

Supplementary Information for the paper:
The nature of epidemic thresholds in networks

Marian Boguñá

*Departament de Física Fonamental, Universitat de Barcelona,
Martí i Franquès 1, 08028 Barcelona, Spain*

Claudio Castellano

*Istituto dei Sistemi Complessi (ISC-CNR),
via dei Taurini 19, I-00185 Roma, Italy and
Dipartimento di Fisica, “Sapienza” Università di Roma, P.le A. Moro 2, I-00185 Roma, Italy*

Romualdo Pastor-Satorras

*Departament de Física i Enginyeria Nuclear,
Universitat Politècnica de Catalunya,
Campus Nord B4, 08034 Barcelona, Spain*

(Dated: July 23, 2013)

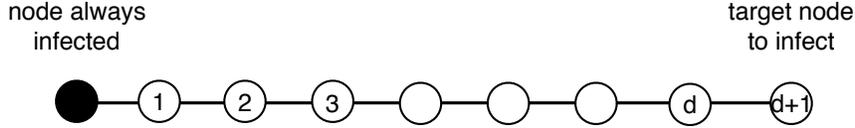


FIG. SI-1: Scheme of the infectious events mediated by chains of nodes.

I. NUMERICAL SIMULATIONS

The SIS dynamics is simulated with a continuous time dynamics as follows: During the course of the simulation, we keep track of the number of infected nodes $N_I(t)$ and the number of active links $E_A(t)$, where an active link is defined as a link emanating from an infected node (notice that links connecting two infected nodes will appear twice in this list). At each step, with probability $p_r = N_I(t)/[N_I(t) + \lambda E_A(t)]$, a randomly chosen infected node is turned susceptible whereas, with probability $1 - p_r$, an active link is chosen at random and if one of the two nodes attached to the link is susceptible, then this node is turned infected. After this procedure, time is updated as $t \rightarrow t + 1/[N_I(t) + \lambda E_A(t)]$ and the list of infected nodes and active links recomputed. An equivalent algorithm keeps a list of active links as those connecting one infected node and one susceptible. The advantage of this latter method is that infectious attempts always end up with a susceptible node being infected, which is not the case with the first method. In this work, all simulations are backed up independently with the two methods.

II. ESTIMATION OF THE INFECTIVE RATE $\bar{\lambda}(d_{ij}, \lambda)$

This rate is the inverse of the average time infected node i takes to infect node j when they are separated by a distance d_{ij} in the original graph. We consider the process of transmission of the infection between the two nodes mediated by a one dimensional chain of length d_{ij} . Consider a one dimensional chain of $d + 2$ nodes with the leftmost node always infected, as indicated in Fig. SI-1. Let $T(d + 1)$ the average time the node at the rightmost position takes to get infected for the first time. Let $\tilde{T}(d)$ the average time between two consecutive infectious events (after the first one) of the node at distance d . Because we are in a chain and the source of the infection is the leftmost node, the node at distance $d + 1$ can only get infected for the first time by its left neighbor. Once this node is infected, the probability that the target node gets the infection before

its left neighbor recovers is simply given by

$$p = \frac{\lambda}{1 + \lambda} \quad (\text{SI-1})$$

the node at distance $d + 1$ can get infected right after its left node gets infected for the first time or after the second time, and so on. The probability that the node gets infected right after its left neighbor gets infected for the n -th time is

$$\text{Prob}(n) = p(1 - p)^{n-1} \quad (\text{SI-2})$$

On the other hand, the average time elapsed in this case $T_n(d + 1)$ is

$$T_n(d + 1) = T_n(d) + (n - 1)(1 + \tilde{T}(d)) + \lambda^{-1}. \quad (\text{SI-3})$$

Combining these two results, we get the equation for the average infection time, $T(d) = \sum_n \text{Prob}(n)T_n(d)$,

$$T(d + 1) = T(d) + \frac{2}{\lambda} + \frac{1}{\lambda}\tilde{T}(d) \quad (\text{SI-4})$$

with the initial conditions $T(1) = \tilde{T}(1) = \lambda^{-1}$. In the limit of low infectious rate, that is, $\lambda \ll 1$, we can approximate $\tilde{T}(d) \approx T(d)$ and we get a closed recursive equation for $T(d)$, whose solution is

$$T(d) = \frac{1}{\lambda} \left[(1 + 2\lambda) \left(1 + \frac{1}{\lambda}\right)^{d-1} - 2\lambda \right] \approx \frac{1}{\lambda} e^{(d-1)\ln(1+\frac{1}{\lambda})} \quad (\text{SI-5})$$

We then conclude that the infective rate is

$$\bar{\lambda}(d_{ij}, \lambda) \approx \lambda e^{-b(\lambda)(d_{ij}-1)}, \text{ with } b(\lambda) = \ln \left(1 + \frac{1}{\lambda}\right) \quad (\text{SI-6})$$

In the case of small-world random graphs, the average internode topological distance depends only on the degree of the nodes as [2]

$$d_{k,k'} = 1 + \frac{\ln \left(\frac{N\langle k \rangle}{kk'}\right)}{\ln \kappa}. \quad (\text{SI-7})$$

Inserting this expression into the effective infective rate we get

$$\lambda_{k,k'} = \lambda \left[\frac{kk'}{\langle k \rangle N} \right]^{\frac{b(\lambda)}{\ln \kappa}}. \quad (\text{SI-8})$$

We test numerically this relationship by keeping a single node of degree k always infected and computing the time it takes to infect for the first time any other node in the network. The rate $\lambda_{k,k'}$ is obtained by inverting the value of this time for the first infectious event, averaged over all nodes of degree k' . Equation (SI-8) predicts that plotting $[\lambda_{k,k'}/\lambda]^{\ln \kappa/b(\lambda)}$ vs k' a linear behavior must be found, and this turns out to agree with the outcome of simulations in a network with $\gamma = 3.5$ (see Fig. SI-2).

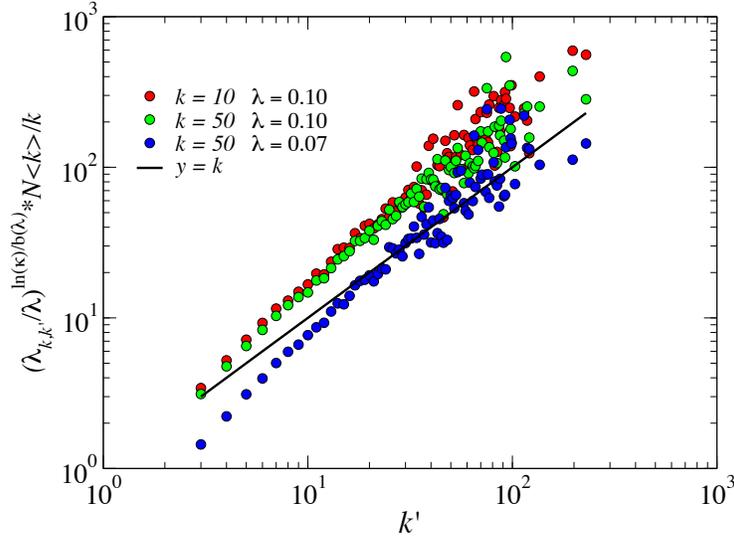


FIG. SI-2: Plot of $[\lambda_{k,k'}/\lambda]^{\ln \kappa/b(\lambda)} \cdot N\langle k \rangle/k$ vs k' for various values of λ and k . The prediction of Eq. (SI-8) is the straight solid line.

III. ESTIMATION OF THE RECOVERY RATE $\delta(k, \lambda)$

This rate can be estimated as the inverse of the survival time of an infection starting at the center of a star of degree k . Unfortunately, the exact solution to this problem is rather involved (see [3] for an exact treatment). Here, we present an approximation based on the discretization of the process in time units of $\mu^{-1} = 1$. Consider the following cycle: initially, the center of the star –the hub– is infected whereas leaf nodes are susceptible. The probability that a leaf node is infected when the hub recovers is

$$p_{in} = \frac{\lambda}{2 + \lambda}. \quad (\text{SI-9})$$

Then, by the time the hub recovers, there are n infected leaf nodes with probability

$$\text{Prob}(n|k) = \binom{k}{n} p_{in}^n (1 - p_{in})^{k-n} \quad (\text{SI-10})$$

the probability that at least one of these n infected nodes infects the hub again before they recover is

$$\sum_{n=1}^k \text{Prob}(n|k) [1 - (1 - p_{in})^n] = 1 - (1 - p_{in}^2)^k. \quad (\text{SI-11})$$

The average time to complete the cycle is 2. The probability that the outbreak goes through a sequence of m complete cycles and then dies is

$$(1 - p_{in})^k [1 - (1 - p_{in}^2)^k]^m; m = 0, 1, \dots \quad (\text{SI-12})$$

and the time elapsed $(2m + 1)$. However, an outbreak can also die in the middle of the cycle, that is, when infected leaves recover before infecting the hub again. The probability that the outbreak goes through a sequence of m complete cycles and dies in the middle of the $m + 1$ cycle is

$$[(1 - p_{in}^2)^k - (1 - p_{in})^k] [1 - (1 - p_{in}^2)^k]^m; m = 0, 1, \dots \quad (\text{SI-13})$$

The average elapsed time is in this case $(2m + 2)$. Putting these pieces together, the effective recovery rate can be approximated as

$$\delta^{-1}(k, \lambda) = \frac{2}{(1 - p_{in}^2)^k} - \frac{1}{(1 + p_{in})^k}. \quad (\text{SI-14})$$

For low infectious rates $p_{in} \ll 1$, this can be approximated as

$$\delta^{-1}(k, \lambda) \approx 2e^{-\lambda^2 k/4}. \quad (\text{SI-15})$$

As we have mentioned at the beginning of this section, the previous calculations provide only an approximation to the true recovery rate. This is so because we have considered the process as discretized in time whereas the real process takes place at continuous time. Nevertheless, we expect that Eq. (SI-15) captures the correct functional dependence. To check this result, we have performed simulations of the SIS model on star graphs, starting from a state with only the hub infected and computed the average time needed to reach the absorbing healthy state. Fig. SI-3 shows the average lifetime for fixed values of λ as a function of the degree k , where the exponential trend predicted by our calculations is clearly visible.

IV. DERIVATION OF EQ. (4)

Let us consider Eq. (3) in the main paper, namely

$$\frac{d\rho_k(t)}{dt} = -\bar{\delta}(k, \lambda)\rho_k(t) + \lambda N \left[\frac{k}{N\langle k \rangle} \right]^{\frac{b(\lambda)}{\ln \kappa}} \sum_{k'} k'^{\frac{b(\lambda)}{\ln \kappa}} P(k')\rho_{k'}(t)[1 - \rho_k(t)]. \quad (\text{SI-16})$$

It is obvious that the absorbing state $\rho_k(t) = 0$ is a fixed point of the dynamics. Therefore, we conclude that an endemic state exists whenever the solution $\rho_k = 0$ is dynamically unstable.

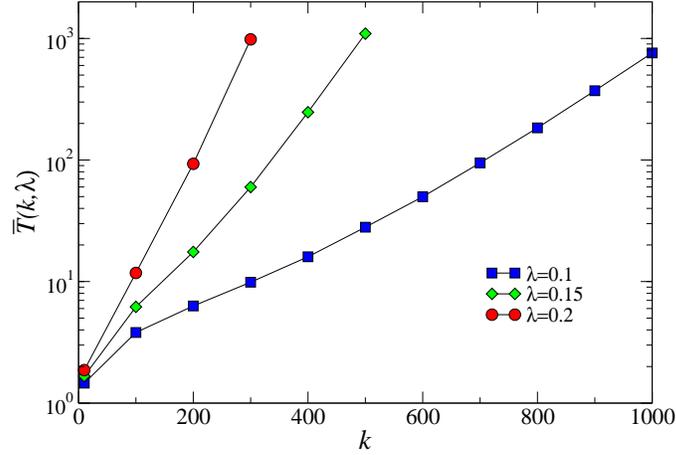


FIG. SI-3: Average lifetime of star graphs with the hub originally infected as a function of the star degree and different values of λ .

Following this idea, we linearize Eq. (SI-16) around $\rho_k = 0$, i.e.,

$$\frac{d\rho_k(t)}{dt} \approx -\bar{\delta}(k, \lambda)\rho_k(t) + \lambda N \left[\frac{k}{N\langle k \rangle} \right]^{\frac{b(\lambda)}{\ln \kappa}} \sum_{k'} k'^{\frac{b(\lambda)}{\ln \kappa}} P(k')\rho_{k'}(t). \quad (\text{SI-17})$$

This equation can be written in matrix form as

$$\frac{d\rho_k(t)}{dt} \approx \sum_{k'} \mathbb{M}_{kk'} \rho_{k'}(t), \quad (\text{SI-18})$$

where

$$\mathbb{M}_{kk'} = -\bar{\delta}(k, \lambda)\delta_{kk'} + \lambda N \left[\frac{k}{N\langle k \rangle} \right]^{\frac{b(\lambda)}{\ln \kappa}} k'^{\frac{b(\lambda)}{\ln \kappa}} P(k'). \quad (\text{SI-19})$$

The stability of the absorbing state is then controlled by the maximum eigenvalue of matrix \mathbb{M} , that is, the maximum Λ_m solution of the eigenvalue problem $\mathbb{M}\vec{u} = \Lambda\vec{u}$. In this way, $\Lambda_m = 0$ defines the threshold between the absorbing and endemic phases. The eigenvalue problem can be rewritten as

$$\frac{\lambda N}{\Lambda + \bar{\delta}(k, \lambda)} \left[\frac{k}{N\langle k \rangle} \right]^{\frac{b(\lambda)}{\ln \kappa}} \sum_{k'} k'^{\frac{b(\lambda)}{\ln \kappa}} P(k')u_{k'} = u_k, \quad (\text{SI-20})$$

which gives us the explicit dependence of u_k on k . Using this result, the eigenvalues satisfy the equation

$$\sum_k P(k) \frac{\lambda N}{\Lambda + \bar{\delta}(k, \lambda)} \left[\frac{k^2}{N\langle k \rangle} \right]^{\frac{b(\lambda)}{\ln \kappa}} = 1. \quad (\text{SI-21})$$

By setting $\Lambda = 0$ and recalling that $\tau(k, \lambda) = \bar{\delta}(k, \lambda)^{-1}$, we recover Eq. (4) in the main paper.

V. NON SMALL-WORLD NETWORKS

Let us assume that the average distance is given by

$$d = 1 + \alpha N^\beta. \quad (\text{SI-22})$$

Plugging this expression in Eq. (1) in the main paper, we obtain a coarse description of the dynamics as

$$\frac{d\rho_k(t)}{dt} = -\bar{\delta}(k, \lambda)\rho_k(t) + \lambda N e^{-\alpha N^\beta \frac{b(\lambda)}{\ln \kappa}} \sum_{k'} P(k')\rho_{k'}(t)[1 - \rho_k(t)]. \quad (\text{SI-23})$$

By repeating the same analysis performed in the previous section, we conclude that the critical infection rate satisfies the equation

$$1 = \lambda N e^{-\alpha N^\beta \frac{b(\lambda)}{\ln \kappa}} \sum_k P(k)\tau(k, \lambda). \quad (\text{SI-24})$$

Assuming again that $P(k)$ decays slower than an exponential, this equation can be approximated as

$$1 = \frac{\lambda}{a(\lambda)} e^{a(\lambda)k_{max} - \alpha N^\beta \frac{b(\lambda)}{\ln \kappa} - \ln\left[\frac{1}{NP(k_{max})}\right]}. \quad (\text{SI-25})$$

From this equation it is easy to see that for any fixed value of λ , if k_{max} grows faster than N^β , there exists a size N such that the exponent in this equation starts growing with the system size and, therefore, the right hand side in this equation will eventually grow above 1. The logical consequence is that, in this case, the epidemic threshold goes to zero as N goes to infinity.

VI. ROBUSTNESS WITH RESPECT TO THE COVERAGE THRESHOLD C

To determine whether a given realization of the SIS process is endemic, we have used the condition that the coverage is larger than a fixed threshold value $C = 0.5$. To check that different assumptions do not qualitatively change the results, we performed some numerical tests. In Fig. SI-4, we consider the effect of changing C for an Erdős-Rényi graph of average degree $\langle k \rangle = 5$ and different sizes N . The numerical estimate of the threshold rapidly converges to the expected value $\lambda_c = 1/5$ for both values of C considered, while the height of the peak grows with an exponent independent of C . In Fig. SI-5, we perform the same analysis for a UCM graph with $\gamma = 3.5$ and $k_{min} = 3$, obtaining similar results. As the system size N is increased, the estimated thresholds decrease and the peak heights increase in a perfectly analogous way. Both figures confirm that the behavior of the model is robust with respect to the arbitrary choice of the coverage threshold C .

VII. FURTHER CHARACTERIZATION OF THE EPIDEMIC TRANSITION

In this section we provide some additional insight into the transition marked by the peak of the average lifetime of finite realizations $\bar{T}(\lambda, N)$. In Fig. SI-6, we compare the curves already plotted in the top left of Fig. 1 of the main paper with the analogous curves computed for a star graph with k_{max} leaves. It is clear that the occurrence of the peak in the full network is not due only to the star graph centered around its hub. The latter sustains alone the activity only for small λ . The transition occurs at higher values of λ , for which the lifetime is exceedingly larger.

In Fig. SI-6 we plot, as a function of N , the value of the average coverage of the whole network at the critical value $\lambda_{max}(N)$ and at the critical value $\lambda_{max}^*(N)$ for the star graph centered around the hub of degree k_{max} . It turns clearly out that at the critical point λ_{max} the coverage assumes a finite value in the thermodynamical limit, while it vanishes for $\lambda = \lambda_{max}^*$. For $\lambda = \lambda_{max}^*$ the hub and its neighbors are fully covered, yet the epidemics does not escape from the hub and its neighbors and it is thus localized. At the transition point λ_{max} instead the epidemics leaves the

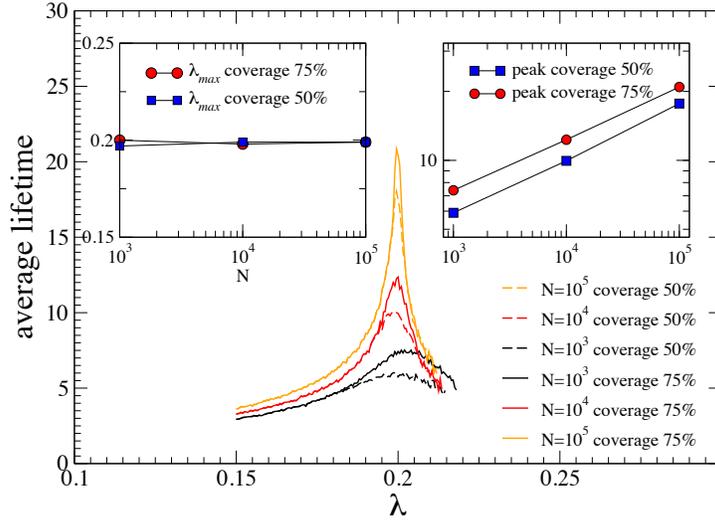


FIG. SI-4: Main: average lifetime $\bar{T}(\lambda, N)$ for finite realizations as a function of λ for the SIS model on an Erdős-Rényi graph of average degree $\langle k \rangle = 5$. Left inset: position of the peak marking the estimate of the numerical threshold as a function of the system size N . Right inset: height of the peak as a function of N .

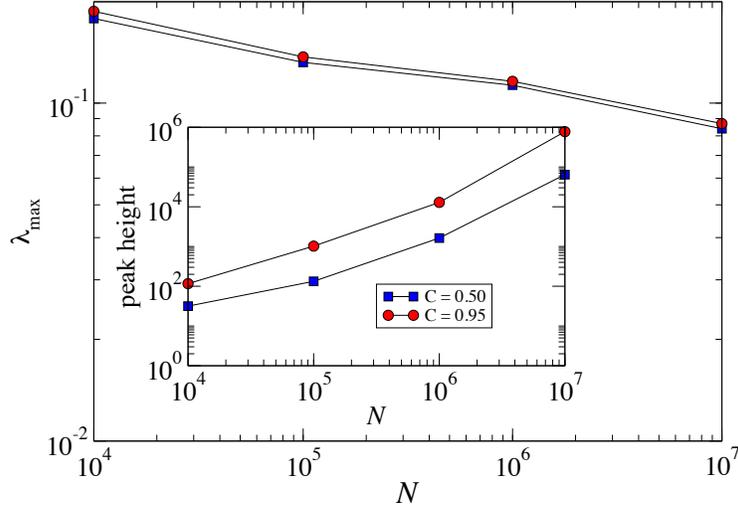


FIG. SI-5: Main: value of the epidemic threshold λ_{max} estimated numerically as the position of the peak of $\bar{T}(\lambda, N)$ as a function of N for a UCM network with $\gamma = 3.5$ and $k_{min} = 3$. Inset: height of the peak of $\bar{T}(\lambda, N)$ for the same system.

hub and affects the whole population, leading to a truly endemic state.

-
- [1] J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge University Press, Cambridge, 1999).
- [2] J. A. Hołyst, J. Sienkiewicz, A. Fronczak, P. Fronczak, and K. Suchecki, *Phys. Rev. E* **72**, 026108 (2005).
- [3] E. Cator and P. Van Mieghem, *Phys. Rev. E* **87**, 012811 (2013).

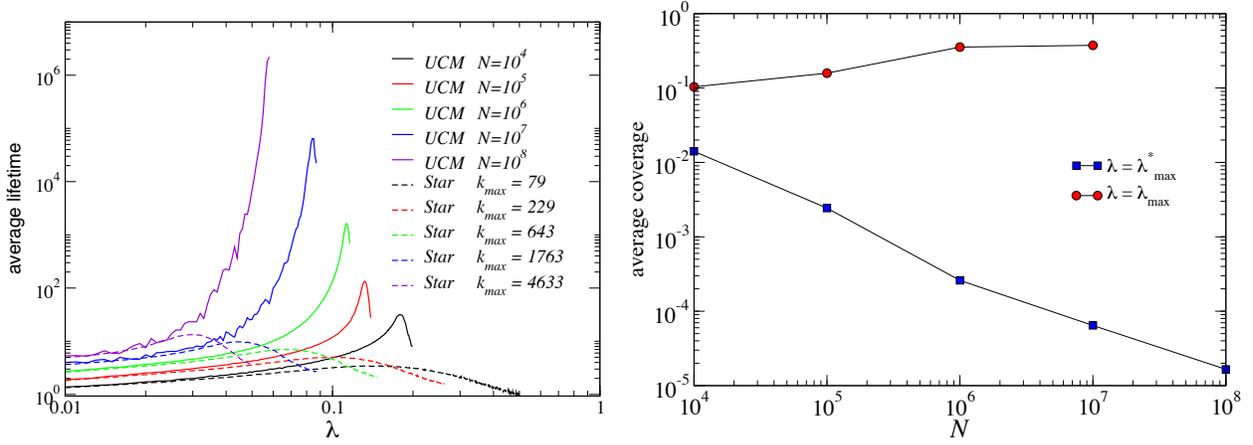


FIG. SI-6: Left: Plot of the average lifetime of finite realizations $\bar{T}(\lambda, N)$ for UCM networks with $\gamma = 3.5$ and $k_{min} = 3$ and increasing values (bottom to top) of the system size N (solid lines). Dashed lines are the same quantity computed for a star graph made of $k_{max} + 1$ nodes, where the values of k_{max} are equal to the largest degree in the UCM networks. Right: Plot of the average coverage for a UCM network with $\gamma = 3.5$ and $k_{min} = 3$ for two different values of λ , corresponding to the thresholds for the whole network (λ_{max}) and for the star graph centered around the largest hub (λ_{max}^*).