COOPERATIVE EPIDEMIC SPREADING IN COMPLEX NETWORKS

Xavier Roderic Hoffmann
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.

Most epidemic spreading models assume memoryless systems and statistically independent infections. Nevertheless, many real-life cases are manifestly time-sensitive and may be strongly correlated. We study the effect of non-Markovian stochastic dynamics on the SIS model, in random and scale-free networks, and propose a novel microscopic description to account for cooperation. Initial exploratory simulations yield promising results, calling for further research.

I. INTRODUCTION

Humanity has been dealing with the burden of disease for a long time. For almost as many years, scientists have attempted to describe the outbreak and spread of illnesses in order to evaluate inoculation or isolation plans and to devise strategies to reduce their mortality rates. Epidemic modeling is one of the main tools used to study the spreading mechanisms, to predict the evolution of an outbreak and to gauge containment protocols [1].

The first systematic epidemiological study dates back to 1662, when John Grant tried to quantify the causes of death of various diseases [1]; his work, however, did not contemplate their transmission. This would have to wait until 1766, when Daniel Bernoulli created a mathematical model for the spreading of smallpox, concluding that universal inoculation could significantly increase life expectancy [2, 3]. These and other efforts preceded the development during the 19th century of germ theory, which marked the beginning of modern theoretical epidemiology. The comprehensive understanding of the biology behind epidemic outbreak and spreading allowed for more sophisticated and accurate modeling [4], ultimately culminating in the seminal works by Kermack and McKendrick (1927, 1932 and 1933).

Kermack and McKendrick's modeling scheme can be summarized as follows: i) divide the population into discrete compartments (or states), ii) translate all biological properties of the disease into mathematical parameters and iii) specify the rules that govern transitions between states [5]. The possibility of including a wide array of factors (e.g. age, birth, death, migration, immunity, vaccination, etc.) has yielded a vast variety of models, often quite sophisticated [6]. Because of its simplicity and versatility it is still the starting point for most of current research in the field of epidemic modeling.

Traditionally the dynamics of disease spreading were described by deterministic differential equations. With the use of tools borrowed from statistical mechanics, scientists were able to introduce stochastic features, further elaborating their models [7]. Later studies showed that the structure of human interactions was central to the problem and networks were brought into play [8, 9]. Treating network dynamics analytically is often quite complicated and sometimes even impossible. Epidemic modeling, therefore, has resorted to computational simulation, now an essential tool in any research effort.

Recently, technological improvements have enhanced computational performance substantialy, permitting highly precise, big-scale simulations. Moreover, the widespread use of mobile and wifi technologies throughout society has enabled the acquisition of detailed information about real-life networks. These advances have prompted many researchers to advocate the use of models as real-time predictive tools, specially in the domain of public health policy [10, 11]. Furthermore, epidemic modeling can be extended to describe other phenomena, such as the spreading of information, rumors, cultural norms and social behavior [12].

This essay is limited to the study of the susceptible-infectious-susceptible (SIS) model. Section II reviews the basic definitions and terminology, describes the results for infinite and limited contact and discusses the computational approach. The effects of equipping the system with memory are explored in Section III, presenting first an algorithm to simulate non-Markovian processes, then introducing heterogeneous transmissions and finally allowing infectors to cooperate.

II. STANDARD SIS MODEL

A. Elementary features

Consider a population of N members, each of which can be in one of two states: infectious (I) or susceptible (S), prone to be infected. The traditional mass-action approximation ignores the structure of contact networks and assumes that each individual has an equal chance, per unit time, of coming into contact with every other. In this case the disease spreads randomly with a per-individual rate β, meaning that each healthy agent has, on average, β contacts with randomly chosen others per unit time. Nevertheless, the infection is transmitted only when a susceptible person has contact with an infected one. On the other hand, disease stricken agents recover spontaneously at a constant average rate η.

This model does not account for birth, death nor migration, i.e. the population remains fixed. Furthermore, recovery from the illness does not confer any long lasting

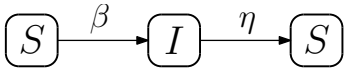

Fig. 1: Flow chart for the SIS model, indicating the transitions between states and their respective rates.

immunity. Some sexually transmitted infections, such as gonorrhea, present these characteristics [7].

The basic macroscopic measurement is the density of infected individuals, $\rho(t) = N_I(t)/N$, also called the prevalence of the disease. Given an initial number of infected agents N_0, and after passing through a transient stage, the system evolves towards a stationary state. The late-time value of the prevalence, ρ_∞, depends solely on the spreading ratio, $\lambda = \beta/\eta$. This behavior is typically represented in a phase diagram, with control parameter λ and order parameter ρ_∞, and presents a continuous phase transition at $\lambda_c = 1$. When $\lambda < \lambda_c$ the outbreak dies out exponentially fast and all agents are disease-free ($\rho_\infty = 0$); this phase is called absorbing or healthy. For $\lambda > \lambda_c$ there is a non-vanishing fraction of infected agents ($\rho_\infty > 0$), defining an active or endemic phase. In the jargon of epidemiology, the critical point λ_c receives the name of epidemic threshold [8].

B. Network topologies

The fully-mixed approach assumed in the previous section is obviously not a realistic representation of the workings of our world. In real life, most people have a set of acquaintances, neighbors, coworkers, etc. whom they meet with some regularity and all other members of the population can safely be ignored. This contact structure can be represented as a network and its topology has a strong effect on the spreading of diseases. The recent availability of huge quantities of data-sets has enabled scientists to classify real-world networks according to their statistical characteristics. In addition, a variety of models have been developed to recreate their properties. Our study is limited to two of these models: random and heavy-tailed networks.

A network is basically a collection of M links between N nodes, node j having degree k_j, i.e. it is connected to k_j other nodes. The structure of these connections is encoded in the degree distribution $P(k)$. Random networks are usually typified by N and the average degree $\langle k\rangle$. Two nodes are connected with probability $p = \langle k\rangle/(N-1)$, giving an expected number of links $\langle M\rangle = \langle k\rangle N/2$. In the limit $N \gg 1$ and constant $\langle k\rangle$, the degree distribution can be approximated by a Poisson function, $P(k) = e^{-\langle k\rangle}\langle k\rangle^k/k!$, reflecting the global homogeneity of the network. On the other hand, heavy-tailed or scale-free networks have a degree distribution $P(k) \sim k^{-\gamma}$, with typically $2 < \gamma < 3$. This endows them with a much richer topology, presenting interesting features such as local clustering and the existence of hubs. A usual method of generating such networks is the configuration model [13].

Modeling epidemic spreading on networks is based on the same principles as for the mass-action approximation, the main difference being that contact between agents is now limited. Therefore the transmission rate β is redefined as the probability per unit time that the disease transmits from an infected individual to a susceptible one, given they are neighbors. The behavior of ρ (if unambiguous, hereupon we drop the subindex $_\infty$) is similar to the fully-mixed case, although the value of λ_c varies depending on the network topology. Analytical mean-field studies show that $\lambda_c = 1/\langle k\rangle$ for random networks [14] and vanishes ($\lambda_c \to 0$) in scale-free networks [15].

C. Markovian stochastic simulations

The state of the system changes when an infected individual recovers or when a susceptible agent is infected, yielding a sequence of events that constitute a mixture of temporal point processes. These processes can be assumed memoryless, with future occurrences predicted only on the system's present state. The inter-event time distribution for process j is then $\psi_j(\tau) = \upsilon_j e^{-\upsilon_j \tau}$, with υ_j its constant occurrence rate. This means that the next event of process j will occur after an interval of length τ with probability $\psi_j d\tau$. The global dynamics can be simulated using Markovian stochastic algorithms able to generate statistically exact realizations, such as the seminal method developed by Gillespie [16, 17].

At a certain time t the system counts N_I infected agents, each representing a possible recovery event. All these recoveries are equivalent, i.e. they have the same rate η. Hence, infected nodes recover, on average, after $\langle\tau\rangle_{\text{rec}} = 1/\eta$ units of time. To count the number of possible infection events N_A, the concept of active link is introduced, i.e. a link between an infected individual and a susceptible one. Since healthy agents may have none or more than one infected neighbors, in general $N_A \neq N_S$, where N_S is the number of susceptible agents. Once again, all active links are equivalent and transmit the infection with the same rate β, therefore the average transmission time is $\langle\tau\rangle_{\text{tra}} = 1/\beta$.

The goal is to find the time until the next event τ and whether this event is a recovery or a transmission. Exploiting the properties of Poisson point processes, in the limit $N \gg 1$ the Gillespie algorithm gives $\tau = \Omega^{-1}$, $\Pi_{\text{rec}} = \eta N_I \Omega^{-1}$ and $\Pi_{\text{tra}} = \beta N_A \Omega^{-1}$ with $\Omega = \eta N_I + \beta N_A$. At each iteration two uniform random numbers are sampled, $u_1, u_2 \in U(0,1)$. If $u_1 < \Pi_{\text{rec}}$, u_2 is used to choose the infected node that recovers; otherwise the latter draws the active link that transmits. Before moving on to the next iteration, the system's state is updated, as is the time, $t \to t + \tau$.

We ran the algorithm on a random network with $N = 9926$ and $\langle k\rangle = 5.05$, and on a scale-free network with $N = 10^4$, $\gamma = 2.4$ and $\langle k\rangle = 4.93$. Without loss of generality we used $\eta = 1$, thus $\beta = \lambda$. Our simulations started at $\lambda = 1$, initially infecting all nodes. The system was allowed to relax towards the steady state during the first $\Delta t = 50$ units of time, after which the prevalence was averaged over another interval Δt. Then we decreased the spreading ratio with $\Delta\lambda = 0.05$ and repeated the relaxing-averaging procedure. If the difference in prevalence for consecutive values of λ exceeded $\Delta\rho = 0.025$ we estimated the slope and decreased $\Delta\lambda$

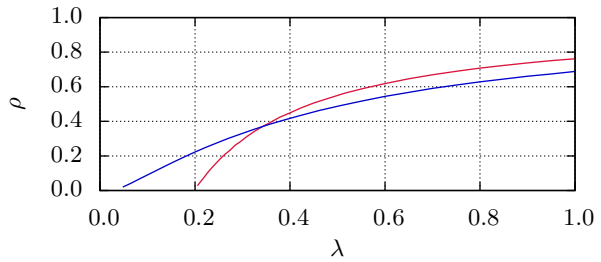

Fig. 2: Markovian stochastic simulation. Phase diagrams for the random (red) and scale-free (blue) networks described in II C.

accordingly; in addition Δt was increased by 20%. The simulation was halted as soon as the system reached the absorbing state. Figure 2 shows the results.

As usual, a computational approach has its limitations. The most notable effect is the enhancement of the stochastic fluctuations due to the system's finite size, specially near the critical point. This may cause the system to become trapped in an absorbing state, even for values of λ well above the epidemic threshold. Although the scale-free network is markedly affected by this phenomenon, overall the results are in good agreement with theoretical values [14, 15].

III. MEMORY ENDOWED SYSTEMS

A. Non-Markovian stochastic simulations

In order to equip the system with memory, an algorithm capable of simulating non-Markovian processes is needed. For this study we have chosen the generalized non-Markovian Gillespie algorithm introduced in [18].

Consider a set of N_T statistically independent discrete stochastic processes, each with an inter-event time distribution $\psi_j(\tau)$. At a certain moment in time t, process j last occurred t_j units of time ago. Let $\phi(\tau, i|\{t_j\})$ denote the joint probability that the next event taking place occurs at time $t+\tau$ and corresponds to process i, conditioned by the set of elapsed times $\{t_j\}$. This probability density can be expressed as

$$\phi(\tau, i|\{t_j\}) = \frac{\psi_i(\tau + t_i)}{\Psi_i(\tau + t_i)} \Phi(\tau|\{t_j\}) \ , \tag{1}$$

where

$$\Phi(\tau|\{t_j\}) = \prod_{j=1}^{N_T} \frac{\Psi_j(\tau + t_j)}{\Psi_j(t_j)} \tag{2}$$

is the survival probability of τ, i.e. the conditional probability that no event takes places before $t+\tau$, and $\Psi_j(\tau) = \int_\tau^\infty \psi_j(z)\, dz$ the survival probability of process j, i.e. the probability that the time until its next event is longer than τ. Once the occurrence time τ is known, the probability that the next occurring event belongs to process i is given by

$$\Pi(i|\tau, \{t_j\}) = \frac{\omega_i(\tau + t_i)}{\sum_{j=1}^{N_T} \omega_j(\tau + t_j)} \ , \tag{3}$$

with $\omega_j(\tau+t_j) = \psi_j(\tau+t_j)/\Psi_j(\tau+t_j)$ the instantaneous hazard rate of process j.

In the limit $N_T \gg 1$, $\Phi(\tau|\{t_j\})$ is close to zero except when $\tau \sim 0$; assuming $\psi_j(\tau)$ analytical permits an expansion in small τ. Substitution in (2) gives $\Phi(\tau|\{t_j\}) = e^{-\tau\Omega(\{t_j\})}$, where $\Omega(\{t_j\}) = \sum_{j=1}^{N_T} \omega_j(t_j)$, and setting $\tau = 0$ in (3) yields

$$\Pi(i|\{t_j\}) = \frac{\omega_i(t_i)}{\Omega(\{t_j\})} \ . \tag{4}$$

The hypothesis that $\psi_j(\tau)$ is analytical is not always valid. To overcome these singular cases the last event (with $t_{\text{last}} = 0$) is removed from the list of possibly occurring processes, implying that the same event can not happen twice in a row. Although this restriction is not present in the real dynamics, the probability is negligible for sufficiently large systems.

Yet again two uniform random numbers, $u_1, u_2 \in U(0,1)$, are required at each iteration: one to sample the occurrence time, $\tau = -\ln(u_1)/\Omega(\{t_j\})$, and the other to choose a process from the discrete distribution (4).

B. Heterogeneous transmissions

Section II assumed all transmission processes to be equivalent at all times, i.e. the probability of transmitting the infection was the same for all active links at any given instant, regardless of the time that had passed since their activation. It is reasonable to think that, for example, a link that has been active for a long time is more prone to transmitting the infection than a link that has been activated more recently. In this section we explore the effect of time-sensitive transmissions, both on the system's late-time properties as on its dynamics.

While recoveries are still exponentially distributed with rate η, inter-event times for transmissions are now governed by a Weibull distribution, $\psi_{\text{tra}}(\tau) = \alpha\mu^\alpha \tau^{\alpha-1} e^{-(\mu\tau)^\alpha}$. This density is chosen because of its versatility: for $\alpha > 1$ it presents a peak, resembling a bell curve, $\alpha = 1$ corresponds to a Poisson distribution and for $\alpha < 1$ it has power-law-like fat tails. The average transmission time is then $\langle\tau\rangle_{\text{tra}} = \mu^{-1}\Gamma(1+\alpha^{-1})$ and the instantaneous transmission rate for active link j

$$\omega_j(t_j) = \alpha\mu^\alpha t_j^{\alpha-1} \ , \tag{5}$$

with t_j the elapsed time since its activation.

An active link A $\to$ B (A infected, B susceptible) can reach this configuration from two different scenarios: i) both nodes were originally healthy and A becomes infected by one of its neighbors other than B or ii) both nodes were originally infected and B recovers. In situation i) it is clear that the active link is new, hence its elapsed time is set to zero, $t_{\text{AB}} = 0$. In situation ii) one can apply the same reasoning and set $t_{\text{AB}} = 0$ (rule 0). However, another may argue that, being A the active infector, the elapsed time is $t_{\text{AB}} = t_{\text{A}}$ (rule A).

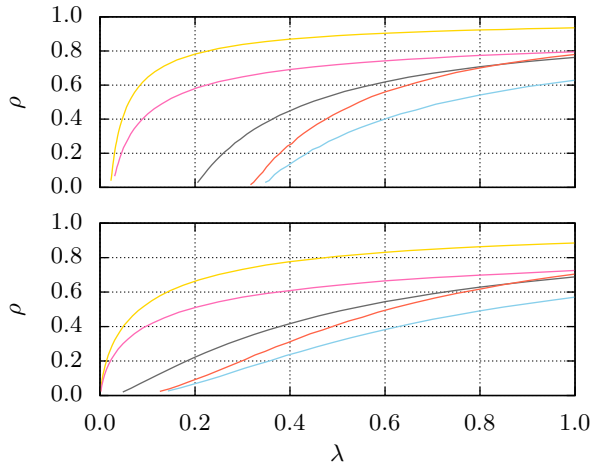

Fig. 3: Heterogeneous transmissions. Phase diagrams for the random (top) and scale-free (bottom) networks described in II C. $\alpha = 0.5$, rule 0 (yellow) and rule A (pink); $\alpha = 1.5$, rule 0 (blue) and rule A (orange); Markovian (gray).

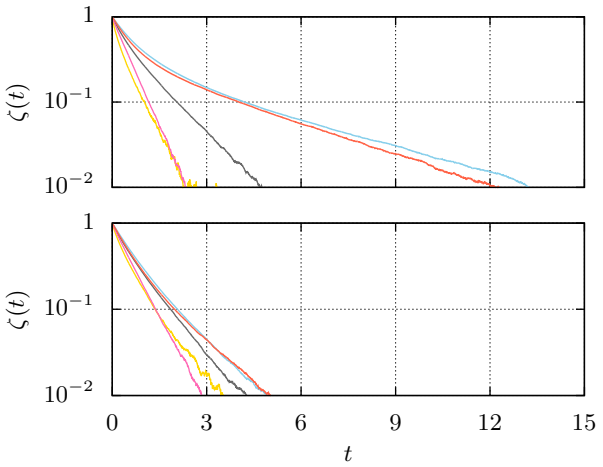

Fig. 4: Heterogeneous transmissions. Dynamics for the random (top) and scale-free (bottom) networks described in II C, for $\lambda = 0.4$. Color code indicated in Fig. 3.

The spreading ratio is redefined as $\lambda = \langle\tau\rangle_{\rm rec}/\langle\tau\rangle_{\rm tra}$, from which $\mu = \eta\lambda\Gamma(1+\alpha^{-1})$. This definition is consistent with the one given for the Markovian case.

Simulations were run on the networks and following the procedure described in Section II C, for $\alpha = 0.5$ and $\alpha = 1.5$, using both rules. As reported in [18] and shown in Figure 3, non-Markovian transmissions significantly modify the value of the critical point in the random network. The most notable effects on the scale-free network are i) a steeper approach to the critical point for $\alpha < 1$ and ii) the possible existence of a non-vanishing epidemic threshold for $\alpha > 1$.

There are also important differences between the two rules, for the same values of α. When $\alpha < 1$ the prevalence for rule 0 is always above that for rule A, and vice versa for $\alpha > 1$. This is evident from the rate (5), a decreasing (increasing) function of t_j when $\alpha < 1$ ($\alpha > 1$), therefore the average transmission period, conditioned by the elapsed time, for rule 0 is shorter (longer) than in the case of rule A, yielding a higher (lower) prevalence.

In order to further elucidate the distinctness between mechanisms we analyzed the system's evolution towards the steady state. With $\lambda = 0.4$, all nodes were initially infected and the system relaxed during $\Delta t = 25$ units of time; results were averaged over 100 independent runs. Figure 4 shows the reduced excess of prevalence

$$\zeta(t) = \frac{\rho(t) - \rho_\infty}{1 - \rho_\infty} \ . \tag{6}$$

The evolution is markedly different for distinct values of α, but surprisingly not so for the two rules. Overall the dynamics are much quicker and more homogeneous on the scale-free network. Fluctuations in the steady state are of order 10^{-3} and their behavior appears to vary substantially between mechanisms. Their study, however, extends the scope of this essay.

C. Cooperative infectors

Up to this point, active links were unaware of the fact that their targets may have been in contact with other infected nodes, hence all transmissions were considered statistically independent. An alternative view is that healthy agents change their state due to the joint effort of all its infected neighbors. Here we modify the model to allow infectors to cooperate.

Infectious nodes spread the disease homogeneously towards all their healthy neighbors, at constant rate ν. Susceptible nodes gather these toxins and become infected following a distribution $\psi_{\rm inf}(\kappa)$, with κ the amassed viral charge. This means that a healthy agent will become infected after being exposed to κ units of contagion with probability $\psi_{\rm inf}d\kappa$. At a given moment a susceptible node has amassed $K = \nu T$ viral charge, with T the total exposition time, and k^I of its neighbors are infected. If the system remains unaltered during an interval τ, this node will accumulate an additional $\kappa = \nu k^I \tau$. The inter-event time distribution can then be expressed as $\tilde{\psi}_{\rm inf}(\tau)d\tau = \psi_{\rm inf}(\kappa)d\kappa$, with survival probability $\tilde{\Psi}_{\rm inf}(\tau) = \Psi_{\rm inf}(\kappa)$ and instantaneous rate

$$\omega_{\rm inf}(\tau + T) = \frac{\tilde{\psi}_{\rm inf}(\tau + T)}{\tilde{\Psi}_{\rm inf}(\tau + T)} = \nu k^I \frac{\psi_{\rm inf}(\kappa + K)}{\Psi_{\rm inf}(\kappa + K)} \ . \tag{7}$$

Using a Weibull distribution for $\psi_{\rm inf}(\kappa)$ yields an average infection time $\langle\tau\rangle_{\rm inf} = (\mu\nu)^{-1}\Gamma(1+\alpha^{-1})$ and in the limit $N \gg 1$ we may set $\tau = \kappa = 0$; consequently, the instantaneous infection rate for susceptible node j is $\omega_j(T_j) = k_j^I \alpha(\nu\mu)^\alpha T_j^{\alpha-1}$. When the last of its infected neighbors recovers, its exposition time T_j is reset to zero. Recoveries remain exponentially distributed and the spreading ratio is now defined as $\lambda = \langle\tau\rangle_{\rm rec}/\langle\tau\rangle_{\rm inf}$, wherefrom $\mu = \lambda\eta\nu^{-1}\Gamma(1+\alpha^{-1})$. Once again, this definition correctly recovers the Markovian case.

We ran simulations for $\alpha = 0.5$ and $\alpha = 1.5$, following the same procedures and using the same networks as be-

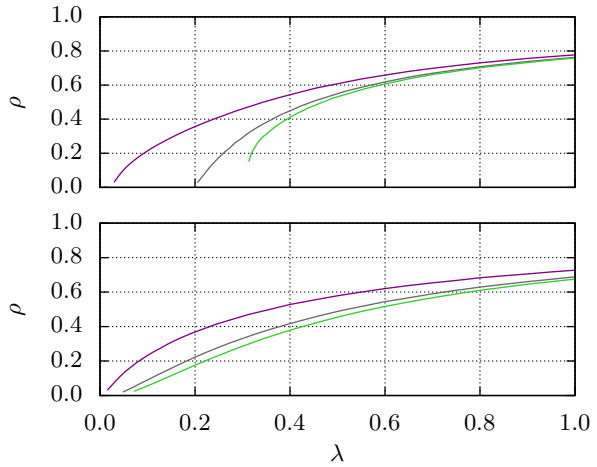

Fig. 5: Cooperative infectors. Phase diagrams for the random (top) and scale-free (bottom) networks described in II C. $\alpha = 0.5$ (purple), $\alpha = 1.5$ (green), Markovian (gray).

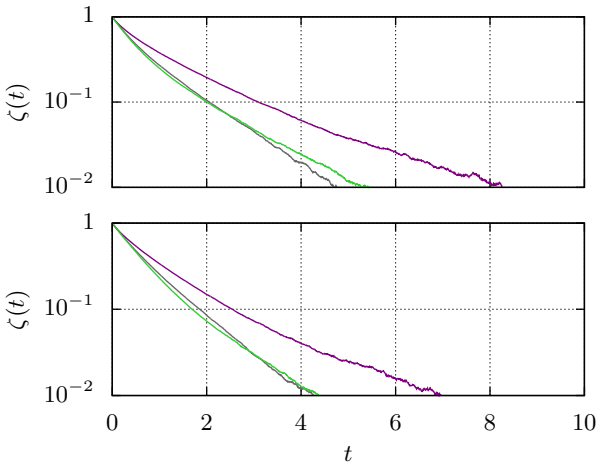

Fig. 6: Cooperative infectors. Dynamics for the random (top) and scale-free (bottom) networks described in II C, for $\lambda = 0.4$. Color code indicated in Fig. 5.

fore. Without loss of generality we used $\eta = \nu = 1$. The late-time prevalence is shown in Figure 5 and Figure 6 shows the reduced excess of prevalence (6) for $\lambda = 0.4$.

Both the steady state and the dynamical evolution differ from the mechanisms presented in Section III B. Noteworthy features include i) the possible existence of a discontinuous phase transition in the random network (for $\alpha = 1.5$) and ii) the similarity in the scale-free network between the Markovian case and that for $\alpha = 1.5$, evident in the phase diagram as in the dynamics. Contrary to the results for heterogeneous transmissions, the evolution is overall alike in both networks but strikingly slower for $\alpha < 1$ than for $\alpha > 1$.

IV. CONCLUSIONS & OUTLOOK

The SIS model for epidemic spreading has been endowed with memory. Modifying the microscopic dynamics has permitted to incorporate non-independent transmissions and cooperative infections. Insofar as analytical treatment of non-Markovian stochastic processes is impossible, computational simulations were used throughout. Preliminary results show that time-sensitivity has a significant effect both on macroscopic properties as on the temporal evolution.

Nonetheless, additional exhaustive analysis is required. For example, the use of bigger systems could clarify the possible existence of an epidemic threshold in scale-free networks and more sophisticated techniques (e.g. susceptibility and coverage) may elucidate whether a discontinuous phase transition appears in random networks. Moreover, exploring the vast landscape of network topologies and metrics might further improve our understanding of epidemic spreading.

ACKNOWLEDGMENTS

Special thanks to my advisor Marián Boguñá, for his continued guidance and counsel. Helena and Pere's useful comments were much appreciated.

[1] Daley, D. J., and Gani, J. *Epidemic Modeling: An Introduction* (Cambridge University Press, New York, 2005)
[2] Hethcote, H. W. *SIAM Rev.* **42** (4): 599 (2000).
[3] Bernoulli, D., and Blower, S. *Rev. Med. Virol.* **14**: 275 (2004).
[4] Brauer, F., and Castillo-Chavez, C. *Mathematical Models in Population Biology and Epidemiology* (Springer, New York, 2001).
[5] Kermack, W. O., and McKendrick, A. G. *Proc. R. Soc. Lond. A* **115**: 700 (1927).
[6] Anderson, R. M., and May, R. M. *Infectious Diseases of Humans* (Oxford University Press, Oxford, 1992).
[7] Keeling, M. J., and Rohani, P. *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, Princeton, 2008).
[8] Newman, M. E. J. *Networks: An Introduction* (Oxford University Press Inc., New York, 2010).
[9] Jackson, M. O. *Social and Economic Networks* (Princeton University Press, Princeton, 2010).
[10] Nishiura, H. *Eur. J. Epidemiol.* **26**: 583 (2011).
[11] Tizzoni, M. et al. *BMC Med.* **10**: 165 (2012).
[12] Vespignani, A. *Nat. Phys.* **8**: 32 (2012).
[13] Albert, R., and Barabási, A.-L. *Rev. Mod. Phys.* **74**: 47 (2002).
[14] Pastor-Satorras, R. et al. arXiv:1408.21701 (2014).
[15] Boguñá, M. et al. *Phys. Rev. Lett.* **111**: 068701 (2013).
[16] Gillespie, D. T. *J. Comput. Phys.* **22**: 403 (1976).
[17] Gillespie, D. T. *J. Phys. Chem.* **81**: 2340 (1977).
[18] Boguñá, M. et al. *Phys. Rev. E* **90**: 042108 (2014).