates :

$$\mathcal{A}^S f(x; \eta) := (f(x, I) - f(x, S)) \cdot \int_{\mathcal{X}} \eta(dy, I) w(x, y),$$
$$\mathcal{A}^I f(x; \eta) := (f(x, S) - f(x, I)) \cdot \gamma(x).$$

Existence and uniqueness result

Existence and uniqueness

Provided $(A1)$ is granted, there exists a unique solution η to the problem (2) among the measurable functions from $[0, T]$ to the set of positive measures on $\mathcal{X} \times \{S, I\}$.

$$\eta_t(dx) := \mu(dx) \cdot ((1 - u(t,x)) \cdot \delta_S(de) + u(t,x) \cdot \delta_I(de)),$$

where u is the unique solution to (1) with initial condition u_0, u_0 being the density of $\eta_0(I, dx)$ with respect to $\mu(dx)$.

Recall : interaction kernel on a sampled graph

Rule $(G1)$ for $(\eta_t^{(n)})_{t\geq 0}$ with $(\gamma^{(n)}, w_E^{(n)}, w_I^{(n)})$:

- Remission of i at rate $\gamma^{(n)}(X_i)$
- For infections, either no connection
  or transmission from j to i at rate $w_I^{(n)}(X_i, X_j)$.
- Contact with probability $w_E^{(n)}(X_i, X_j)$,

given the (X_i)'s.

Rule $(G0)$ with $(\gamma^{(n)}, w^{(n)})$ is the mean-field case,
i.e. rule $(G1)$ with $(\gamma^{(n)}, 1, w^{(n)})$.

Core assumption on the main result

(A2) : Convergence of the parameters

- $n \cdot w_I^{(n)} \cdot w_E^{(n)}$ converges to w in the uniform norm.
- $A_1^{(n)}(w_I^{(n)})$ converges to 0, where :

$$A_1^{(n)}(w_I^{(n)}) = \int_{\mathcal{X}} \mu^{(n)}(dx) \int_{\mathcal{X}} \mu^{(n)}(dy) \ \min\{w_I^{(n)}(x,y), 1\}$$
$$:= (1/n^2) \sum_{\{i,j \leq n\}} \ \min\{w_I^{(n)}(X_i, X_j), 1\}.$$

- + classical boundedness and convergence properties

Main result

Main convergence result

Assume $(A1)$ (for γ, w and $\eta_0^{(n)}$) and $(A2)$ (for $(\gamma^{(n)}, w_E^{(n)}, w_I^{(n)})$).
Then, $(\eta^{(n)})_n$ obeying rule $(G1)$ converges
in the Skorokhod space $\mathcal{D}([0,T], \mathcal{M}_+(\mathcal{X} \times \{S, I\}))$
to the solution η of problem (2).

Previous studies 1/2 :

- PhD of *Dronnier with Delmas and Zitt* ('22-23) :
  Study of the limiting dynamics,
  with the scope of modeling targeted vaccination strategies
- *Perkins ('99)* : reference for the tightness criterion
- *Fournier & Méléard ('04)* : generic reference
  of large population limits for mean-field kernels
  of individual interactions, i.e. $w_E^{(n)}(x, y) \equiv 1$

Previous studies 2/2 :

- *Kuehn & Throm ('19) :*
  Graphon limit of a system of ODEs (with dense connections $w_E^{(n)} = O(1)$)
- *Billiard & Leman & Rey & Tran ('22) :*
  Two step convergence, first from individual-based model to a system of ODEs, then graphon limit (also with dense connections $w_E^{(n)} = O(1)$)
- *Keliger & Horváth & Takács ('22) :*
  Graph limit of individual-based models, $w_I^{(n)}(x, y) \equiv \epsilon^{(n)}$ (scaled graphon structure), piecewise Lipschitz kernel

Core assumption on sparsity

Interest of $w_I^{(n)}$ as a type function ?

$$(1/n^2)\sum_{\{i,j\leq n\}} \min\{w_I^{(n)}(X_i, X_j), 1\} \to 0$$

appears the most efficient average to exploit.
Note for instance that rare highly contagious contacts
can be neglected.
Typical example that is covered, with density coefficient $\alpha \in (0, 1]$:

$$w_I^{(n)}(x, y) = \frac{w_I(x, y)}{n^\alpha} \; ; \quad w_E^{(n)}(x, y) = \frac{w_E(x, y)}{n^{1-\alpha}},$$

where $w = w_I \cdot w_E$.
Note the average degree : $d(x) \approx n \cdot \int_{\mathcal{X}} w_E^{(n)}(x, y)\mu(dy) = O(n^\alpha)$

Summary

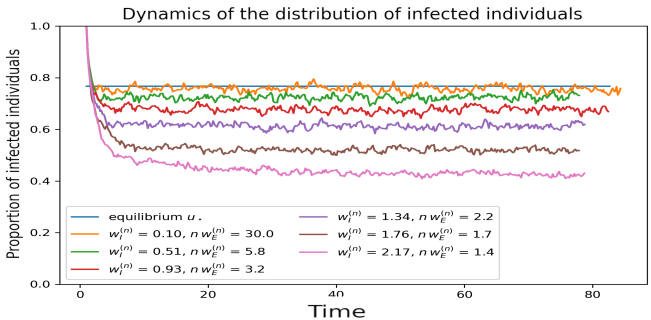
Context of simulations

As a starter, $\mathcal{X}$ is a singloton, so $\dot{u}_t = w\, u_t\,(1 - u_t) - \gamma\, u_t$.
Unless otherwise stated :

- $n = n_0 = 2000$: number of individuals
- $U = 1$: proportion initially infected
- $\gamma = 0.7$: remission rate
- $w = 3 \Rightarrow$ growth rate of the starting epidemic : $w - \gamma = 2.3$
- $w_I^{(n_0)} = 1.2 \qquad (w_I^{(n)} = (n/n_0)^{-\alpha} w_I)$
- thus $w_E = w/w_I = 2.5 \qquad (w_E^{(n)} = n^{\alpha-1} n_0^{-\alpha} w_E)$
- $\alpha = 0$: very sparse graph $\Rightarrow w_E$, average degree, of order 1

Discrepancy with the limit, varying sparsity

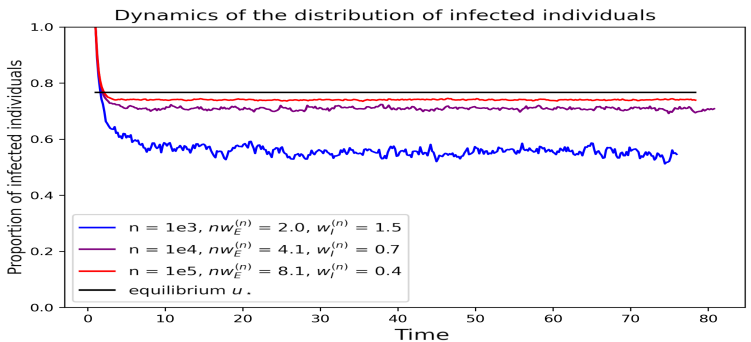
The larger is w_I, the larger is the discrepancy below the limit :



Comparison for large time to the equilibrium u_* of the ODE.

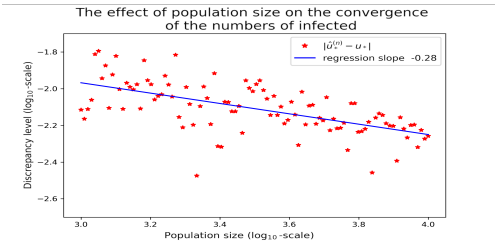
Deviation from the limit, large population size ?

This deviation is vanishing with increasing population sizes, even for intermediate levels of sparsity :

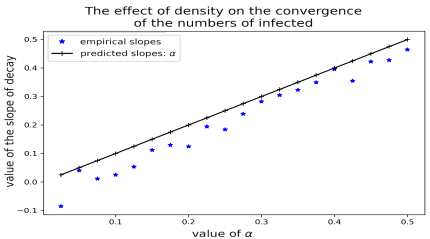


here $\alpha = 0.3$, i.e. $w_I^{(n)} = 1.2 \times (n/n_0)^{-0.3}$.

Deviation from the limit, large population size ?

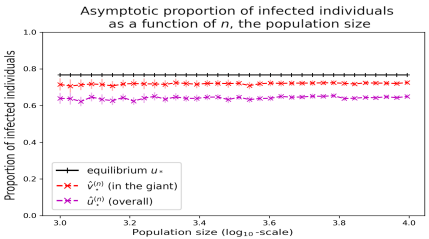


$|\hat{u}_*^{(n)} - u_*|$ for growing population size n and regression slope, for $\alpha = 0.3$. $R^2 \approx 0.38$ is the regression coefficient corresponding to the proportion of variance captured by the prediction with a slope of -0.3 (and adjusted averages).

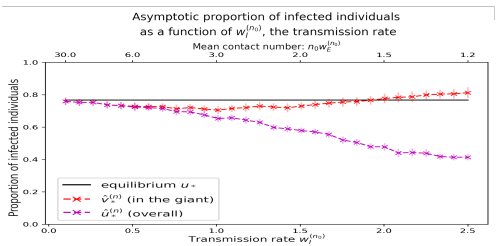


Slopes of the log-log regressions of $|\hat{u}_*^{(n)} - u_*|$ vs. n, for various values of α. $R^2 \approx 0.97$ is the regression coefficient corresponding to the proportion of variance captured by the prediction of slopes given by exactly $-\alpha$.

Deviation from the limit in large population size ?



Temporal average of the proportion of infected individuals for growing population size n (in log-scale), with $w_I^{(n)} = 1.2$ and $n \cdot w_E^{(n)} = 2.5$.



Temporal average of the proportion of infected individuals as a function of $w_I^{(n_0)}$, with fixed population size n_0 and $w_E^{(n_0)} \cdot w_I^{(n_0)} = 3/n_0$.

Summary

The main results

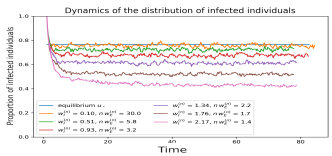
- The limiting dynamics on the trait space
- General interaction kernel

Numerics

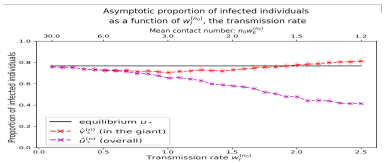
- Discrepancy with the limit
- Very sparse interactions

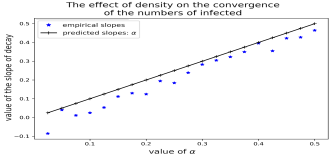
Perspectives

- More general infection history $\Rightarrow$ adapt the coupling approach to extend other mean-field scenarii
- A more elaborated graph structure ?
- Optimize the design of regulation strategies
- Functional Central Limit Theorem
- Investigation of the heterogeneity in the sparse situation



I thank you for your kind attention !





References 1/2

- Billiard, S., Leman, H., Rey, T., Tran, V.C. ; Continuous limits of large plant-pollinator random networks and some applications ; MathematicS In Action, special issue Mathématiques et Biologie. V. Bansaye, E. Kuhn and Ph. Moireau eds.
- Decreusefond, L., Dhersin, J-S., Moyal, P., Tran, V.C. ; Large graph limit for an SIR process in random network with heterogeneous connectivity ; Ann. Appl. Probab., V.22, N.2, pp. 541 - 575 (2012)
- Delmas, J.F., Dronnier, D., Zitt, P.A. ; An Infinite-Dimensional SIS Model, J. Diff. Eq., V.313, pp.1-53 (2022)

References 2/2

- Britton, T., Pardoux, E. ; Stochastic epidemic models with inference, Lectures notes in Mathematics 2555, Springer (2019)
- Bollobas, B., Janson, S., Riordan, O. ; The phase transition in inhomogeneous random graphs ; Random Structures & Algorithms, V.31, N.1, pp.3–122 (2007)
- Junyu Cao and Mariana Olvera-Cravioto ; Connectivity of a general class of inhomogeneous random digraphs. Random Structures & Algorithms, V.56, pp.722–774 (2020)

Appendix

Summary

Coupling objective

Coupling procedure

Under the same assumptions as for our Main Result, there exists a coupling with the following relation between the processes $(\eta^{(n)})_n$ obeying rule $(G1)$ with resp. $(\gamma^{(n)}, w_E^{(n)}, w_I^{(n)})$ and the processes $(\widetilde{\eta}^{(n)})_n$ obeying rule $(G0)$ with resp. $(\gamma^{(n)}, n \cdot w_E^{(n)} \cdot w_I^{(n)})$, with the same initial condition $(\eta_0^{(n)})_n$:

$$\mathbb{E}\left(\sup_{\{t\leq T\}} \left\|\eta_t^{(n)} - \widetilde{\eta}_t^{(n)}\right\|_{TV}\right) \leq C\, A_1^{(n)}(w_I^{(n)}).$$

Coupling techniques

- Selction procedure of the active edges :
  Sample $(V_{(i,j)})_{1\leq i<j\leq n}$ independent uniform on $[0,1]$.

$$i \sim j \text{ if } V_{(i,j)} \leq w_E^{(n)}(X_i, X_j).$$

- Dynamics of $\widetilde{\eta}^{(n)}$: graph structure forgotten
  after each infection event, then resampled.
- independence structure between the edges :
  crucial for the coupling.
  $V_{(i,j)}$ into account in the dynamics of $\widetilde{\eta}^{(n)}$
  on the "first" stimulation of edge (i,j).

Localization techniques

Through this coupling :

- discrepancies between $\eta^{(n)}$ and $\widetilde{\eta}^{(n)}$ can be generated each time an edge (i,j) is selected for the second time by $Q_I^{(n)}$ potentially changing the state of the target individual differently.
- then required that either $V_{(i,j)}$ or v is less than $w_E^{(n)}(X_i, X_j)$.
- differences propagate to other sites each time an edge activated by $Q_I^{(n)}$ involves a source individual j whose state is distinct between the two processes.
- the only two mechanisms by which new sites may differ between $\eta^{(n)}$ and $\widetilde{\eta}^{(n)}$.

Localization techniques

We will localize these perturbations via the process $(\xi_t^{(n)})$ defined as follows, measure-valued on $[\![1,n]\!]$ and increasing in time.

$$\xi_t^{(n)}(dm) = \iint^{(t)} \delta_i(dm) \cdot 1_{\left\{\xi_{s-}^{(n)}(\{i\})=0\right\}} \cdot \Big[\big(1_{\{N_{s-}(i,j)\geq 1\}} \cdot 1_{\left\{V_{(i,j)} \wedge v \leq w_E^{(n)}(X_i,X_j)\right\}}\big) \vee \big(1_{\left\{\xi_{s-}^{(n)}(\{j\})=1\right\}} \cdot 1_{\left\{v \leq w_E^{(n)}(X_i,X_j)\right\}}\big)\Big] d\widehat{Q}_I^{(n)}.$$

Crucial properties of $\xi_t^{(n)}$

$\xi_t^{(n)}$ is equal to 1 on its support $\Xi_t^{(n)} := Supp(\xi_t^{(n)})$.
For all $t > 0$, for all $i \notin \Xi_t^{(n)}$, $E_t^i = \widetilde{E}_t^i$.

Control of the discrepancies

Control of the discrepancies

Under the assumptions of our Main Result,

$$\mathbb{E}(|\Xi_T^{(n)}|/n) = O(r^2 T^2 e^{rT} A_1^{(n)}(w_I^{(n)})).$$

Each increase of $|\Xi_T^{(n)}|$ is decomposed as

- either the creation of a new root
- or a propagation of uncertainty

$$\xi_t^{(n)}(dm) = \iint^{(t)} \delta_i(dm) \cdot 1_{\left\{\xi_{s-}^{(n)}(\{i\})=0\right\}} \cdot \Big[\big(1_{\{N_{s-}(i,j)\geq 1\}} \cdot 1_{\left\{V_{(i,j)} \wedge v \leq w_E^{(n)}(X_i,X_j)\right\}}\big) \vee \big(1_{\left\{\xi_{s-}^{(n)}(\{j\})=1\right\}} \cdot 1_{\left\{v \leq w_E^{(n)}(X_i,X_j)\right\}}\big)\Big] d\widehat{Q}_I^{(n)}.$$

Control of the number of roots

Technical definition of the set $\mathcal{R}_T^{(n)}$ of roots, involving $N_T^{(n)}(i,j) \geq 2$.

$$N_t^{(n)}(i',j') := \iint^{(t)} \left(1_{\{i=i',j=j'\}} 1_{\left\{u \leq w_l^{(n)}(X_i,X_j)\right\}} + 1_{\{i=j',j=i'\}} 1_{\left\{u \leq w_l^{(n)}(X_j,X_i)\right\}}\right) dQ_l.$$

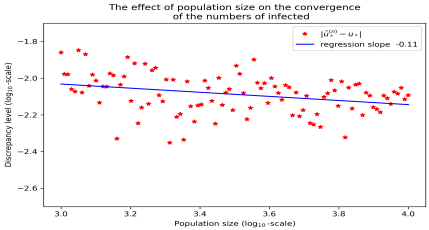
Upper-bound on the average number of roots

Under the assumptions of our Main Result :

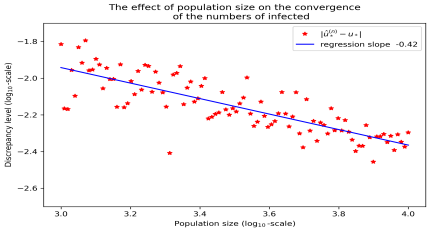
$$\mathbb{E}\left[Card(\mathcal{R}_T^{(n)})\right] \leq 4n\, r\, T^2 A_1^{(n)}(w_l^{(n)}).$$

Summary

Various values of α

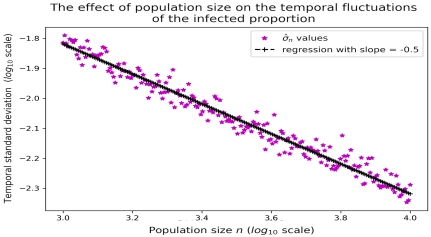


$\alpha = 0.1.$

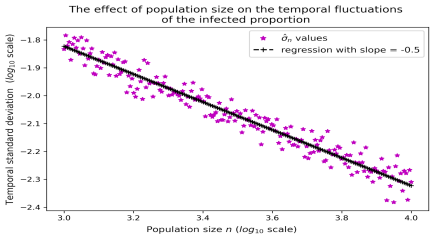


$\alpha = 0.45.$

Temporal fluctuations



$\alpha = 0.3.$



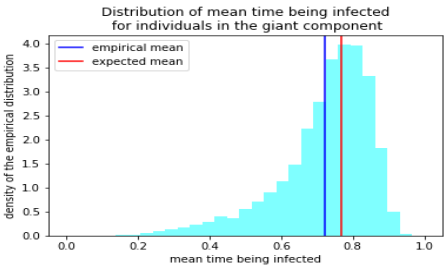
$\alpha = 0.$

Temporal standard deviation $\hat{\sigma}_n$ of the proportion of infected individuals, and comparison with the decay $1/\sqrt{n}$, when $w_I^{(n)} \asymp n^{-\alpha}$. Each star point is obtained from a single run.

Typical infection duration

- The average infection duration for a given individual is obtained by considering the proportion of observations, over the time interval $[20, 80]$,
  where the individual is notified as infected.
- I validated that the time grid evolves in a regular way
  (it is given by a sampling of a sequence with very many exponential inter-times).
- The average over the individuals should compare
  to the proportion of infected individuals
  (provided it is in equilibrium between each time step).

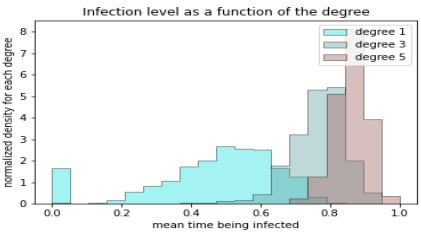
Distribution of infection duration



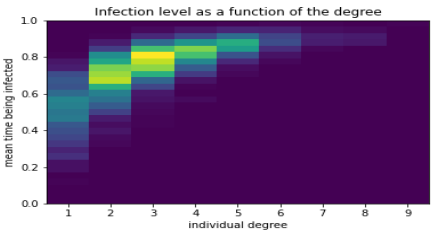
For these simulations, a clear difference is visible between the empirical mean and the expected one depending on the equilibrium value of the graphon SIS.
The effect of belonging to the giant component is already taken into account in the histogram.

Distribution of infection duration, depends on the (small) degrees

This infection duration strongy depends on the degree, although this information is not complete.



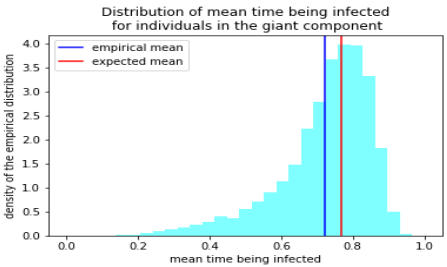
for different degrees



2D Histogram

(figure obtained with $w_I = 1.13$)

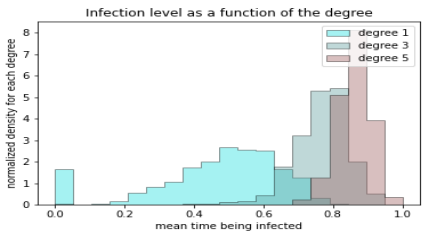
Distribution of infection duration



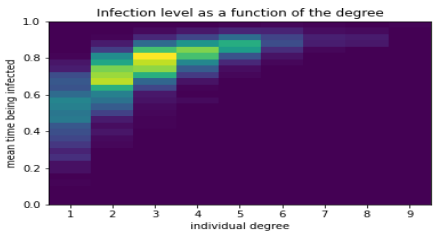
For these simulations, a clear difference is visible between the empirical mean and the expected one depending on the equilibrium value of the graphon SIS.
The effect of belonging to the giant component is already taken into account in the histogram.

Distribution of infection duration, depends on the (small) degrees

This infection duration strongy depends on the degree, although this information is not complete.



for different degrees



2D Histogram

(figure obtained with $w_I = 1.13$)