The spreading time in SIS epidemics on networks

Zhidong He* P. Van Mieghem*

Delft University of Technology

Abstract

In a Susceptible-Infected-Susceptible (SIS) process, we investigate the spreading time T_m, which is the time when the number of infected nodes in the metastable state is first reached, starting from the outbreak of the epidemics. We observe that the spreading time T_m resembles a lognormal-like heavy-tailed distribution, both for the Markovian and the non-Markovian infection process, which implies that the spreading time can be very long with a relatively high probability. In addition, we show that a stronger virus, with a higher effective infection rate τ or an earlier timing of the infection attempts, does not always lead to a shorter average spreading time $E[T_m]$. We numerically demonstrate that the average spreading time $E[T_m]$ in the complete graph and the star graph scales logarithmically as a function of the network size N for a fixed fraction of infected nodes in the metastable state.

1 Introduction

Epidemic spreading on networks is a ubiquitous process, which can describe the information spreading on social networks [1], emotions [2], biological diseases [3] and failures in networked systems [4]. The Susceptible-Infected-Susceptible (SIS) model is a simple epidemic model where each infected item can be cured, and becomes susceptible again after recovering from the disease. Since the epidemic is a time-dependent spreading process, we are naturally concerned with characteristic times that can be applied to predict or control the spreading process. In spite of the simplicity of the SIS process, unfortunately, only a few results for exact SIS times on a generic graph have been presented [5, p. 460].

We concentrate on the Markovian Susceptible-Infected-Susceptible (SIS) epidemics on a graph, where both the curing and infection processes are Poisson processes [5]. The ratio between the infection rate β and the curing rate δ is called the effective infection rate $\tau = \beta/\delta$. The SIS model features a phase transition [6] around the epidemic threshold τ_c. Viruses with an effective infection rate τ above the epidemic threshold τ_c can infect a sizeable portion of the population on average and stay for a long time in the network. This long period is called the metastable state. A first-order mean-field approximation of the epidemic threshold $\tau_c^{(1)} = 1/\lambda_1(A)$, where $\lambda_1(A)$ is the spectral radius of the adjacency matrix A, was shown [3] [7] to be a lower bound for the epidemic threshold, $\tau_c^{(1)} < \tau_c$.

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

Due to the existence of an absorbing state, which is the overall healthy or disease-free state in the SIS process, any initial infection will ultimately extinguish in any finite graph. The time until the network reaches the all-healthy state is called the extinction time, or alternatively, the time to absorption or the survival time [8]. When the effective infection rate τ is below the epidemic threshold τ_c, the infectious process dies out exponentially fast for sufficiently large time [9], which is called quick die out or *early* extinction. The mean early extinction time is of the order of $\log N$, where N is the number of nodes in the network. A sufficient condition for *slow* extinction is that the effective infection rate τ is above the epidemic threshold τ_c. If the effective infection rate $\tau > \tau_c$, the infection stays very long in any sufficiently large network [10]. The average survival time is dominated by the second largest eigenvalue of the infinitesimal generator of the Markov chain [8] [11].

The extinction time in real-world large graphs is too long, compared to the actually observed time that an epidemic lasts. Therefore, besides the extinction time, we are interested in characteristic times before the absorbing state is reached. Van de Bovenkamp and Van Mieghem [12] showed that the average hitting time to the metastable state can be computed by using a uniformed embedded Markov chain for the complete graph and the star graph. However, the analytic expression of the average hitting time to the metastable state in a generic graph is hard to derive in closed form.

In this paper, we define the spreading time T_m as the time when the number I_m of infected nodes in the metastable state is first reached, starting from one initially infected node. The spreading time indicates the spreading velocity of the SIS process in the early stage and unveils the time-dependent properties of epidemic activity before the metastable state. Based on the simulations, we study the distribution of the spreading time T_m both for the Markovian and non-Markovian infection process, and further investigate the effect of the effective infection rate τ, the network size N and the non-Markovian process on the average spreading time $E[T_m]$.

This paper is organized as follows. Section 2 introduces the definition and determination of the spreading time. We investigate the distribution of the spreading time T_m in Section 3. In Section 4, we further present the effect of the effective infection rate τ, the non-Markovian infection times and the network size N on the average spreading time. Finally, the conclusion are made in Section 4. We define the metastable state and the stability t_s in a SIS process in Appendix A. Appendix B presents the procedure of the simulator for SIS epidemics (SSIS). Appendix C offers an approximate analytic computation of the spreading time T_m in a complete graph K_N with N nodes.

2 Definition and determination of the spreading time

We first propose a preferred definition of the metastable state and the stability time t_s as follow:
Definition 1: In an epidemic process, the metastable state is reached at the stability time t_s, which is the smallest time obeying $\frac{d\overline{y}(t)}{dt}\big|_{t>t_s} < \epsilon$, where the average fraction of infected nodes is $\overline{y}(t) = \frac{1}{N}E[I(t)]$, with $I(t) \geq 1$ is the number of infected nodes at time t, and ϵ is a small positive real number that needs to be agreed upon.

A more detailed discussion on the determination of the stability time is presented in Appendix A.
Definition 2: The spreading time T_m is defined as the first time when the number $I_m = I(t_s)$ of the infected nodes in the metastable state is reached, starting from one initially infected node.

Specifically, the probability distribution of the spreading time T_m in the graph G with N nodes

follows

$$\Pr[T_m \le t] = \sum_{n=1}^{N} \Pr[T_m \le t | I(t_s) = n] \Pr[I(t_s) = n] \tag{1}$$

Thus, the average spreading time $E[T_m]$ follows from (1) as

$$E[T_m] = \sum_{n=1}^{N} E[T_{H_n}] \Pr[I(t_s) = n] \tag{2}$$

where the hitting time $T_{H_n} = T_m\big|_{I(t_s)=n}$ is the first time when the Markov chain reaches the state with n infected nodes. After differentiating both sides of (1) with respect to t, we obtain the probability density function (pdf) of the spreading time $f_{T_m}(t)$:

$$f_{T_m}(t) = \sum_{n=1}^{N} f_{T_m}(t | I(t_s) = n) \Pr[I(t_s) = n] \tag{3}$$

Physically, the spreading time T_m describes the spreading velocity in the early stage of the spreading process, which depends on the local topology around the initial spreaders. The spreading time is of practical significance to control the epidemic because it gives an upper-bound to organize the immunization actions to eradicate the epidemic. After T_m time units, the epidemic approximates the metastable state and already infected a substantial part of the population. Thus, the action of control is preferred to be taken earlier than the average spreading time $E[T_m]$.

Due to the limitation of the analytical methods, an event-driven simulator SSIS (see Appendix B) for the SIS spreading process on a network is implemented based on the Gillespie algorithm [13] to estimate the spreading time. For an unaltered graph and a fixed effective infection rate τ, the epidemic begins with one initially infected node and lasts for the period of t_{limit} time units which is ensured to be long enough to make the spreading process reach the metastable state but not the absorbing state. We record every time point t_k when the kth event happens, as well as the corresponding number of the infected nodes $i(t_k)$ after the kth event. Assume that $0 < t_1 \le t_2 \le \cdots \le t_m < t_{limit}$, then m events have occurred on the timeline before the time limit t_{limit}. After identifying the metastable state and the stability time t_s (see Appendix B), we then determine the spreading time t_m in each realization. The spreading time can be determined from the time t_m when the number of infected nodes $i(t_m)$ first equals to the number $i(t_s)$ of infected nodes at the stability t_s of the metastable state. The random variable T_m corresponds to the spreading time t_m in all realizations *without* extinction probability. Figure 1 illustrates the estimation scheme of the spreading time t_s in a complete graph K_{50}, which also shows the Gaussian-like distribution of the number of infected nodes in the metastable state.

3 Distribution of the spreading time T_m

The hitting time T_{H_i} is the first time when the Markov process reaches the state with i infected nodes, starting from one initial spreader. The epidemic process in the complete graph K_N is a birth and death process, and the average hitting time $E[T_{H_i}]$ from one initial spreader can be analytically derived [12] as

$$E[T_{H_i}] = \sum_{j=1}^{i-1} \sum_{k=0}^{i-j-1} \frac{(N-i+k)!\tau^{j+k-i}}{j(N-j)!}. \tag{4}$$

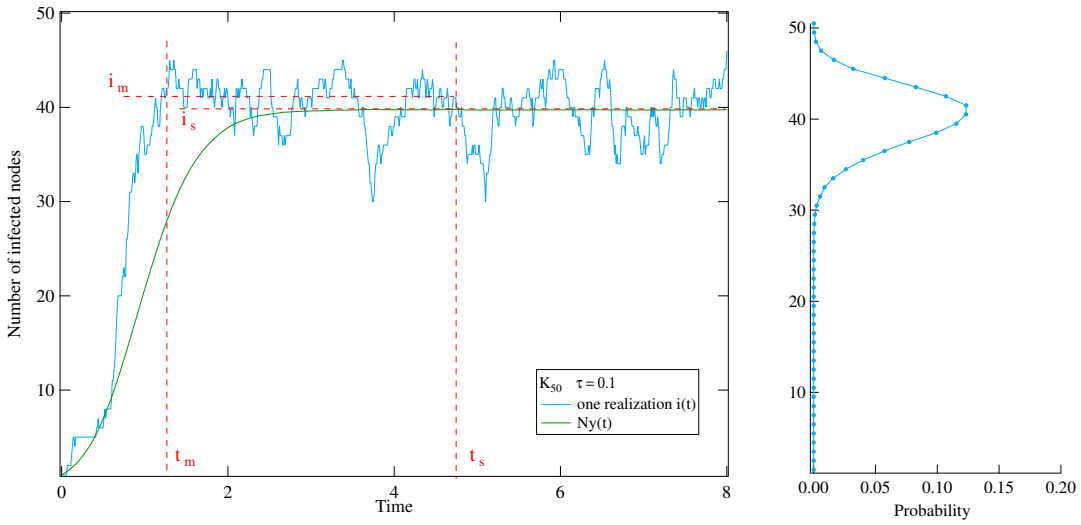

Figure 1: Illustration of the estimation scheme of the stability time t_s in the prevalence via SSIS and the spreading time t_m for one realization $i(t)$. The distribution of the number of infected nodes in the metastable state is shown in the right subgraph. The green line represents the average number of infected nodes with time based on 10^6 realizations.

in the modified SIS (MSIS) model [14], where the absorbing state is removed in MSIS Markovian chain. However, a hitting time analysis is only attractive when the spreading process can be described as a simple, analytically tractable Markov chain [12].

Figure 2 exemplifies the average hitting time $E[T_{H_i}]$, from one initial spreader, as a function of the fraction $y = \frac{i}{N}$ of the infected nodes in the complete graph K_{50} with different effective infection rate τ. NIMFA approximates the average number of infected nodes in the metastable state for a complete graph K_N with N nodes as $i_s = \lfloor N\left(1 - \frac{1}{\tau(N-1)}\right)\rfloor$. When the effective infection rate τ is above the epidemic threshold τ_c, the average hitting time $E[T_{H_i}]$ exhibits two different regimes in the average fraction y of infected nodes as shown in Figure 2. In Regime 1, where $y < \frac{i_s}{N}$, the average hitting time $E[T_{H_i}]$ increases exponentially-like as $ae^{\kappa y}$, where the rate κ decreases with the effective infection rate τ. In Regime 2, where $y > \frac{i_s}{N}$, the average hitting time $E[T_{H_i}]$ increases faster than an exponential function.

Figure 2 suggests that the average hitting time $E[T_{H_n}]$ scales approximately exponentially with the number n of infected nodes around the average number $E[I(t_s)]$ of infected nodes in the metastable state. Assuming that the hitting time T_{H_n} with small variance is correlated to the number n of infected nodes $T_{H_n} \propto e^{\kappa n}$, the spreading time can be regarded as the random variable $T_m(I(t_s)) \approx e^{\kappa I(t_s)}$, where the number of the infected nodes $I(t_s)$ is approximately a Gaussian-like random variable [14] with probability density function $\Pr\left[I\left(t_s\right) = n\right] \approx \frac{1}{\tilde{\sigma}\sqrt{2\pi}} \exp\left[-\frac{(n-\tilde{\mu})^2}{2\tilde{\sigma}^2}\right]$. Therefore, we may infer that the pdf of the spreading time by

$$f_{T_m}\left(t\right) \approx \frac{1}{\kappa t\tilde{\sigma}\sqrt{2\pi}} \exp\left[-\frac{(\frac{1}{\kappa}\log t - \tilde{\mu})^2}{2\tilde{\sigma}^2}\right] = \frac{1}{\sigma t\sqrt{2\pi}} e^{-\frac{(\log t - \mu)^2}{2\sigma^2}}, \tag{5}$$

which is a lognormal distribution by replacing $\mu = \kappa\tilde{\mu}$ and $\sigma = \kappa\tilde{\sigma}$.

The conclusion agrees with the analysis in Appendix C. The hitting time to a state with d infected nodes starting from state 1 is distributed as the sum of d independent exponential random variables

each with a rate η_k. We approximate the distribution of the spreading time T_m in the complete graph by assuming that the d rates are the same in the complete graph.

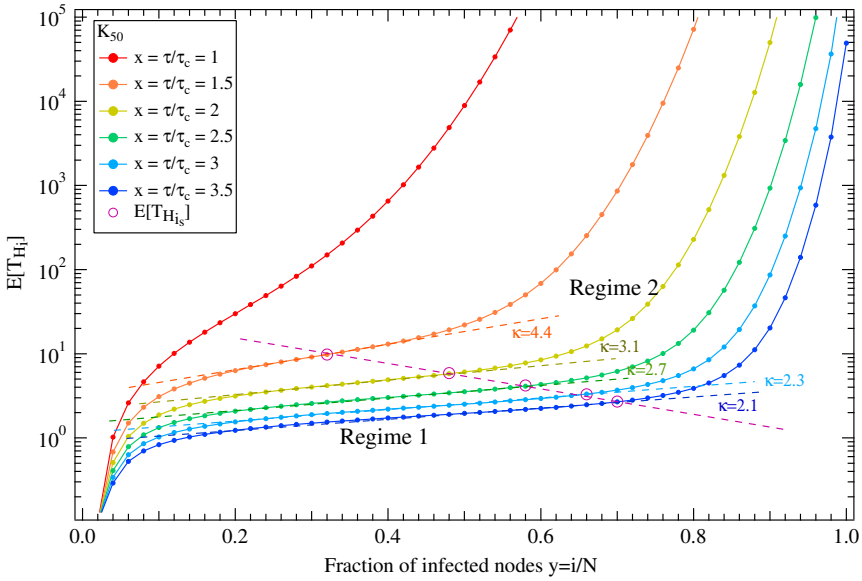

Figure 2: The average hitting time $E[T_{H_i}]$ to the state with i infected nodes in the complete graph K_{50} with different effective infection rate τ, given that there is one initially infected node. The average fraction of infected nodes in the metastable state via NIMFA is marked.

First, we show the spreading time T_m started from one initially infected node in two typical graph including a complete graph K_{50} and a star $K_{1,49}$ with $N = 50$ nodes. Figure 3 and Figure 4 show the spreading time T_m for two values of normalized effective infection rate $x = \tau/\tau_c$ on a log-log scale. The positive skewness of the distribution, derived in Figure 3 – 4, means that the average spreading time $E[T_m]$ is above the mode of the spreading time, which is caused by the rapidly increasing average hitting time $E[T_{H_n}]$ in (2), when the number of infected nodes n exceeds the average number i_s of infected nodes in the metastable state. Comparing the distributions with different normalized effective infection rate x in Figure 3 and Figure 4, the probability of the small value of the spreading time T_m decreases or even disappears with increasing effective infection rate τ. The analytic approximation (12), (13) and Figure 19 accentuates the asymmetry of the distribution of the spreading time T_m.

Further, Figure 5 – 7 show the distributions of the spreading time T_m in an Erdős-Rényi (ER) random graph, a rectangle lattice with $N = 50$ nodes and a BA (Barabási-Albert) power law graph with $N = 1000$ nodes, respectively, where the distribution of the spreading time is influenced by the position of the initially infected spreader and the effective infection rate τ. The parameters of the lognormal pdf are fitted by nonlinear least squarese estimation. Taking the lognormal distribution as a reference distribution in the quantile-quantile plots, we find the spreading time fits the lognormal pdf well when the value of the spreading time is not very large, but deviates in the tail, with a heavier tail than the lognormal distribution. Comparing with Figure (3) with with 5×10^7 realizations, we observe that the head and tail of the distribution fits lognormal pdf better with less realizations (5×10^5 realizations), especially for the value of spreading time with relatively high probability We observe that the more regular the graph is, the better the lognormal pdf fits the distribution of spreading time T_m. That regularity agrees with the governing rule of a lognormal, as the limit distribution of a sum

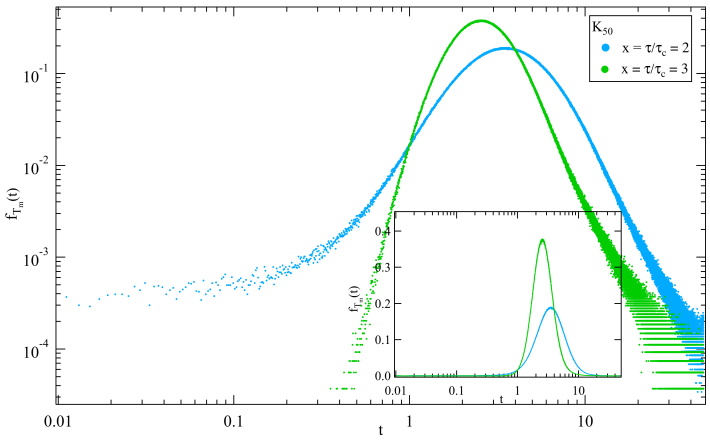

Figure 3: The distribution of spreading time T_m in the complete graph K_{50} with the effective infection rate $x = \tau/\tau_c = 2$, which is based on more than 5×10^7 realizations. Both the axises are on log-scale while only x-axis in the subgraph is on log-scale.

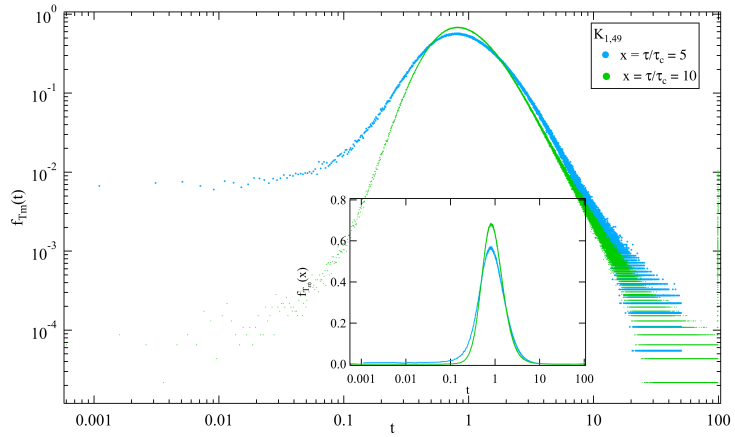

Figure 4: The distribution of spreading time T_m in the star $K_{1,49}$ with the effective infection rate $x = \tau/\tau_c = 5$, which in based on more than 10^7 realizations. Both the axises are on log-scale while only x-axis in the subgraph is on log-scale.

of the logarithm of random variable that each does not differ much [5]. We also mark the stability time t_s via simulation in Figure 5 – 6, which shows that the stability time t_s lies closely to the tail of the distribution of the spreading time T_m, and is larger than the average spreading time $E[T_m]$. We observe that the head and tail of the distribution fits lognormal pdf better with less realizations, especially for the value of spreading time with relatively high probability.

The infection time is exponentially distributed in the classic Markovian SIS process. More generally, we here investigate the spreading time T_m in a non-Markovian process, which is more common in real-world situations, such as information spread in online social networks and real diseases with incubation periods [15]. We assume that the infection and curing processes are independent in a non-Markovian SIS model, where the curing process is still Poissionian with rate δ, and the infection process at each node infects its neighbors in a time T that is Weibullean, with the pdf

$$f_T(x) = \frac{\alpha}{b}\left(\frac{x}{b}\right)^{\alpha-1} e^{-(x/b)^{\alpha}}. \tag{6}$$

In order to compare the Weibull with the exponential distribution, we fix the average infection time to $\frac{1}{\beta}$, so that $b = \left(\Gamma\left(1+\frac{1}{\alpha}\right)\beta\right)^{-1}$. Thus, the shape parameter α tunes the power-law start and the

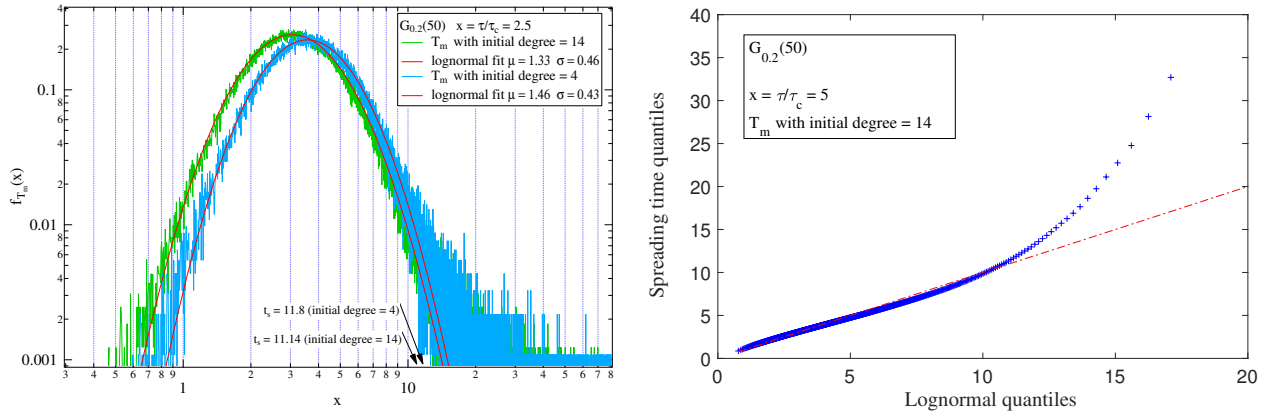

Figure 5: The distribution of the spreading time T_m in a connected ER random graph $G_{0.2}(50)$ with 50 nodes started from the initial spreader with $d_{initial} = 4$ and $d_{initial} = 14$, $x = \tau/\tau_c = 5$. The histogram is based on 5×10^5 realizations.

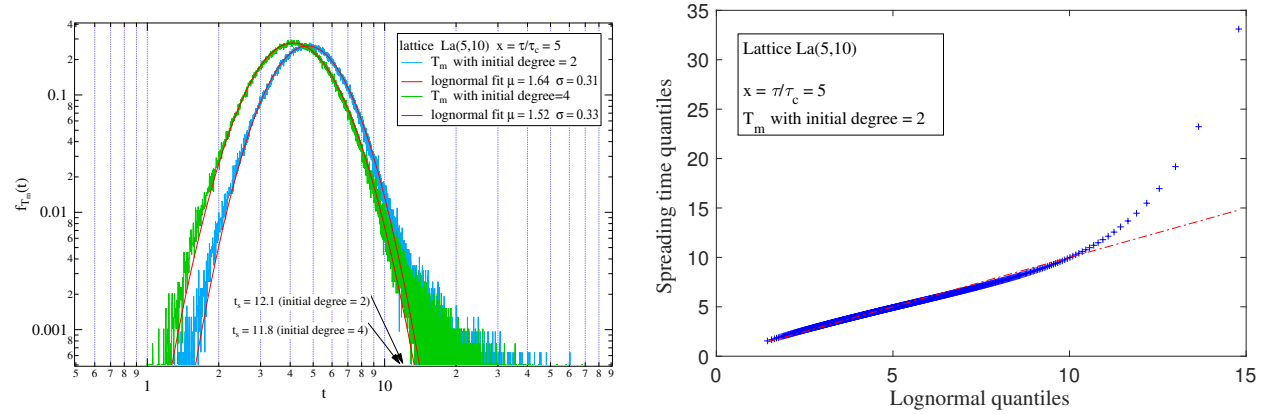

Figure 6: The distribution of the spreading time T_m in a grid $La(5, 10)$ with 5×10 nodes started from the initial node with $d_{initial} = 2$ and $d_{initial} = 4$, $x = \tau/\tau_c = 5$ where $\tau = 1.35$. The histogram is based on 5×10^5 realizations.

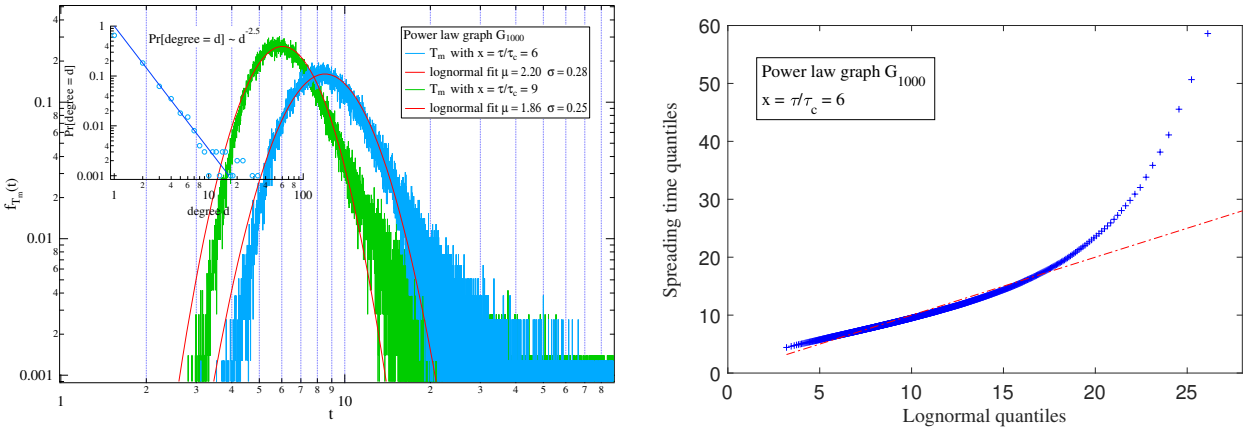

Figure 7: The distribution of the spreading time T_m in a power law graph G_{1000} with 1000 nodes starting from one initial node. The histogram is based on 2×10^5 realizations.

tail of the Weibull distributions with the same mean infection time $E[T] = \frac{1}{\beta}$.

Figure 8 and 9 show the distribution of spreading time T_m as a function of the shape parameter α in a complete graph and a star graph. The pdf of the spreading time remains heavy-tailed, and

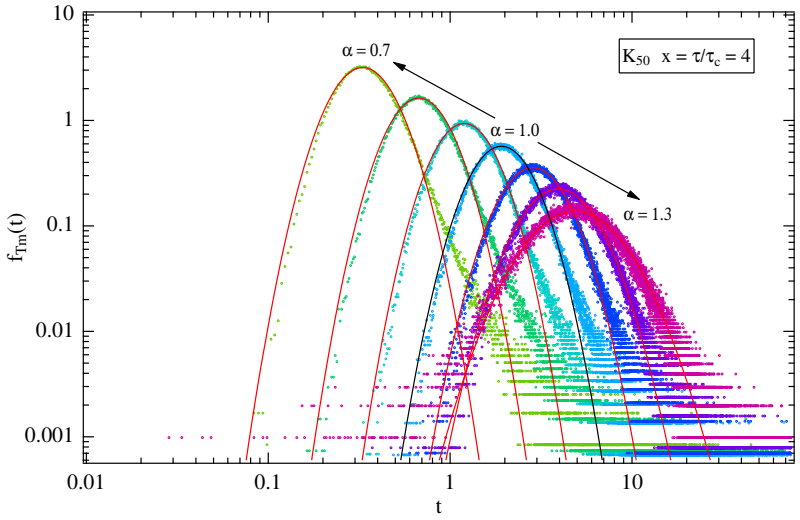

Figure 8: The distribution of the spreading time T_m with the different shape parameter α in the complete graph K_{50} with the effective infection rate $x = \tau/\tau_c = 4$. The exponential case ($\alpha = 1$) is indicated in black. The histograms are based on more than 5×10^5 realizations.

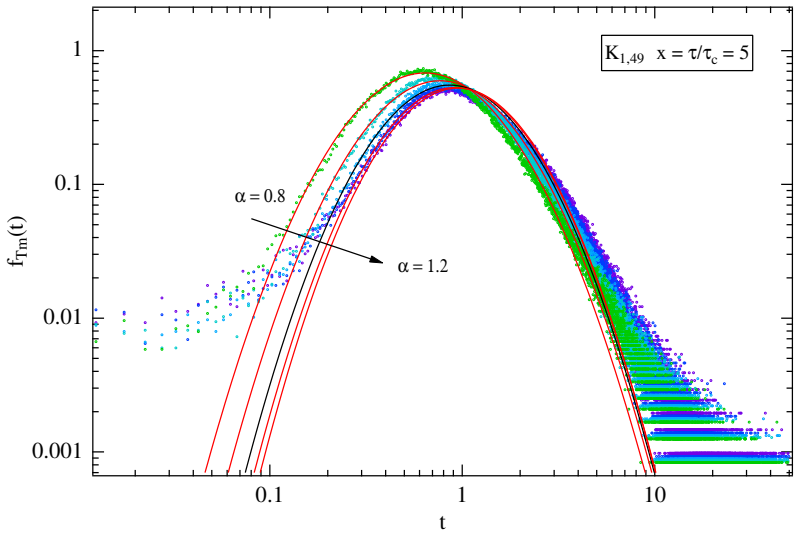

Figure 9: The distribution of the spreading time T_m with the different shape parameter α in the star graph $K_{1,49}$ with the effective infection rate $x = \tau/\tau_c = 5$. The exponential case ($\alpha = 1$) is indicated in black. The histograms are based on more than 5×10^5 realizations.

the shape parameter α shifts the mode of the pdf of the spreading time. The tail of the distribution of the spreading time tends to a lognormal pdf better with the increasing shape parameter α in the complete graph.

The heavy-tailed distribution of the spreading time T_m means that the real average spreading time $E[T_m]$ is larger than the average hitting time $E[T_{H_{i_s}}]$ when the average number of infected nodes in the metastable state is reached. The spreading time has a higher probability to encounter large value resulting in the outliers, and the sample mean usually underestimates the population mean.

4 The average spreading time $E[T_m]$ in SIS epidemic on networks

4.1 Effect of the effective infection rate on $E[T_m]$

We study the average spreading time $E[T_m]$ as a function of the effective infection rate τ in a SIS process, started from a same initially infected node. Figure 10 and Figure 11 illustrate the function of the average spreading time $E[T_m]$ with the effective infection rate τ in a complete graph and a star. The average spreading time $E[T_m]$ is not monotonic with the effective infection rate τ but exhibits a maximum, which means that a stronger virus may not lead to a shorter average spreading time $E[T_m]$.

To better explain the above phenomenon, we define the spreading capacity as $c = \frac{E[I_m]}{E[T_m]}$, which approximately describes the average number of nodes that can be infected in a time unit in the early state of the spreading. Thus, a higher effective infection rate leads to a smaller reciprocal of the spreading capacity $1/c$, which describes the average time units to infect per node. Meanwhile, the average number of infected nodes $E[I_m]$ in the metastable state increases with the effective infection rate τ in a network when the effective infection rate is above the epidemic threshold τ_c. Therefore, the average spreading time $E[T_m]$, which is represented by $E[T_m] = \frac{E[I_m]}{c}$, is influenced by $E[I_m]$ and the spreading capacity c simultaneously, exhibits the property of non-monotony with the effective infection rate τ. The sub-graphs of Figure 10 and Figure 11 illustrate the reciprocal of the spreading capacity $1/c$ and the average number of infected node $E[I_m]$ in the metastable state as a function of the effective infection rate τ.

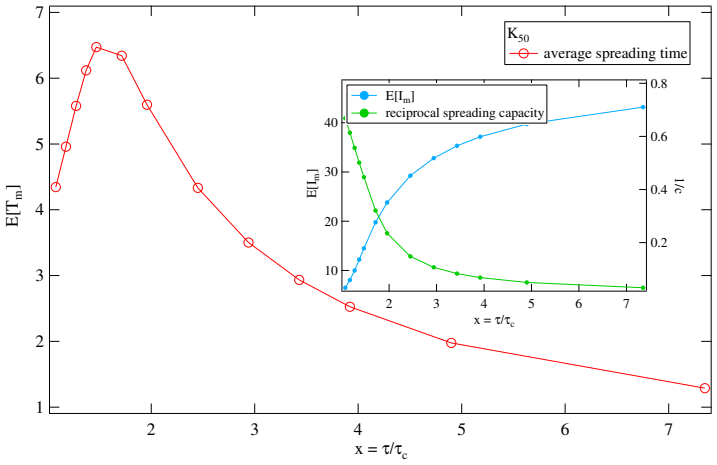

Figure 10: The average spreading time $E[T_m]$ as a function of the effective infection rate $x = \tau/\tau_c$ in a complete graph K_{50}. The subgraph illustrates the average number $E[I_m]$ of infected nodes in the metastable state and the the reciprocal of the spreading capacity $1/c$ with the normalized effective infection rate $x = \tau/\tau_c$.

4.2 Effect of the shape parameter α on $E[T_m]$

We now investigate the effect of the shape parameter α in the Weibull-distributed infection time with pdf (6) on the average spreading time $E[T_m]$, where the Markovian infection process is a special case with $\alpha = 1$. As discussed in Section 4.1, the average spreading time depends on the spreading capacity c and the average fraction $y(t_s)$ of infected nodes in the metastable state, both of which are influenced by the shape parameter α.

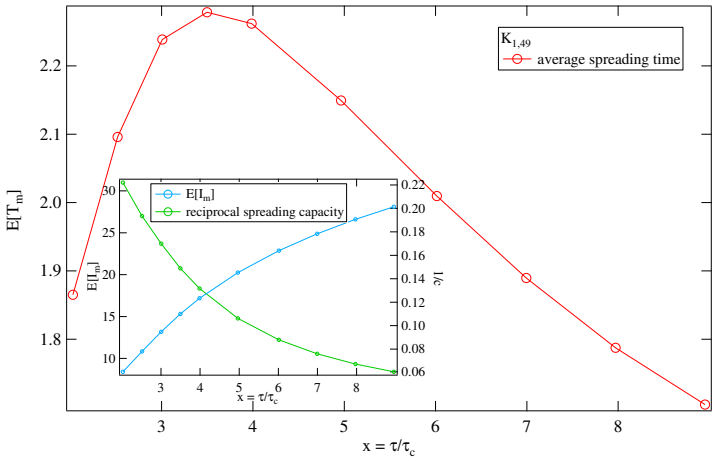

Figure 11: The average spreading time $E[T_m]$ with $x = \tau/\tau_c$ in a star graph $K_{1,49}$ with 49 leaves, started from the center of the graph. The subgraph illustrates the average number $E[I_m]$ of infected nodes in the metastable state and the the reciprocal of the spreading capacity $1/c$ with the normalized effective infection rate $x = \tau/\tau_c$.

The average number of infection attempts during a recovery time is a physically more general description than the effective infection rate in non-Markovian epidemics [15]. Considering the distribution of the infection attempts over an infectious period of a node, the occurrence of events is not uniformly distributed over an interval when the infection process is non-Markovian. For $\alpha < 1$, the infection events tend to happen earlier than the Poission-distributed events (for $\alpha = 1$) with high probability, while for $\alpha > 1$, the infection events tend to happen later. Therefore, the timing of the infection attempts relative to the curing time of a node influences the epidemics process even for the same average number of expected infection attempts [8]. Physically, the reciprocal of the spreading capacity $1/c$, which describes the average time units to infect per node before the metastable state, also increases for a higher α.

Figure 12 shows that the average fraction $y(t_s)$ of infected nodes in the metastable state depends on both the effective infection rate τ and the shape parameter α. Specifically, the average fraction $y(t_s)$ of infected nodes in the metastable state decreases with a higher parameter α for a same effective infection rate τ. Figure 13 suggests that $\log(\tau) \sim \frac{\log(Ny(t_s))}{\alpha}$ for the same number $Ny(t_s)$ of infected nodes in the metastable state, which implies that $\tau^\alpha \sim y(t_s)$ in the complete graph when $\tau < 1$. This relation is consistent with the conclusion that the epidemic threshold $\tau_c(\alpha)$ in the non-Markovian SIS epidemics scales as $(\tau_c^{(1)})^{\frac{1}{\alpha}}$, where $\tau_c^{(1)} = \tau_c(1)$ is the epidemic threshold in the Markovian SIS model [7]. As Figure 13 shows in the star graph, the Weibull shape factor α barely influences the fraction $y(t_s)$ of infected nodes in the metastable state when the effective infection rate $\tau \geq 1$.

Figure 14 shows that, both in the complete graph and the star, the average spreading time $E[T_m]$ does not always increase monotonically with the shape parameter α, but exhibits a maximum when the effective infection rate τ is small. For a higher α, the timing of the infection attempts is postponed while the fraction of infected nodes in the metastable state decreases. These two factors leads to the non-monotonicity of the average spreading time $E[T_m]$ with the shape parameter α, and implies that increasing the parameter α may not shorten the average spreading time $E[T_m]$.

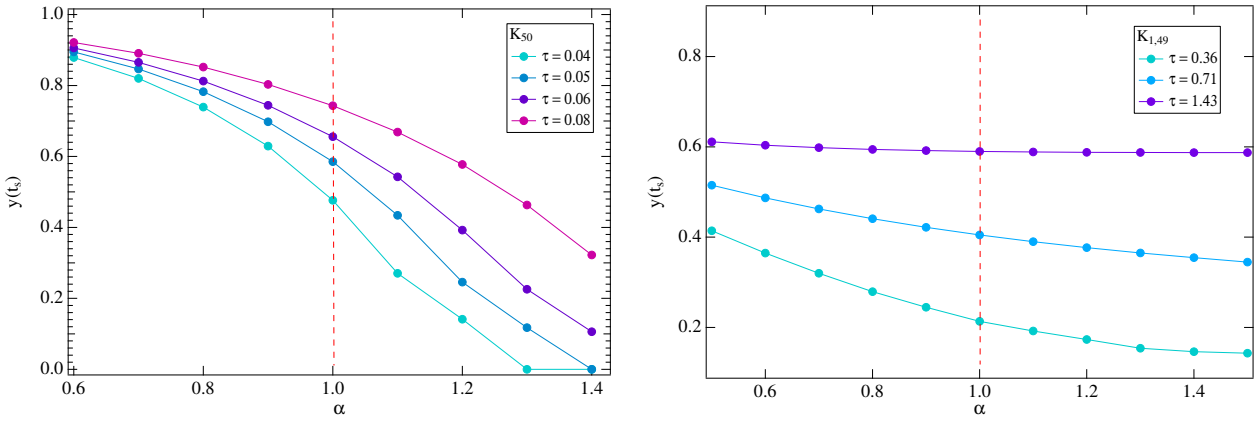

Figure 12: The average fraction of infected nodes in the metastable state for the same τ in the non-Markovian SIS process in a complete graph K_{50} and a star graph $K_{1,49}$.

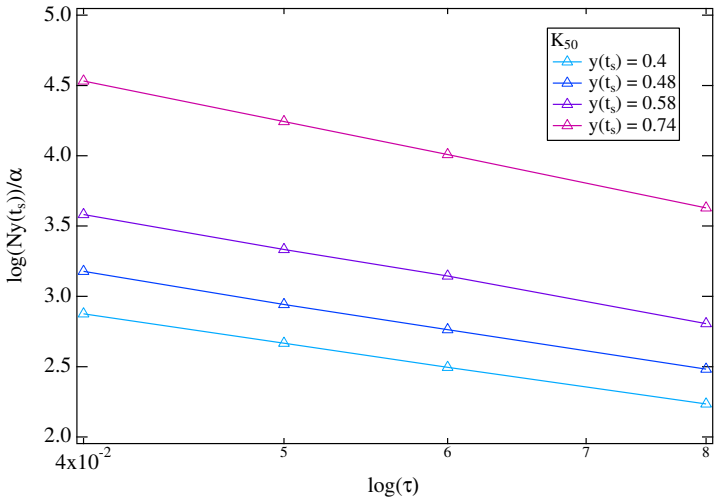

Figure 13: The reciprocal of the parameter α as a function of $log(\tau)$ in the complete graph K_{50} for the same fraction of infected nodes in the metastable state.

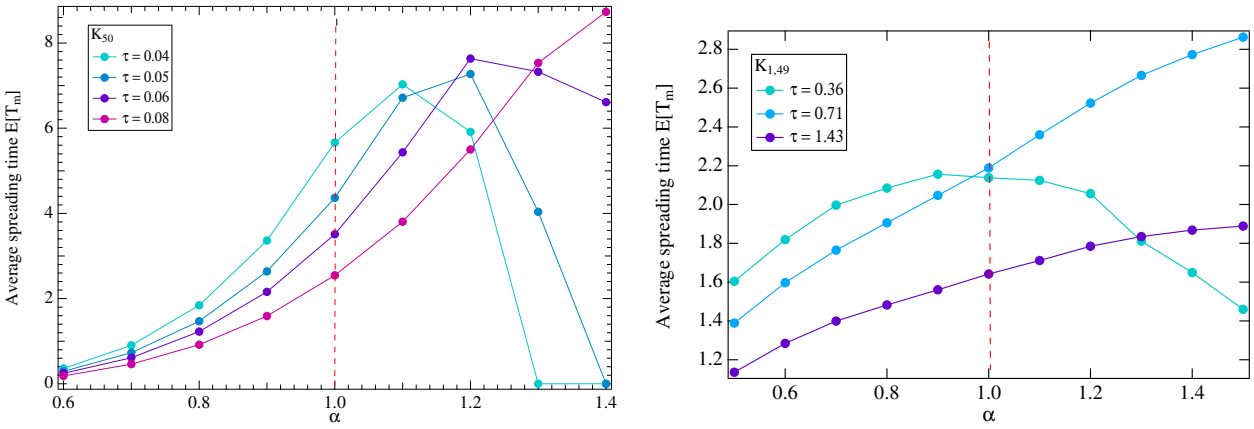

Figure 14: The average spreading time $E[T_m]$ as a function of the parameter α for the same effective infection rate τ in the complete graph K_{50} and in the star graph $K_{1,49}$

4.3 Effect of the network size on $E[T_m]$

We now investigate the effect of the network size N on the average spreading time $E[T_m]$. Figure 15 – 17 show the average spreading time $E[T_m]$ starting from one initially infected node as a function of the network size for a complete graph K_N, a star $K_{1,N}$, and an ER random graph $G_p(N)$. Referring to the average fraction of infected nodes $y^{(1)}(t_s) = 1 - \frac{1}{(N-1)\tau}$ in the metastable state in a complete graph with N nodes via NIMFA [14], we can estimate the effective infection rate $\tau = \frac{1}{(1-y(t_s))(N-1)}$ for a fixed average fraction $y(t_s)$ of infected nodes in the metastable state. Similarly, in an ER graph, the effective infection rate $\tau = \frac{1}{(1-y(t_s))(N-1)p}$ for a fixed fraction $y(t_s)$ of infected nodes in the metastable state is estimated by the NIMFA approximation $y^{(1)}(t_s) = 1 - \frac{1}{(N-1)p\tau}$, where the link probability $p = \frac{2\log N}{N}$. The effective infection rate τ in a star is estimated by the NIMFA approximation [14] that $y^{(1)}(t_s) = \frac{N-\tau^{-2}}{N+1}\left\{\frac{1}{\tau^{-1}+1} + \frac{1}{\tau^{-1}+N}\right\} \approx \frac{\tau}{1+\tau}$ when $N \gg \tau$.

We ignore the curing events and consider a Susceptible-Infected (SI) process in the complete graph. The average time when I_m nodes are infected [5] follows $\sum_{n=1}^{I_m} \frac{1}{\tau n(N-n)}$, where $y(t_s) = \frac{I_m}{N} \approx 1 - \frac{1}{N\tau}$ is fixed. Thus, we obtain

$$E[T_m] \approx \sum_{n=1}^{I_m} \frac{1}{\tau n(N-n)} = \frac{2}{\tau N} \sum_{n=1}^{I_m} \frac{1}{n} \sim 2(1-y(t_s)) \log(y(t_s)N), \tag{7}$$

which scales logarithmically with the network size N. For an SIS process, Figure 15 – 17 show that the average spreading time $E[T_m]$ via simulation approximately scales logarithmically as $a\log(N) + b$ for different fractions $y(t_s)$ of infected nodes in the metastable state in a complete graph, an ER random graph and a star. The process needs more time to infect a same fraction of nodes in a network with a larger size. The slope a of the fit is larger for a smaller fraction $y(t_s)$ of infected nodes in the metastable state, which means the average spreading time $E[T_m]$ tends to increase more quickly with the network size N when the fraction $y(t_s)$ of infected nodes in the metastable state is smaller. We observe the similar trend of the average spreading time with the network size N in the complete graph and the ER random graph. Actually, when the link density p in an ER random graph is above the critical link density $p_c = \frac{\log N}{N}$, the graph is already dense and follows similar behaviour as the complete graph [5].

5 Conclusion

We define the spreading time as the time when the number of infected nodes in the metastable state is first reached, starting from the outbreak of a spreading.

We investigated the distribution of the spreading time. The average hitting time $E[T_{H_i}]$ to the state i around the average number of infected nodes in the metastable state approximates an exponential function, where the number of infected nodes in the metastable state resembles a Gaussian-like distribution. Thus, we observe that the spreading time T_m resembles a lognormal-like heavy-tailed distribution, which is exhibited both in the Markovian and the non-Markovian infection process.

We further investigated the properties of the average spreading time. Because the number of infected nodes in the metastable state and the spreading capacity are influenced by the effective infection rate simultaneously, the average spreading time $E[T_m]$ is not necessarily monotonous with

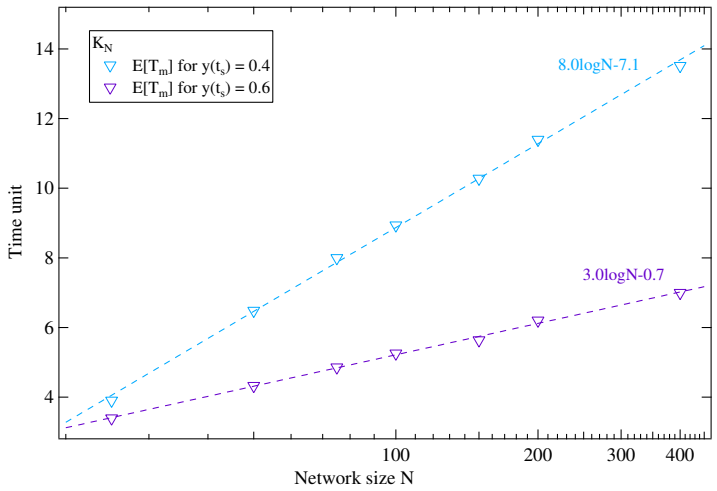

Figure 15: The average spreading time $E[T_m]$ starting from one initially infected node as a function of the network size in the complete graph K_N.

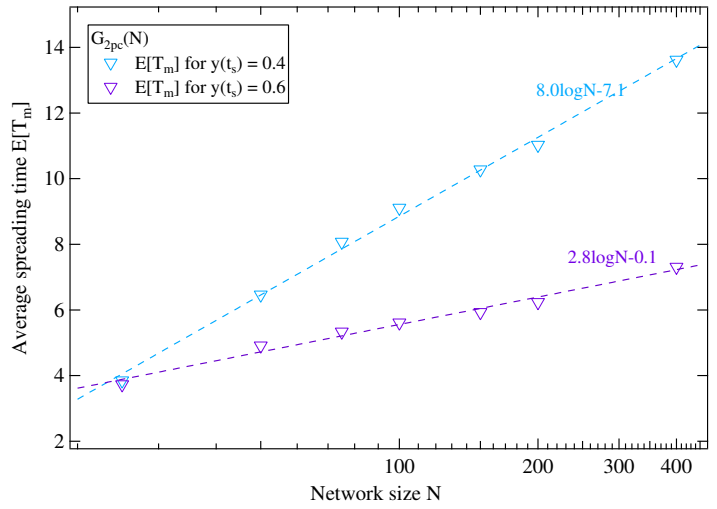

Figure 16: The average spreading time $E[T_m]$ starting from one initially infected node with the average degree as a function of the network size in the ER random graph $G_{2p_c}(N)$.

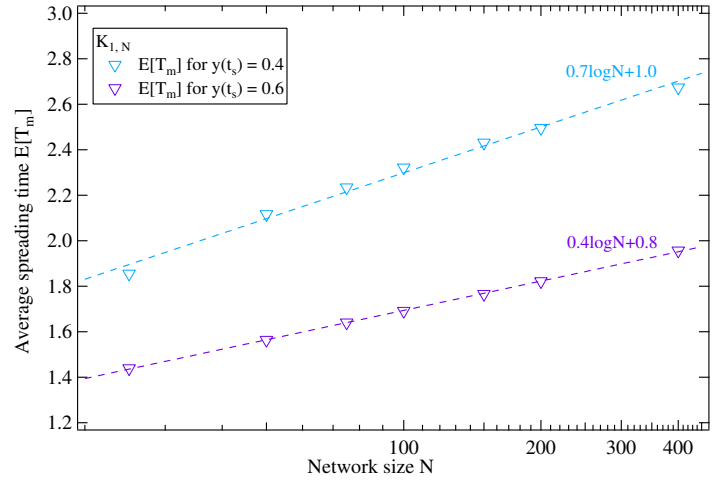

Figure 17: The average spreading time $E[T_m]$ starting from the center as a function of the network size in the star graph $K_{1,N}$.

the effective infection rate τ but exhibits a maximum, which means that a higher effective infection rate τ may not lead to a shorter average spreading time $E[T_m]$. Similarly, both the fraction of infected nodes in the metastable state and the timing of the infection attempts are influenced simultaneously by the parameter α of the Weilbullean infection times, which leads to non-monotonicity of the average spreading time $E[T_m]$ with the shape parameter α. Finally, we showed that the average spreading time $E[T_m]$ scales logarithmically as a function of the network size N, given that the average fraction $y(t_s)$ of infected nodes in the metastable state is fixed.

Acknowledgement

This research is supported by the China Scholarship Council (CSC). We thank for M.Märtens, H.Cetinay and X.Wang for useful comments.

References

[1] R. Pastor Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks," *Reviews of modern physics*, vol. 87, no. 3, p. 925, 2015.

[2] A. L. Hill, D. G. Rand, M. A. Nowak, and N. A. Christakis, "Emotions as infectious diseases in a large social network: the sisa model," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 277, no. 1701, pp. 3827–3835, 2010.

[3] P. Van Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," *IEEE/ACM Transactions on Networking*, vol. 17, no. 1, pp. 1–14, 2009.

[4] L. Buzna, K. Peters, and D. Helbing, "Modelling the dynamics of disaster spreading in networks," *Physica A: Statistical Mechanics and its Applications*, vol. 363, no. 1, pp. 132–140, 2006.

[5] P. Van Mieghem, *Performance analysis of complex networks and systems*. Cambridge University Press, 2014.

[6] C. Castellano and R. Pastor-Satorras, "Thresholds for epidemic spreading in networks," *Physical Review Letters*, vol. 105, no. 21, p. 218701, 2010.

[7] P. Van Mieghem and R. Van de Bovenkamp, "Non-markovian infection spread dramatically alters the susceptible-infected-susceptible epidemic threshold in networks," *Physical Review Letters*, vol. 110, no. 10, p. 108701, 2013.

[8] R. Van de Bovenkamp and P. Van Mieghem, "Survival time of the susceptible-infected-susceptible infection process on a graph," *Physical Review E*, vol. 92, no. 3, p. 032806, 2015.

[9] P. Van Mieghem, "Approximate formula and bounds for the time-varying susceptible-infected-susceptible prevalence in networks," *Physical Review E*, vol. 93, no. 5, p. 052312, 2016.

[10] P. Van Mieghem, "Decay towards the overall-healthy state in sis epidemics on networks," *arXiv preprint arXiv:1310.3980*, 2013.

[11] J. R. Artalejo, "On the time to extinction from quasi-stationarity: A unified approach," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 19, pp. 4483–4486, 2012.

[12] R. Van de Bovenkamp and P. Van Mieghem, "Time to metastable state in sis epidemics on graphs," in *Tenth International Conference on Signal-Image Technology and Internet-Based Systems (SITIS)*. IEEE, 2014, pp. 347–354.

[13] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *The journal of physical chemistry*, vol. 81, no. 25, pp. 2340–2361, 1977.

[14] E. Cator and P. Van Mieghem, "Susceptible-infected-susceptible epidemics on the complete graph and the star graph: Exact analysis," *Physical Review E*, vol. 87, no. 1, p. 012811, 2013.

[15] E. Cator, R. Van de Bovenkamp, and P. Van Mieghem, "Susceptible-infected-susceptible epidemics on networks with general infection and cure times," *Physical Review E*, vol. 87, no. 6, p. 062816, 2013.

A Appendix A: Determination of the metastable state and the stability time

We define a Bernoulli random variable $X_i(t) \in \{0,1\}$ as the infectious state of node i, where $X_i(t) = 1$ indicates that node i is infected and $X_i(t) = 0$ indicates that node i is susceptible at time t. The prevalence $y(t) = \frac{1}{N}E[I(t)]$ of an SIS process is the expected fraction of infected nodes at time t, where $I(t) = \sum_{i=1}^{N} X_i(t)$ is the number of infected nodes. We present several definitions of the metastable state in the SIS process on finite graphs derived from the prevalence $y(t)$ in this section.

Definition 1(a): In an epidemic process, the metastable state is reached at the stability time t_s, which is the smallest time obeying $\frac{dy(t)}{dt}\big|_{t=t_s} = 0$.

It seems reasonable to define the start of the metastable state when the prevalence $y(t)$ reaches its first extremum. However, the SIS prevalence $y(t)$, started from multiple initial spreaders, may pass multiple extrema in the transient regime in a specific network, which demonstrates that Definition 1(a) is not precise. In addition, as shown in Figure 18, the prevalence $y(t)$ may monotonically decreases when the average number of infected nodes in the metastable state is smaller than the number of the initially infected nodes. Therefore, this definition may not be adequate for the computation of the spreading, starting from multiple initially infected nodes.

Definition 1(b): In an epidemic process, the metastable state is reached at the stability time t_s, which is the smallest time obeying $\frac{dy(t)}{dt}\big|_{t=t_s} = 0$, and $|y(t) - y(t_s)| \leq \epsilon$ for $\forall t > t_s + \alpha E[T_{absorbing}]$, where $0 < \alpha < 1$. The positive real numbers α and ϵ need to be agreed upon.

To remedy the defect of Definition 1(a), we try to bound the prevalence $y(t)$ in an interval around the the fraction $y(t_s)$ of infected node at the stability time t_s. However, the prevalence $y(t)$ will inevitably exceed the bound because the prevalence will reach an absorbing state $y(t) = 0$ finally. Therefore, it is hard to determine the two parameters ϵ and α that allows $|y(t) - y(t_s)| \leq \epsilon$ for $\forall t > t_s + \alpha E[T_{absorbing}]$.

Definition 1(c): In an epidemic process, the metastable state is reached at the stability time t_s, which is the smallest time obeying $y(t_s|i) = y(t_s|N)$.

Because we cannot bound the prevalence $y(t)$ in the metastable state, we consider to use the prevalence $y(t|N)$ started from $I_0 = N$ initial nodes as a reference curve and locate the start of the metastable state as the intersection point in time with the prevalence $y(t|I_0 = i)$, where the prevalence $y(t|N)$ with all initial spreaders converges fastest to the metastable state.

Figure 18 shows that there exists a gap between the prevalences $y(t|1)$ and $y(t|N)$ in the metastable state due to the different probability of extinction, which means that the prevalence started from a different number of initial nodes will not intersect before the absorbing state. Figure 18 also shows that the difference between the two prevalences $y(t|1)$ and $y(t|N)$ becomes narrower with the time, which implies that the decreasing rate of the prevalence is also influenced by the initial infection condition. We expect that all the prevalences $y(t|I_0)$ with $I_0 \in (1, 2, \ldots, N)$ initially infected nodes will meet only in the absorbing state, which demonstrates the infeasibility to locate the metastable state by the intersection of the prevalence curves.

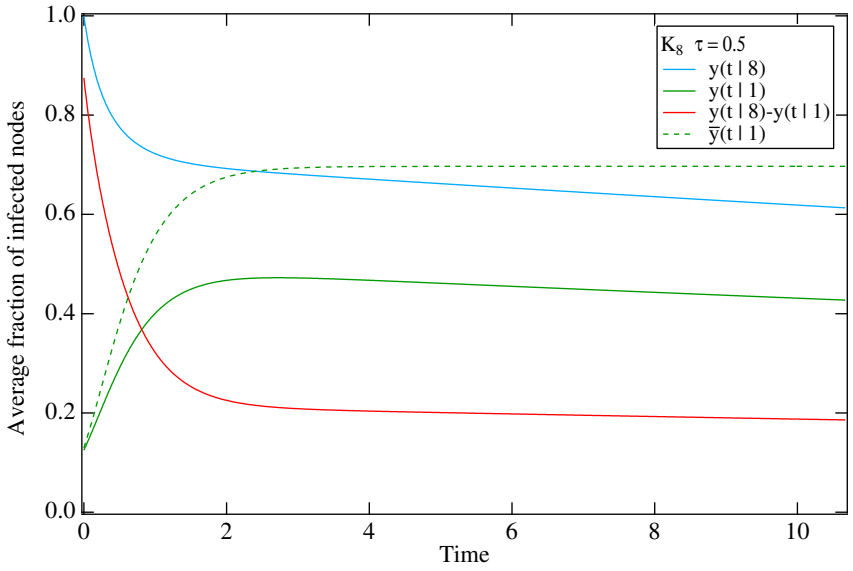

Figure 18: The exact prevalences $y(t|1)$ started from one infected node and the exact prevalence $y(t|8)$ started from all infected nodes in a complete graph K_8 with the effective infection rate $\tau = 0.5$. The red line represents the difference between the two prevalences $y(t|1)$ and $y(t|8)$. The green dash line represents the prevalence $\overline{y}(t|1)$ excluding early extinction probability.

Definition 1(d): In an epidemic process, the metastable state is reached at the stability time t_s, which is the first time obeying $\frac{dy(t)}{dt}\big|_{t>t_s} < 0$.

Definition 1(d) means that the last extremum of the prevalence $y(t)$ is located as the start of the metastable state, and the average fraction of infected nodes monotonically decreases after the stability time t_s. The prevalence $y(t)$ is the average fraction of infected nodes, which includes the realizations that die out early as well as the realizations that reach the metastable state, thus

$$y(t) = y(t)\big|_{I(t)>0} \Pr[I(t) > 0] + y(t)\big|_{I(t)=0} \Pr[I(t) = 0] \tag{8}$$

The fraction $y(t)\big|_{I(t)>0}$ approximates gradually the fraction $y(t)$ with the decreasing extinction probability $\Pr[I(t) = 0]$. However, the extinction probability $\Pr[I(t) = 0]$ is hard to estimate in a general network mathematically.

Definition 1(e): In an epidemic process, the metastable state is reached at the stability time t_s,

which is the smallest time obeying $\frac{d\overline{y}(t)}{dt}\big|_{t>t_s} < \epsilon$, where the average fraction of infected nodes is $\overline{y}(t) = \frac{1}{N}E[I(t)]$, with $I(t) \geq 1$ is the number of infected nodes at time t, and ϵ is a small positive real number that needs to be agreed upon.

In the above definition, we introduce the prevalence

$$\overline{y}(t) = y(t)\big|_{I(t)>0} = \frac{y(t)}{1 - \Pr[I(t) = 0]} \tag{9}$$

subject to the condition that the process does not die out, where $\Pr[I(t) = 0]$ is the extinction probability. The prevalence $\overline{y}(t)$ excluding early extinction, as illustrated in Figure 18, tends to stay almost constant instead of decaying as the prevalence $y(t)$ after reaching the extremum. We consider that the metastable state starts when the prevalence $\overline{y}(t)$ stays almost constant. Actually, the prevalence $\overline{y}(t)$ excluding early extinction is a monotonically increasing function, which only stays constant when $t \to \infty$, as follows from general Markov theory [5]. The prescribed stringent parameter ϵ can be determined as a small value. Definition 1(e) is also consistent with the definition of the quasi-stationary state, which leads to the almost steady average number of infected nodes without extinction realizations.

In summary, we choose Definition 1(e) as our preferred definition of the metastable state and the stability time t_s in this paper.

B Appendix B: Simulation for a SIS process on networks

There are two kinds of events in the SIS process which are infection events and curing events. All the events are marked on a same timeline and are handled by the order of their time after the beginning of the simulation. The process of SSIS (Simulation for SIS epidemics) is described by Algorithm 1.

Algorithm 1 Simulation for SIS epidemics

1) Initialize the graph G_N with N nodes, and set the initial condition including the infection rate β, the curing rate δ, the initial spreader(s) I_0, the time limit t_{limit}.

2) Insert the initial infected events in the timeline with time $t_0 = 0$ for the initial infected nodes.

3) Find the earliest un-handled events in the timeline, denoted $EVT_n(t)$ the event for node n at time t.

4) If $EVT_n(t)$ is an infection event,

a) node n is infected, and the number of infected nodes $i(t) = i(t) + 1$;

b) the new curing event $EVT_n(t^{'})$ for node n is added in the time line after an exponentially distributed random time interval with mean $1/\delta$, thus $t^{'} = t + ramdom(1/\delta)$;

c) Each susceptible neighbor of node n is allocated an infection time as $t^{''} = t + ramdom(1/\beta)$. If the infection times for these neighbors $t^{''} < t^{'}$, the infection event of these neighbors are added on the timeline with time $t^{''}$.

5) If $EVT_n(t)$ is an curing event, node n will return to the susceptible state, and the number of infected nodes $i(t) = i(t) - 1$.

6) If the time of event $t > t_{limit}$, the simulation finishes. Else, return step 3).

We set the parameter $\epsilon = 0.01$ in Definition 1(e) and run the SSIS repeat for an unaltered graph,

a fixed effective infection rate τ and the same initial condition. The prevalence can be obtained by $\overline{y}(t) = \frac{1}{N}E[I(t)]$, where the random variable $I(t)$ denotes the number of infected nodes $i(t)$ in all realizations. Then we determine the stability time t_s as the first time when $\frac{\overline{y}(t_s+\Delta t)-\overline{y}(t_s)}{\Delta t} < \epsilon$, where the time sample interval $\Delta t = 0.01$ in our simulations.

C Appendix C: Analytic approximation of the spreading time T_m

The spreading time T_m, given that initially $I(0) = i$ nodes are infected, can be defined as

$$f_{T_m}(t) = \sum_{n=0}^{\infty} f_{T_m}(t|I(t_s) = n) \Pr[I(t_s) = n]$$

where we have omitted in the condition that $I(0) = i$ to not overload the notation. In the sequel, we confine to the case of $i = 1$ initially infected node.

Define the vector $w(t) = (X_1(t), X_2(t), \ldots, X_N(t))$ and $X_k(t) \in \{0, 1\}$ is the infectious state of node k at time t. Physically, $\Pr[T \leq x|I(\theta) = n, w(0)]$ is the probability that the hitting time, given that the process started with an initial condition $w(0)$, reaches a state in the Markov process that contains precisely n infected nodes. If the effective infection rate τ is sufficiently above the epidemic threshold τ_c, then $\Pr[I(t_s) = n] \approx \frac{1}{\tilde{\sigma}\sqrt{2\pi}} \exp\left[-\frac{(n-\tilde{\mu})^2}{2\tilde{\sigma}^2}\right]$ is close to a Gaussian (see e.g. [5, p. 474] and [14]). Generally, the hitting time in a Markov process consists of a random sum of exponentially distributed random variables, because the time between events in any continuous-time Markov process is exponentially distributed [5, p. 210]. Since the process is assumed to start with one infected node $I(0) = 1$ and since each Markov event can at best increase the number of infected nodes by one, we approximate the hitting time as a sum S_n of n independent exponential random variables. Hence, we assume that the spreading process is as efficient as a shortest path approach, which is justified for an effective infection rate τ is sufficiently above the epidemic threshold τ_c

Let us investigate this claim and study

$$f_{T_m}(t) = \frac{1}{\tilde{\sigma}\sqrt{2\pi}} \sum_{n=0}^{\infty} f_{S_n}(t) \exp\left[-\frac{(n-\tilde{\mu})^2}{2\tilde{\sigma}^2}\right] \tag{10}$$

Using the general expression [5, p. 45] for the pdf of a sum of independent random variables each with rate η_k, we obtain

$$f_{T_m}(t) = \frac{1}{\tilde{\sigma}\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{zt} \prod_{k=1}^{n} \frac{\eta_k}{z+\eta_k} \exp\left[-\frac{(n-\tilde{\mu})^2}{2\tilde{\sigma}^2}\right] dz$$

which we rewrite, for $c > 0$,

$$f_{T_m}(t) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{zt} \left\{ \frac{1}{\tilde{\sigma}\sqrt{2\pi}} \sum_{n=0}^{\infty} \exp\left(-\frac{(n-\tilde{\mu})^2}{2\tilde{\sigma}^2} + \sum_{k=1}^{n} \log\frac{\eta_k}{z+\eta_k}\right)\right\} dz$$

where

$$\varphi_{T_m}(z) = \frac{1}{\tilde{\sigma}\sqrt{2\pi}} \sum_{n=0}^{\infty} \exp\left(-\frac{(n-\tilde{\mu})^2}{2\tilde{\sigma}^2} + \sum_{k=1}^{n} \log\frac{\eta_k}{z+\eta_k}\right) \tag{11}$$

is the probability generating function of the spreading time T_m.

Since the sum in (11) is hard to compute, we confine ourselves to the case where all rates $\eta_k = \eta > 0$ are the same. Directly from the definition (10) of the spreading time and using the Erlang distribution [5, p. 45], we obtain

$$f_{T_m}(t) = \frac{\eta e^{-\eta t}}{\tilde{\sigma}\sqrt{2\pi}} \sum_{n=0}^{\infty} \frac{(\eta t)^{n-1}}{(n-1)!} \exp\left[-\frac{(n-\tilde{\mu})^2}{2\tilde{\sigma}^2} \right]$$

and

$$f_{T_m}(t) = \frac{\eta e^{-\eta t}}{\tilde{\sigma}\sqrt{2\pi}} \left\{ \exp\left[-\frac{(n-\tilde{\mu})^2}{2\tilde{\sigma}^2} \right] + \sum_{n=0}^{\infty} \frac{(\eta t)^n}{n!} \exp\left[-\frac{(n+1-\tilde{\mu})^2}{2\tilde{\sigma}^2} \right] \right\} \tag{12}$$

which illustates that $f_{T_m}(0) = \frac{\eta}{\tilde{\sigma}\sqrt{2\pi}} \exp\left[-\frac{(\tilde{\mu}-1)^2}{2\tilde{\sigma}^2} \right]$. On the other hand, the probability generating function in (11) reduces to

$$\varphi_{T_m}(z) = \frac{1}{\tilde{\sigma}\sqrt{2\pi}} \sum_{n=0}^{\infty} s^n \exp\left[-\frac{(n-\tilde{\mu})^2}{2\tilde{\sigma}^2} \right]$$

where $s = \frac{\eta}{z+\eta}$. Further, we approximate the sum by an integral

$$\varphi_{T_m}(z) \approx \frac{1}{\tilde{\sigma}\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{w \log s} \exp\left[-\frac{(w-\tilde{\mu})^2}{2\tilde{\sigma}^2} \right] dw$$

because $\tilde{\mu}$ is sufficiently large (for a sufficiently large effective infection rate τ) allowing to extend the lower boundary from 0 to $-\infty$. Using the Laplace transform [5, p. 43] gives

$$\varphi_{T_m}(z) \approx \exp\left(\frac{\tilde{\sigma}^2}{2} (\log s)^2 + \tilde{\mu} \log s \right)$$

which is "lognormal-like" in s. With $s^{-1} = 1 + \frac{z}{\eta}$, the approximate pgf becomes

$$\varphi_{T_m}(z) \approx \exp\left(-\frac{\tilde{\mu}^2}{2\tilde{\sigma}^2} \right) \exp\left(\frac{\tilde{\sigma}^2}{2} \left(\log\left(1 + \frac{z}{\eta} \right) - \frac{\tilde{\mu}}{\tilde{\sigma}^2} \right)^2 \right)$$

Inverse Laplace transformation leads to another representation of the probability density function

$$f_{T_m}(t) \approx e^{-\frac{\tilde{\mu}^2}{2\tilde{\sigma}^2}} \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{zt} \exp\left(\frac{\tilde{\sigma}^2}{2} \left(\log\left(1 + \frac{z}{\eta} \right) - \frac{\tilde{\mu}}{\tilde{\sigma}^2} \right)^2 \right) dz$$

We can move the line of integration to $c' = 1 + \frac{c}{\eta}$ and, finally, we arrive at the complex integral

$$f_{T_m}(t) \approx \eta e^{-\eta t} \frac{1}{2\pi i} \int_{c'-i\infty}^{c'+i\infty} e^{z\eta t} z^{-\tilde{\mu}} e^{\frac{\tilde{\sigma}^2}{2} \log^2 z} dz \qquad (c' > 1) \tag{13}$$

where both t and η are non-negative. The integrand $g(z) = e^{z\eta t} z^{-\tilde{\mu}} e^{\frac{\tilde{\sigma}^2}{2} \log^2 z}$ has a branch cut along the negative real axis and is analytic everywhere else, so that $c' > 0$. Since $\lim_{r\to\infty} g\left(re^{i\theta}\right) = 0$ when $\frac{\pi}{2} < \theta < \frac{3\pi}{2}$, we deform the contour to enclose the branch cut from above, circle with radius ε around the origin in clockwise sense, and turn back to infinity under the negative real axis (branch cut). The contour L consists of the vertical integration along the imaginary axis at the real c', then the circle

with $r \to \infty$ with vanishingly small contribution and line around the cut. This contour L encloses a region that is analytic so that the corresponding integral is zero. Thus, we have

$$\begin{aligned}\frac{1}{2\pi i}\int_{c'-i\infty}^{c'+i\infty} e^{z\eta t} z^{-\tilde{\mu}} e^{\frac{\tilde{\sigma}^2}{2}\log^2 z} dz = &-\frac{e^{i\pi}}{2\pi i}\int_{\infty}^{\varepsilon} e^{xe^{i\pi}\eta t}\left(xe^{i\pi}\right)^{-\tilde{\mu}} e^{\frac{\tilde{\sigma}^2}{2}\log^2\left(xe^{i\pi}\right)} dx \\ &- \frac{\varepsilon^{-\tilde{\mu}+1}}{2\pi}\int_{\pi}^{-\pi} e^{\varepsilon e^{i\theta}\eta t} e^{i(-\tilde{\mu}+1)\theta} e^{\frac{\tilde{\sigma}^2}{2}\log^2\left(\varepsilon e^{i\theta}\right)} d\theta \\ &- \frac{e^{-i\pi}}{2\pi i}\int_{\varepsilon}^{\infty} e^{xe^{-i\pi}\eta t}\left(xe^{-i\pi}\right)^{-\tilde{\mu}} e^{\frac{\tilde{\sigma}^2}{2}\log^2\left(xe^{-i\pi}\right)} dx\end{aligned}$$

Simplifying the integrals leads to

$$\begin{aligned}\frac{f_T(t)}{\eta e^{-\eta t}} \approx &\frac{e^{-\pi^2\frac{\tilde{\sigma}^2}{2}}}{\pi}\int_{\varepsilon}^{\infty} e^{-x\eta t} x^{-\tilde{\mu}} e^{\frac{\tilde{\sigma}^2}{2}\log^2 x} \sin\left(\pi\left(\tilde{\sigma}^2\log x + (1-\tilde{\mu})\right)\right) dx \\ &- \frac{\varepsilon^{1-\tilde{\mu}} e^{\frac{\tilde{\sigma}^2}{2}\log^2\varepsilon}}{2\pi}\int_{\pi}^{-\pi} e^{\varepsilon e^{i\theta}\eta t} e^{i(1-\tilde{\mu})\theta} e^{i\theta\tilde{\sigma}^2\log\varepsilon} e^{-\frac{\tilde{\sigma}^2}{2}\theta^2} d\theta\end{aligned}$$

The latter term diverges when $\varepsilon \to 0$, because $e^{\frac{\tilde{\sigma}^2}{2}\log^2 z}$ has an essential singularity at $z = 0$. When we choose $\varepsilon = 1$, then

$$\lim_{\varepsilon\to 1}\frac{\varepsilon^{1-\tilde{\mu}} e^{\frac{\tilde{\sigma}^2}{2}\log^2\varepsilon}}{2\pi}\int_{\pi}^{-\pi} e^{\varepsilon e^{i\theta}\eta t} e^{i(1-\tilde{\mu})\theta} e^{i\theta\tilde{\sigma}^2\log\varepsilon} e^{-\frac{\tilde{\sigma}^2}{2}\theta^2} d\theta = \frac{1}{2\pi}\int_{\pi}^{-\pi} e^{e^{i\theta}\eta t} e^{i(1-\tilde{\mu})\theta} e^{-\frac{\tilde{\sigma}^2}{2}\theta^2} d\theta$$

and further evaluation finally yields

$$\begin{aligned}\frac{f_T(t)}{\eta e^{-\eta t}} \approx &\frac{e^{-\pi^2\frac{\tilde{\sigma}^2}{2}}}{\pi}\int_{1}^{\infty} e^{-x\eta t} x^{-\tilde{\mu}} e^{\frac{\tilde{\sigma}^2}{2}\log^2 x} \sin\left(\pi\left(\tilde{\sigma}^2\log x + (1-\tilde{\mu})\right)\right) dx \\ &+ \frac{1}{\pi}\int_{0}^{\pi} e^{\eta t\cos\theta} e^{-\frac{\tilde{\sigma}^2}{2}\theta^2} \cos\left(\eta t\sin\theta + (1-\tilde{\mu})\theta\right) d\theta\end{aligned} \tag{14}$$

Numerical integration of (14) agrees excellently[1] with (12). Moreover, (14) is numerically more stable for large t than (12). Figure 19 shows both the analytic approximation (14) and simulations of the spreading time T_m (after proper scaling of the unknown rate η so that the maximum occurs at the same time). The agreement is not convincing, which is most likely due to the confinement to a same rate η for each transition in the Markov process. The variability in rate (instead of all rates the same) will increase the variability in T_m, which will result in a broader bell-shape (i.e. a larger $Var[T_m]$) and likely agree better with simulations. The analytic approximation in Figure 19 accentuates the asymmetry in T_m, which may indicate that the spreading time T_m is only approximately a lognormal distribution.

[1] Actually, nearly all digits (up to 10) are the same!

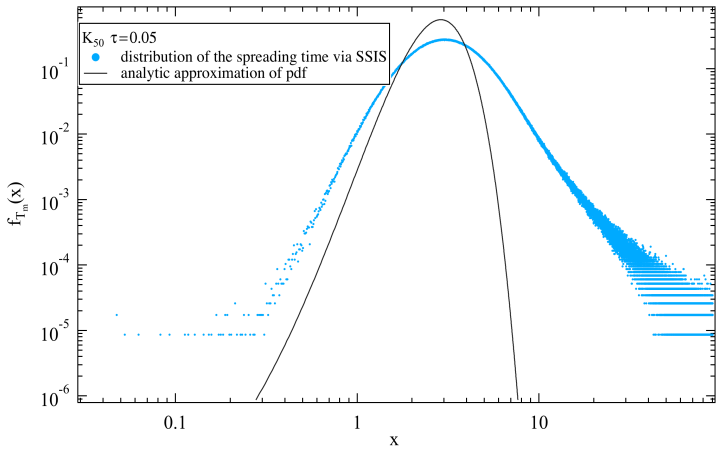

Figure 19: The distribution of the spreading time T_m via SSIS and the analytic approximation f_{T_m} with $\tilde{\mu} = 30$, $\tilde{\sigma} = 4.6$. We set $\eta = 10$ so that the maximum of $f_{T_m}(t)$ occurs the same time as the simulation.