Power-law decay in epidemics is likely due to interactions with the time-variant contact graph

P. Van Mieghem,[*] M. A. Achterberg and Q. Liu

Delft University of Technology
v1. 7 October 2020
v2. 30 October 2020
v3. 1 December 2020

Abstract

Reported COVID-19 data from official health agencies in several countries suggest that the prevalence, the average fraction of infected people, decays over time with power-law tails. Moreover, the prevalence over the entire outbreak in a country seems reasonably well fitted by a lognormal function. The major consequence of the observation is that Markovian modeling on *fixed graphs*, characterized by exponential times such as initial exponential increase and eventual exponential decrease (in addition to exponential infection and recovery times), fails. As a result, an epidemic will last much longer than predicted by simple but widely used, mean-field Markovian SIR (or variants) equations on a fixed graph. Surprisingly, extensive simulations of non-Markovian SIR epidemics on fixed graphs exhibit exponential decays. Thus, also a realistic non-Markovian description fails. Therefore, we are led to conjecture that power-law decay is caused by interactions between the viral transmission and the time-varying, human contact graph.

The current Corona pandemic places virus spread in networks as part of Network Science in the scientific spotlights. The network science [3] definition of a network rests upon the duality between the network's *graph*, also called the structure or topology, and the network's *process*, also called the function or service that runs over the graph. The graph on N nodes is specified [34] by an $N \times N$ adjacency matrix A, where the element $a_{ij} \in \{0, 1\}$ expresses link existence, reflecting a relation or interaction between node i and j. In general, the graph is not fixed, but changes over time. In epidemics, besides the viral transmission process, a human mobility process (HMP) generates the contact graph in two different ways: (a) in temporal networking, the viral process and HMP are independent and (b) in adaptive networking, a third interaction process couples the viral process and HMP. For example, if people sense an infected, they adapt their behavior by avoiding contacts, otherwise they proceed independently of the viral process, e.g. in case of asymptomatics.

One of the simpler viral processes on networks is SIS epidemic spread [38], which consists of two competing processes [30]: an infection and curing or recovery process, characterized by an infection time distribution $f_T(t)$ and a recovery time distribution $f_R(t)$, respectively. Epidemic spread on

---
[*]Faculty of Electrical Engineering, Mathematics and Computer Science, P.O Box 5031, 2600 GA Delft, The Netherlands; *email*: P.F.A.VanMieghem@tudelft.nl

networks belongs to a class of processes with a simple local rule[1], that results via the interactions with the local rules of its neighbors in a complicated, global emergent behavior.
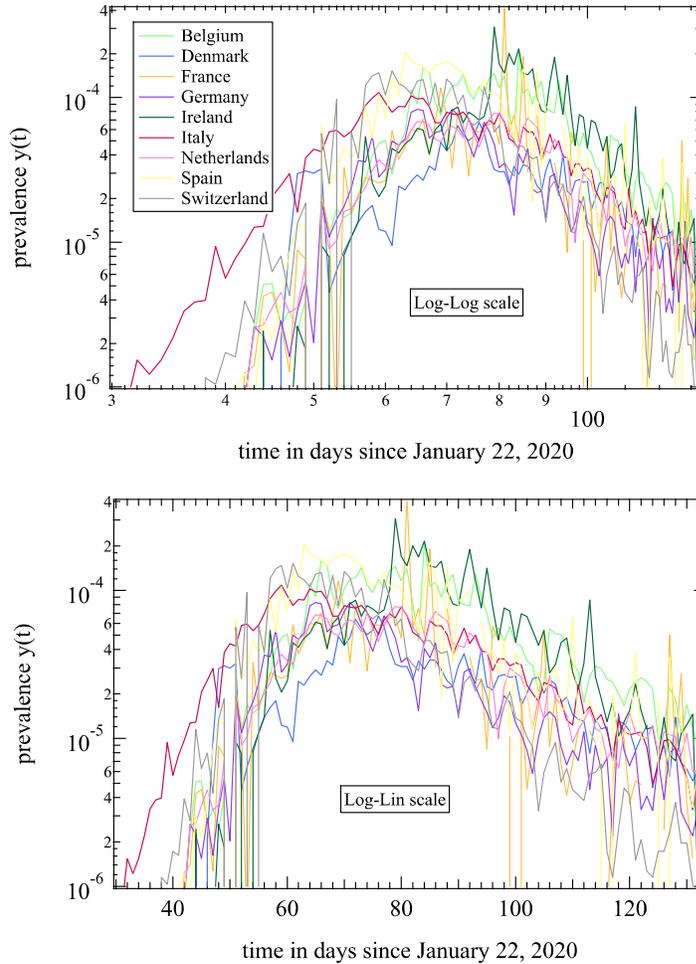
Figure 1: The prevalence (reported from official health authorities) versus time on a log-log (top) and log-lin (bottom) scale in a set of European countries. Only tail data during a "same regime", just after lockdown was released, is considered.

The competition leads to two phases, where either the infection or recovery process is dominant and the separation between the two phases or phase transition is the so-called epidemic threshold. The epidemic threshold τ_c of an epidemic on a network distinguishes, after sufficiently long time in an SIS process, between the overall-healthy network regime and the effective infection regime τ where permanently a non-zero fraction of the nodes is infected. Only in the so-called thermodynamic limit $N \to \infty$, the phase transition occurs in a single point $\tau_{c;\infty}$, whereas for finite graphs, the transition region is smeared out around that point $\tau_{c;\infty}$. The theory of phase transitions [31] is

---

[1] For example, the local rule in SIS epidemics is a simply to program statement: "while infected (until cured), then keep trying to infect your healthy neighbors". The Bernoulli random variables in (2) describe this non-linear rule just as an addition of two multiplications. Other examples of the local-rule-global-emergence-property class (explained in more detail in [41]) are Kuramoto synchronization, swarms of births and fishes, sandpile models, etc.

rich, but complicated. In public reporting, the basic reproduction number R_0 is used in place of the epidemic threshold τ_c. The basic reproduction number R_0, studied in a general setting by [11] and applied to graphs in [32], is defined as the expected number of secondary infected cases produced, in a completely susceptible population, by a typical infective individual. Clearly, R_0 is easier to explain and understand, although R_0 is physically and computationally inferior to the epidemic threshold τ_c (see e.g [22]). Besides the SIS process as basic model for re-infections, the SIR model is the simplest model for a single disease outbreak. A large number of other compartment models [2, 1, 9, 10, 27, 18] exists that detail the viral process, not the graph.

Our conjecture in the abstract arose in an attempt to determine the duration of an epidemic. More precisely, the answer[2] to express "how long an epidemic will last" is given by the tail probability $\Pr\left[Y\left(t\right) \leq y_{\text{acceptable}}\right] < \varepsilon$, where ε is a stringency, the random variable $Y\left(t\right)$ (capital Y) reflects the average number $y\left(t\right)$ (small y) of infected items in a population at time t and $y_{\text{acceptable}}$ is an acceptably low level of the prevalence. Fig. 1 shows several COVID-19 realizations of $Y\left(t\right)$ in several EU countries, while Fig. 2 presents fitted estimates for the probability density function $f_{Y(t)}\left(x\right) = \frac{d}{dx}\Pr\left[Y\left(t\right) \leq x\right]$ for the Netherlands, but other EU countries are remarkably similar. The data is explained in the Appendices. Ideally, if the entire population is measured and tested, then the prevalence $y\left(t\right)$ is a real number, void of uncertainties. From plots like Fig. 2 predictions are made about the initial increase, the highest amount of infections $y(t_p) = 10^{-p}$ (peak at time t_p) and the decay after the peak from which the duration of the epidemic can be obtained from the first point t_a in time where the prevalence $y(t_a) \leq y_{\text{acceptable}}$. Fig. 2 (top) illustrates that most distributions approximate the peak $y(t_p)$ well, while the tail decay of $y(t)$ is puzzling. Indeed, the tails in Fig. 1 can both be fitted by a straight line. Hence, from prevalence data in a limited time interval, both exponential decay $y(t) \approx ae^{-qt}$ and power law decay $y(t) \approx bt^{-u}$ for large t can be concluded. Consequently, the extrapolated duration $t_{a;\exp}$ from exponential and $t_{a;\text{pow}}$ from power law can be hugely different! Accurate prediction is impossible in absense of a theory that specifies the decay law of the prevalence. Fig. 2 (bottom) indicates that a lognormal distribution, possessing power law tails,

$$f_{\text{lognormal}}(t) = \frac{\exp\left[-\frac{(\log t - \mu)^2}{2\sigma^2}\right]}{\sigma t \sqrt{2\pi}} \tag{1}$$

provides the best least-mean square fit. As shown in the Appendix, the observation holds for most other EU countries as well and the fitting parameters $\mu = 3.85 \pm 0.35$ (measured in days) and $0.28 \leq \sigma \leq 0.46$ do not vary much between the countries. A lognormal distribution also occurs in other real spreading processes such as Twitter message [12], social networks [37] and other human behavior [4]. Recent work on COVID-19 has also reported heavy, non-exponential tails. Recently, Maier and Brockmann [23] have reported a *non-exponential increase* of the COVID-19 prevalence in China and argued that quarantine measures can explain non-exponential initial growth. Based on virtually all past epidemics of which data is available and by using rigorous non-parametric statistics, Cirillo and Taleb [7] have demonstrated that the risk of dying in a past epidemic is power-law distributed.

The precise effects that may cause non-exponential asymptotics of the prevalence are unknown and we do not speculate about possible drivers. Instead, we ask "Do we have models that may lead

---

[2]Tail probabilities often occur in telecomunications to assess the quality of service (explained in detail in [33] and [35, p. 2-3]).

to a power-law decay of the prevalence?"
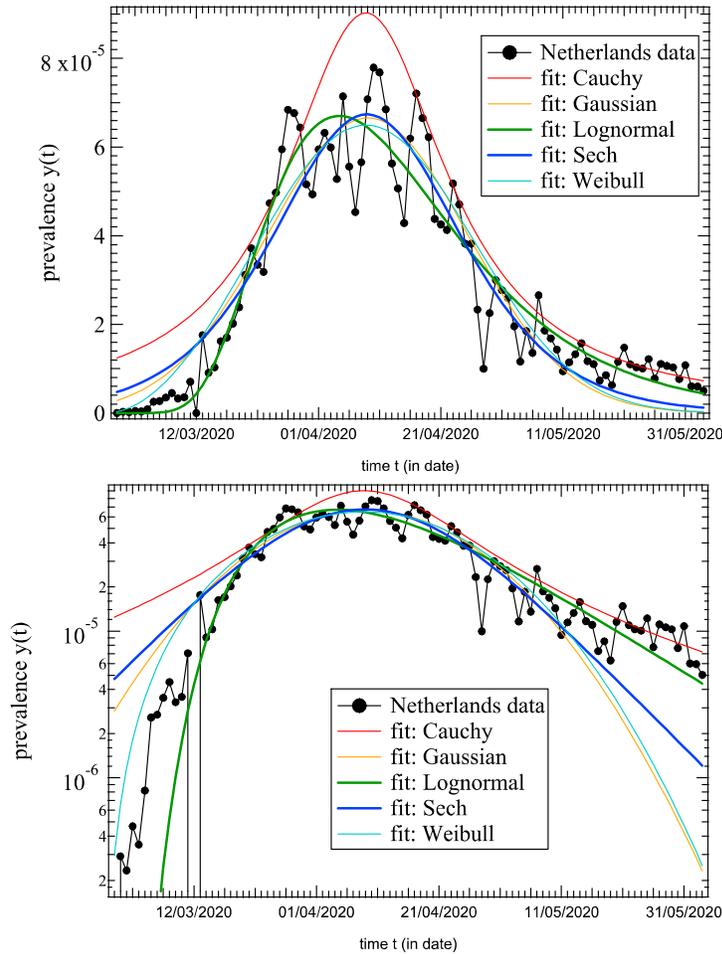
Figure 2: The prevalence $y(t)$ in the Netherlands versus time t (date) for various fitting distributions on both lin-lin scale (top) and log-lin scale (bottom).

Markovian epidemics The majority of studies (see e.g. [27], [18] and [15]) has concentrated on Markovian epidemics, which has allowed us to learn much about the spread of items (biological and digital viruses, emotions, innovations, rumors, etc.) in networks. The Markovian setting, where the present state only depends on the previous state, is analytically tractable. The governing equations of a broad class of epidemics (with any number of compartments) over any fixed graph can be mathematically derived [29]. Confining to SIS, the simplest compartmental model[3], the infectious state $X_i(t)$ of a node i in the network at time t is a Bernoulli random variable, in which $X_i(t) = 1$ if node i is infected, otherwise $X_i(t) = 0$. Both the infection and recovery process are independent Poisson processes and both the infection time T and recovery time R are exponentially distributed with mean $E[T] = \frac{1}{\beta}$ and $E[R] = \frac{1}{\delta}$, where β and δ are the infection and curing rate, respectively. Their ration

---

[3]For simplicity, we limit ourselves to a homogeneous setting where the infection rate $\beta_{ij} = \beta$ of each link (i, j) is the same as well as the recovery rate $\delta_i = \delta$ of each node. The conclusion about the "exponential times" also holds for a full heterogeneous Markovian setting.

$\tau = \frac{\beta}{\delta}$ equals the effective infection rate and measures the infectious strength of the virus. The SIS and SIR epidemic threshold τ_c can be lower bounded [40, 39] by $\tau_c > \frac{1}{\lambda_1} = \tau_c^{(1)}$, the mean-field epidemic threshold, where λ_1 is the largest eigenvalue[4] of the adjacency matrix A. Invoking the Bernoulli property that $E[X_i] = \Pr[X_i = 1]$, the exact SIS governing equation [6] for node i specifies the infection probability of node i as

$$\frac{dE[X_i]}{dt} = E\left[-\delta X_i + \beta(1 - X_i)\sum_{k=1}^{N} a_{ki}X_k\right] \tag{2}$$

where δ, β and a_{ki} can be a function of time. The time-derivative of the infection probability $E[X_i] = \Pr[X_i = 1]$ of a node i consists of the expectation of two competing processes in (2): (a) while node i is infected, i.e. $X_i = 1$, the node i is cured at rate δ and (b) while node i is healthy $X_i = 0$, thus $(1 - X_i) = 1$, all infected neighbors $\sum_{k=1}^{N} a_{ki}X_k$ of node i try to infect the node i with rate β. Rewriting (2) and assuming a fixed graph (where the matrix element a_{ki} is independent of time) as

$$\frac{dE[X_i]}{dt} = -\delta E[X_i(t)] + \beta\sum_{k=1}^{N} a_{ki}E[X_k(t)] - \beta\sum_{k=1}^{N} a_{ki}E[X_i(t)X_k(t)]$$

illustrates that the last non-linear, non-negative term is complicating, but its removal leads to an upper bound in the change of the infection probability $E[X_i]$. When written for all nodes i with $w_i = E[X_i(t)]$ in terms of the $N \times 1$ vector $W = (w_1, w_2, \cdots, w_N)$, we obtain the matrix inequality

$$\frac{dW(t)}{dt} \le (\beta A - \delta I)W(t)$$

whose solution is

$$W(t) \le e^{(\beta A - \delta I)t}W(0)$$

For a characteristic polynomial $\det(A - \lambda I) = \prod_{j=1}^{s}(\lambda - \lambda_j)^{m_j}$, where each eigenvalue λ_j, ordered as $|\lambda_j| < |\lambda_{j-1}|$ for $2 \le j \le s$, is different with multiplicity m_j and $\sum_{j=1}^{s} m_j = N$, the exponential matrix function is [16, p. 116]

$$e^{At} = \sum_{j=1}^{s}\left\{\sum_{k=1}^{m_j} Z_{sk}t^{k-1}\right\}e^{\lambda_s t} \tag{3}$$

where the matrices Z_{sk} are linearly independent constant matrices that are polynomials in A. As reported earlier in [40], [35, p. 457-458], a Markovian epidemic dies out in any graph exponentially fast in time t as $O\left(e^{(\beta\lambda_1 - \delta)t}\right)$ for sufficiently large t when $\tau < \frac{1}{\lambda_1}$. If initially only a few nodes are infected and $\tau > \frac{1}{\lambda_1}$, then the epidemics grows for small times $t > 0$ at most as fast as $O\left(t^{\max_j m_j - 1}e^{(\beta\lambda_1 - \delta)t}\right)$. As reported in [36] even below the epidemic threshold $\tau < \tau_c$, the infection probabilities $E[X_i(t)]$ can initially (for small $t > 0$) increase due to the factors t^{k-1} in (3) when the adjacency matrix A has eigenvalues[5] with multiplicity $m_j > 1$.

In summary, Markovian epidemics on networks are characterized by exponential infection and curing times as well as exponential initial increase above the epidemic threshold τ_c and eventual

---

[4]The largest eigenvalue λ_1 of A is simple and non-negative by the Perron-Frobenius theorem of non-negative matrices [34].

[5]For example, the star graph $K_{1,N-1}$ – whose central node is the extreme super spreader – has the zero eigenvalue of the adjacency matrix A with multiplicity $N - 2$ and the complete graph K_N that has two eigenvalues $\lambda_1 = N - 1$ and $\lambda_j = -1$ for $2 \le j \le N$.

exponential decay in time below τ_c. Both the local rule as the global emergent behavior are exponential in time and "exponential times" are fingerprints of underlying Markovian processes.

Power-law decay of the SIS prevalence due to Griffith's phase theory is reported in [25, 8]. However, Griffith's phase theory of networks requires the existence of very-large-degree hubs or cliques in the network and an effective infection rate τ located at the epidemic threshold τ_c. Since the degree of human contacts are limited and τ is unlikely situated exactly at the epidemic threshold τ_c, the Griffith's phase effect does not explain power-law decay. Moreover, the existence of Griffith's phase may rely on a recurrent infection state as in the SIS model, while a single outbreak is closer modeled by an SIR epidemic (or SIR-like variants).

non-Markovian epidemics Whereas Markovian epidemics are both locally (infection and recovery times) and globally (the prevalence) characterized by an exponential function in time, real epidemics do not possess exponential infection or recovery times (see SI). Consequently, we *hypothesize* that the global dynamics may decay non-exponentially in time. A fundamental question is whether the distribution of "local" infection time T and recovery time R in each node gives rise, on any graph, to the same time-distributions of globally observed quantities as the prevalence, but likely with different parameters (e.g. as μ and σ in the lognormal (1)) due to regional differences. The entire COVID-19 prevalence, well fitted by a lognormal in Fig. 2, seems to suggest that either the infection time T or recovery time R or both may possess a heavy-tailed distribution.

It would be desirable to possess a non-Markovian theory of epidemics on networks that could verify these "from measurements" deduced findings and could offer some order estimate for the prevalence's $y(t)$ decay with time t. When a stochastic process is not Markovian, its mathematical description and analysis is considerably more complex. At the best of our knowledge, the precise SIR or SIS governing equations for any fixed graph are not available, not even in a mean-field approximation[6], although the steady-state of non-Markovian SIS epidemics on any fixed graph can be approximated well by a mean-field approach [5, 38]. In many non-Markovian processes on networks and also in this paper, computer simulations and measurements are often the only resort to investigate its behavior and properties.

Extensive simulations on fixed graphs with different parameters and distributions for infection time T and recovery time R contradict our hypothesis. Fig. 3 shows the prevalence as a function of time in a Barabási-Albert power-law degree graph on $N = 1000$ nodes. Similar plots of other graphs (a lattice graph and an Erdős-Rényi graph) are shown in the Appendix. On each graph type, three combinations are simulated: (a) an exponential infection and curing time (i.e. Markovian epidemics), (b) exponential infection and lognormal curing time and (c) lognormal infection and curing time. Apart from the rather complicated time-dependence of the average fractions of susceptible, infected and removed, the log-lin scale demonstates for sufficiently large time t that *the prevalence decay is exponential*! Other simulations (not shown) with different non-exponential distribution for infection time T and recovery time R confirm that the non-Markovian SIR on a fixed graph decays at least as fast as an exponential function in time t. Of course, simulations – no matter how many – never prove a statement, they can only disprove a theory by showing a counterexample.

---

[6]The general theory in Kermack and McKendrick [17] for a homogeneous population (thus a complete graph) is described by a Volterra-type of integral differential equation and is still a mean-field approximation. However, the extension to arbitrary graphs is lacking, because it is challengingly difficult.
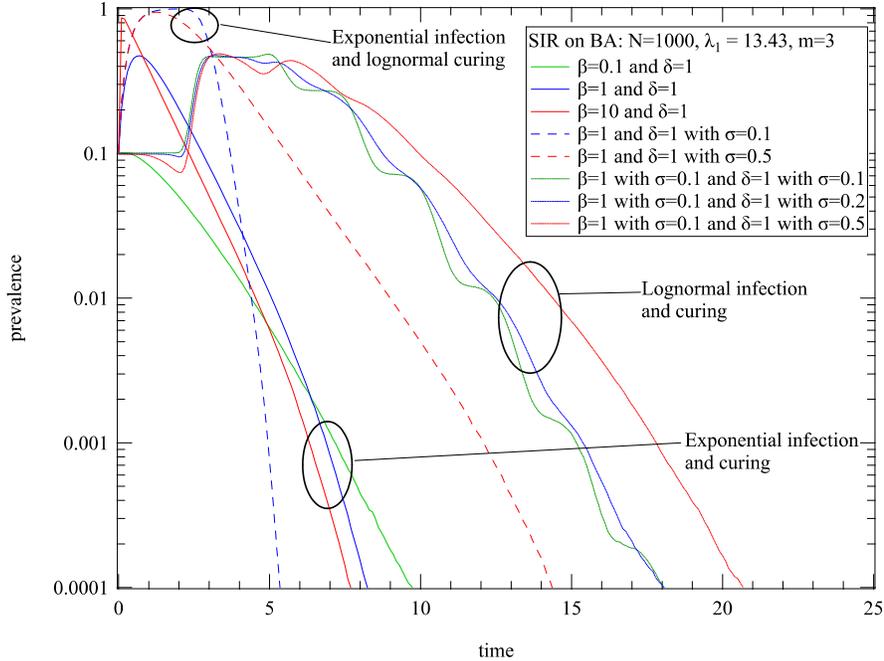
Figure 3: The SIR prevalence versus time on a Barabasi-Albert graph. The dotted curve is the prevalence of the SIR process with lognormal distributed infection and curing time. The dashed curve refers to an exponential infection time and a lognormal curing time. The smooth curve is the prevalence of the Markovian SIR process. All prevalence curves decay exponentially or slightly faster (because the prevalence is almost straight in a log-lin scaled plot). For the non-Markovian SIR process with a lognormal infection time, the prevalence oscillates with time. The reason for the oscillation is that a lognormal distribution infection time introduces a synchronized infection: All initially infected nodes infect susceptible neighbors at a time, which approximately equal to the mean of the lognormal distribution of the infection time. The oscillation of the prevalence will be smoothed out with time.

If a general, non-Markovian epidemic process on a fixed graph cannot explain power-law decay in the prevalence in real epidemics – a conjecture that needs a rigorous proof –, then, as a plausible explanation, the interaction of the epidemic process with the dynamics of the human contact graph may cause a heavy-tailed decay. In classical epidemics, the main attention focuses on the viral process, while the HMP is largely ignored or mainly abstracted by a fixed graph or an averaged topology. Our study hints to the considerable influence of the time-varying contact graph, its governing HMP and possible interaction with the viral process. At present, neither accurate models for HMP, nor the coupling process between HMP and viral process can be written down in governing equations. We conclude that epidemic theory on time-varying (and possibly adaptive) networks is needed as well as models for human mobility processes that produce time-varying contact graphs. We expect a new research avenue in the physics of real epidemics that will combine the three essential processes (viral, HMP and coupling). At last, digital technology to measure the contact network (such as mobile app algorithms [28]) will help in the quest towards understanding and predicting.

References

[1] R. M. Anderson and R. M. May. *Infectious Diseases of Humans: Dynamics and Control.* Oxford University Press, Oxford, U.K., 1991.

[2] N. T. J. Bailey. *The Mathematical Theory of Infectious Diseases and its Applications.* Charlin Griffin & Company, London, 2nd edition, 1975.

[3] A. L. Barabási. *Network Science.* Cambridge University Press, Cambridge, U.K., 2016.

[4] N. Blenn and P. Van Mieghem. Are human interactivity times lognormal? Delft University of Technology, Report20160711 (www.nas.ewi.tudelft.nl/people/Piet/TUDelftReports)*; arXiv:1607.02952*, 2016.

[5] E. Cator, R. van de Bovenkamp, and P. Van Mieghem. Susceptible-Infected-Susceptible epidemics on networks with general infection and curing times. *Physical Review E*, 87(6):062816, June 2013.

[6] E. Cator and P. Van Mieghem. Second order mean-field SIS epidemic threshold. *Physical Review E*, 85(5):056111, May 2012.

[7] P. Cirillo and N. N. Taleb. Tail risk of contagious diseases. *Nature Physics*, 16:606–613, June 2020.

[8] Wesley Cota, Silvio C. Ferreira, and Géza Ódor. Griffiths effects of the susceptible-infected-susceptible epidemic model on random power-law networks. *Physical Review E*, 93:032322, Mar 2016.

[9] D. J. Daley and J. Gani. *Epidemic modelling: An Introduction.* Cambridge University Press, Cambridge, U.K., 1999.

[10] O. Diekmann, H. Heesterbeek, and T. Britton. *Mathematical Tools for Understanding Infectious Disease Dynamics.* Princeton University Press, Princeton, USA, 2012.

[11] O. Diekmann, J. A. P. Heesterbeek, and J. A. J. Metz. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology*, 28:365–382, 1990.

[12] C. Doerr, N. Blenn, and P. Van Mieghem. Lognormal infection times of online information spread. *PLoS ONE*, 8(5):e64349, May 2013.

[13] E. Dong, H. Du, and L. Gardner. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*, 20:533–534, May 2020.

[14] Z. Du, X. Xu, Y. Wu, L. Wang, B. J. Cowling, and L. Meyers. Serial Interval of COVID-19 among Publicly Reported Confirmed Cases. *Emerging Infectious Diseases*, 26:1341–1343, 2020.

[15] G. Ferraz de Arruda, F. A. Rodrigues, and Y. Moreno. Fundamentals of spreading processes in single and multilayer complex networks. *Physics Reports*, 756:1–59, 2018.

[16] F. R. Gantmacher. *The Theory of Matrices*, volume II. Chelsea Publishing Company, New York, 1959.

[17] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society London, A*, 115:700–721, August 1927.

[18] I. Z Kiss, J. C. Miller, and P. L Simon. *Mathematics of network epidemics: from exact to approximate models.* Springer, 2016.

[19] S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine*, 172(9):577–582, 2020. PMID: 32150748.

[20] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S.M. Leung, E. H.Y. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T.Y. Lam, J. T. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung, and Z. Feng. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus - Infected Pneumonia. *New England Journal of Medicine*, 382(13):1199–1207, 2020.

[21] N. M. Linton, T. Kobayashi, Y. Yang, K. Hayashi, A. R. Akhmetzhanov, S. Jung, B. Yuan, R. Kinoshita, and H. Nishiura. Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data. *Journal of Clinical Medicine*, 9(2), 2020.

[22] Q.-H. Liu, M. Ajelli, A. Aleta, S. Merler, Y. Moreno, and A. Vespignani. Measurability of the epidemic reproduction number in data-driven contact networks. *Proceedings of National Academy of Science of the United States of America (PNAS)*, 115(50):12680–12685, November 2018.

[23] B. F. Maier and D. Brockmann. Effective containment explains subexponential growth in recent confirmed Covid-19 cases in China. *Science*, 368:742–746, 15 May 2020.

[24] J. C. Miller and T. Ting. Eon (epidemics on networks): a fast, flexible python package for simulation, analytic approximation, and analysis of epidemics on networks. *Journal of Open Source Software*, 4(44):1731, 2019.

[25] Paolo Moretti and Miguel A Muñoz. Griffiths phases and the stretching of criticality in brain networks. *Nature communications*, 4(1):1–10, 2013.

[26] H. Nishiura, N. M. Linton, and A. R. Akhmetzhanov. Serial interval of novel coronavirus (COVID-19) infections. *International Journal of Infectious Diseases*, 93:284 – 286, 2020.

[27] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Review of Modern Physics*, 87(3):925–979, September 2015.

[28] B Prasse and P. Van Mieghem. Mobile smartphone tracing can detect almost all SARS-CoV-2 infections. *arXiv:2006.14285*, 2020.

[29] F. D. Sahneh, C. Scoglio, and P. Van Mieghem. Generalized epidemic mean-field model for spreading processes over multi-layer complex networks. *IEEE/ACM Transaction on Networking*, 21(5):1609–1620, October 2013.

[30] J. Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Review of Modern Physics*, 48(4):571–585, October 1976.

[31] H. E. Stanley. *Introduction to Phase Transitions and Critical Phenomena*. Oxford University Press, July 1987.

[32] P. van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180:29–48, 2002.

[33] P. Van Mieghem. *Data Communications Networking*. Piet Van Mieghem, ISBN 978-94-91075-01-8, Delft, 2nd edition, 2010.

[34] P. Van Mieghem. *Graph Spectra for Complex Networks*. Cambridge University Press, Cambridge, U.K., 2011.

[35] P. Van Mieghem. *Performance Analysis of Complex Networks and Systems*. Cambridge University Press, Cambridge, U.K., 2014.

[36] P. Van Mieghem. Approximate formula and bounds for the time-varying SIS prevalence in networks. *Physical Review E*, 93(5):052312, 2016.

[37] P. Van Mieghem, N. Blenn, and C. Doerr. Lognormal distribution in the Digg online social network. *European Physical Journal B*, 83(2):251–261, 2011.

[38] P. Van Mieghem and Q. Liu. Explicit non-Markovian SIS mean-field epidemic threshold for Weibull and Gamma infections but Poisson curings. *Physical Review E*, 100(2):022317, August 2019.

[39] P. Van Mieghem, F. D. Sahneh, and C. Scoglio. Exact Markovian SIR and SIS epidemics on networks and an upper bound for the epidemic threshold. *Proceedings of the 53rd IEEE Conference on Decision and Control (CDC2014), December 15-17, Los Angeles, CA, USA*, 2014.

[40] P. Van Mieghem and R. van de Bovenkamp. Non-Markovian infection spread dramatically alters the SIS epidemic threshold in networks. *Physical Review Letters*, 110(10):108701, March 2013.

[41] P. Van Mieghem and R. van de Bovenkamp. Accuracy criterion for the mean-field approximation in SIS epidemics on networks. *Physical Review E*, 91(3):032812, March 2015.

A   Data

We select a set of 13 European countries, where the outbreak of COVID-19 has been reduced sufficiently to distinguish a clear epidemic peak. The time series for the number of cases is obtained from the Dashboard of the Johns Hopkins University [13] that started collecting data from January 22, 2020. We consider the period from January 22 until June 3. In most European countries the pandemic started around end February 2020, followed a couple of weeks later by a lockdown, which was released around the beginning of June. The data thus reflects one "regime" of rise and decay of the prevalence $y(t)$ per country under lockdown.

The power-law exponent of the 13 European countries has been fitted by a linear fit $\log y(t) = a \log t + b$ in the tail region and the slope a, shown in Fig. 4, lies around $a = -5 \pm 2$.
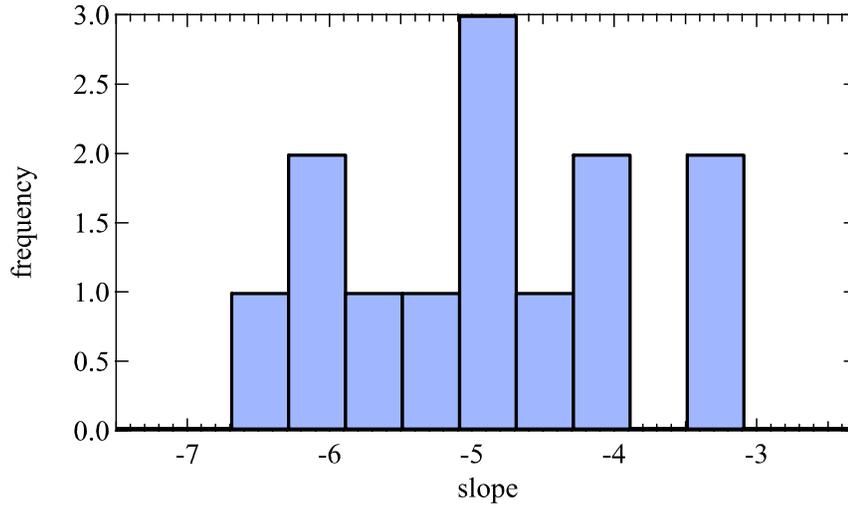
Figure 4: The histogram of slope of $\log y(t)$ versus $\log t$ for 13 countries.

We fit the daily number of reported cases of the prevalence $y(t)$ to the following functions/distributions:

$$\text{Weibull:} \quad \frac{k}{l} \left(\frac{t}{l}\right)^{k-1} e^{-(t/l)^k}$$

$$\text{Gaussian:} \quad \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right]$$

$$\text{Lognormal:} \quad \frac{1}{\sigma t\sqrt{2\pi}} \exp\left[-\frac{(\log(t)-\mu)^2}{2\sigma^2}\right]$$

$$\text{Cauchy:} \quad \frac{g}{\pi[g^2 + (t-\mu)^2]}$$

$$\text{Sech:} \quad \frac{K}{2} \frac{1}{1 + \cosh(-K(t-\mu))}$$

The Sech function follows as the approximate solution of the SIR model [17]. Each distribution function is multiplied by a constant $c \in [0, 1]$ to account for the fact that the area under the curve does not necessarily equal one.

The parameters in each of the above functions are estimated using the nonlinear curve fitting procedure `GlobalSearch` in Matlab. The Lognormal distribution shows the best performance for 10

countries and the Cauchy distribution for 3 countries, as indicated in Table 1. The fitting performance is commonly measured in terms of the R-squared R^2, which is also reported [19, 26, 14, 21, 20] for the Lognormal distribution in Table 1.

Table 1: The countries considered in this work.

country	Log-Log slope	best fit	Lognormal μ	Lognormal σ	Lognormal R^2
Austria	-6.21	Cauchy	3.51	0.28	0.8156
Belgium	-5.11	Lognormal	3.78	0.40	0.8162
Denmark	-3.49	Lognormal	3.89	0.45	0.6779
France	-5.93	Cauchy	3.82	0.30	0.3126
Germany	-4.64	Lognormal	3.69	0.36	0.8808
Ireland	-5.64	Cauchy	3.90	0.28	0.7386
Italy	-4.08	Lognormal	3.78	0.46	0.9402
Netherlands	-4.92	Lognormal	3.81	0.40	0.9035
Serbia	-5.01	Lognormal	3.80	0.33	0.7914
Spain	-4.87	Lognormal	3.68	0.34	0.6940
Switzerland	-6.35	Lognormal	3.55	0.34	0.8768
Turkey	-4.14	Lognormal	3.71	0.44	0.9252
United Kingdom	-3.34	Lognormal	4.17	0.40	0.8677

Many studies investigate the infection time of COVID-19, which is also called the generation time or serial interval. The infection time is defined as "the time from illness onset in a primary case (infector) to illness onset in a secondary case (infectee)". Nishiura et al. (2020) suggest that the infection time is lognormally distributed with median 4.0 days [26] whereas Du et al. (2020) find that the infection time is normally distributed with mean 3.96 days [14]. The incubation time is "the time delay from infection to illness onset". Lauer et al. (2020) observe a lognormally distributed incubation time with mean 5.2 days [19] which is also found by Li et al. (2020) [20], where Linton et al. (2020) find a lognormal distribution with median 5.0 days [21]. [The values for σ in the lognormal are generally not reported.]

Fig. 5 shows the example for the Netherlands and Italy, but more countries are fitted reasonably well by a lognormal function.

B   Simulation

We use the *Epidemic on Networks* (EoN) module [24], a Python package for simulating exact stochastic processes and solving ordinary differential equations of epidemics, to simulate the SIR processes. Both Markovian and non-Markovian SIR processes are simulated on three different networks: a 32 by 31 lattice ($N = 992$ nodes), a Barabási-Albert (BA) network with $N = 1000$ and an Erdős-Rényi (ER) network with $N = 1000$. For Markovian SIR, we set both the infection and curing time exponentially distributed. For the non-Markovian process, two cases are simulated: SIR with exponential infection time and lognormal curing time; SIR with both lognormal infection and curing time. Each prevalence

curve in the figures is obtained by averaging over 100 realizations and in each realization, 10 percent nodes are initially infected. We vary the parameters in each case of simulation and find only exponential decay of the prevalence.

Fig. 6 and Fig. 7 indicates that the prevalence in non-Markovian SIR epidemics has an exponential tail for sufficiently large time.
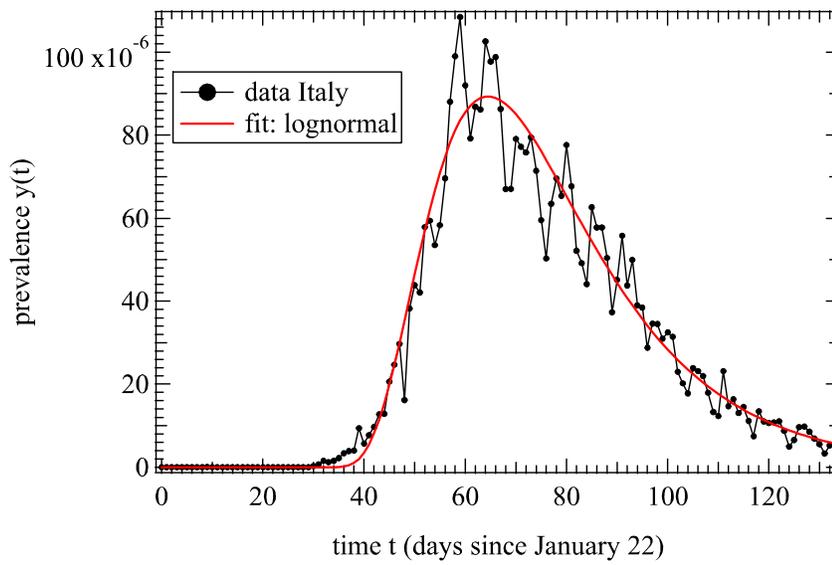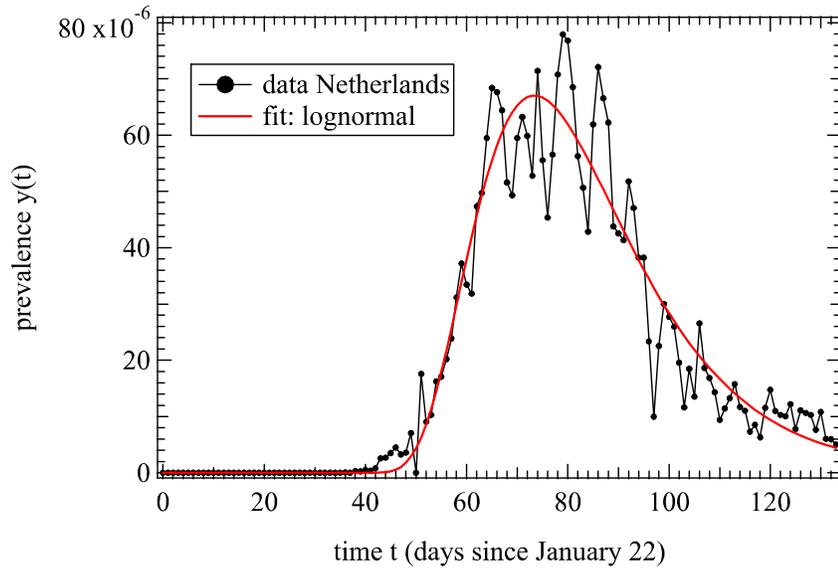
Figure 5: Two examples of the prevalence in the Netherlands and Italy versus time, fitted with a lognormal function.
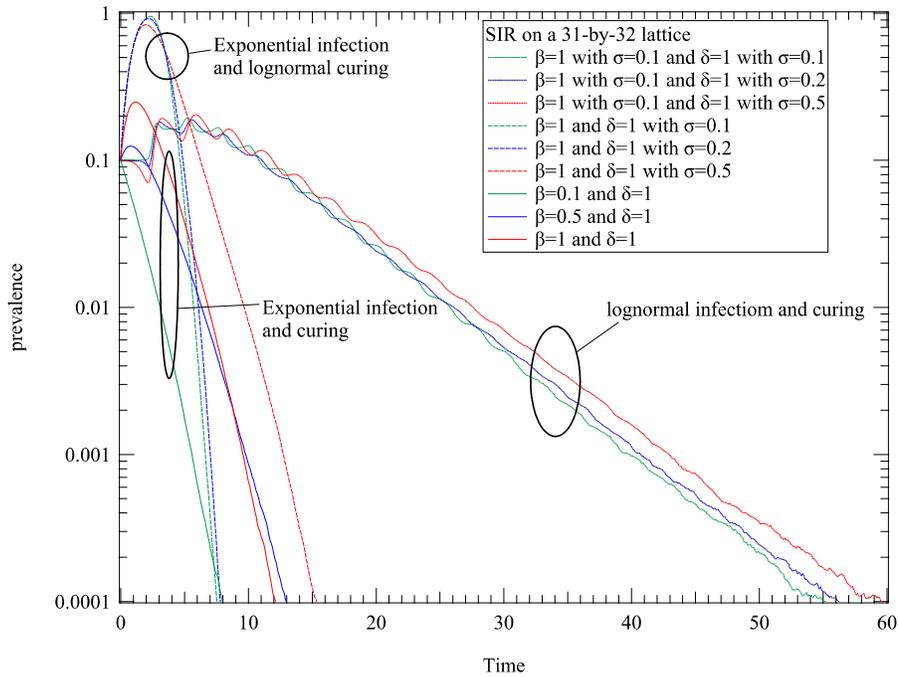
Figure 6: The prevalence of SIR on a 31 by 32 lattice. The dotted curve is the prevalence of the SIR process with lognormal distributed infection and curing time. The dashed curve refers to an exponential infection time and a lognormal curing time. The smooth curve is the prevalence of the Markovian SIR process. All prevalence curves decay exponentially or slightly faster (because the prevalence is almost straight in a log-lin scaled plot). For the non-Markovian SIR process with a lognormal infection time, the prevalence oscillates with time. The reason for the oscillation is that a lognormal distribution infection time introduces a synchronized infection: All initially infected nodes infect susceptible neighbors at a time, which approximately equal to the mean of the lognormal distribution of the infection time. The oscillation of the prevalence will be smoothed out with time. The other plot in Fig. 7 is similar to this figure and differs only in the underlying graph.

14
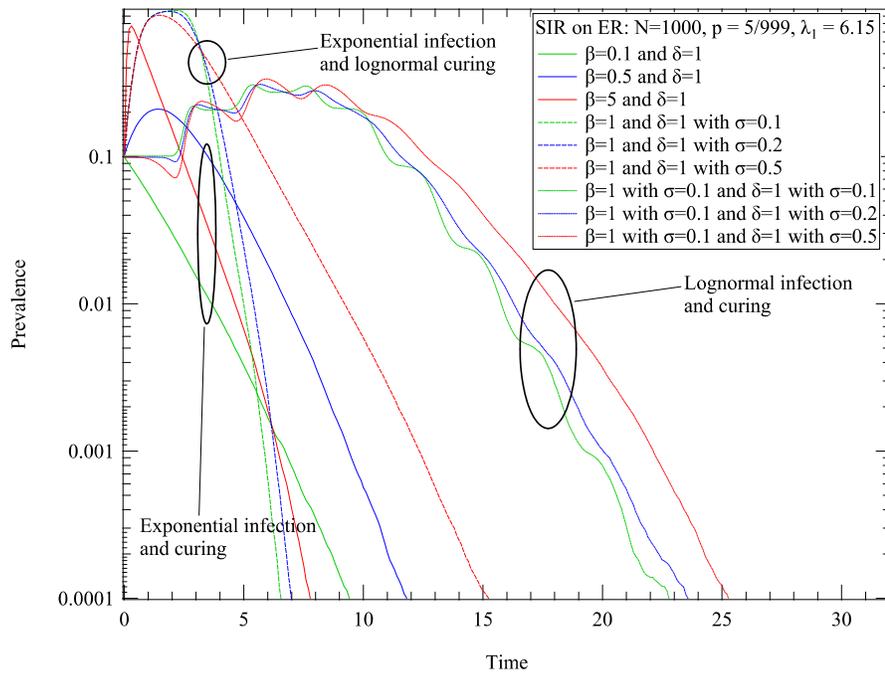
Figure 7: The SIR prevalence versus time on a Erdos-Renyi graph (similar to Fig. 6)