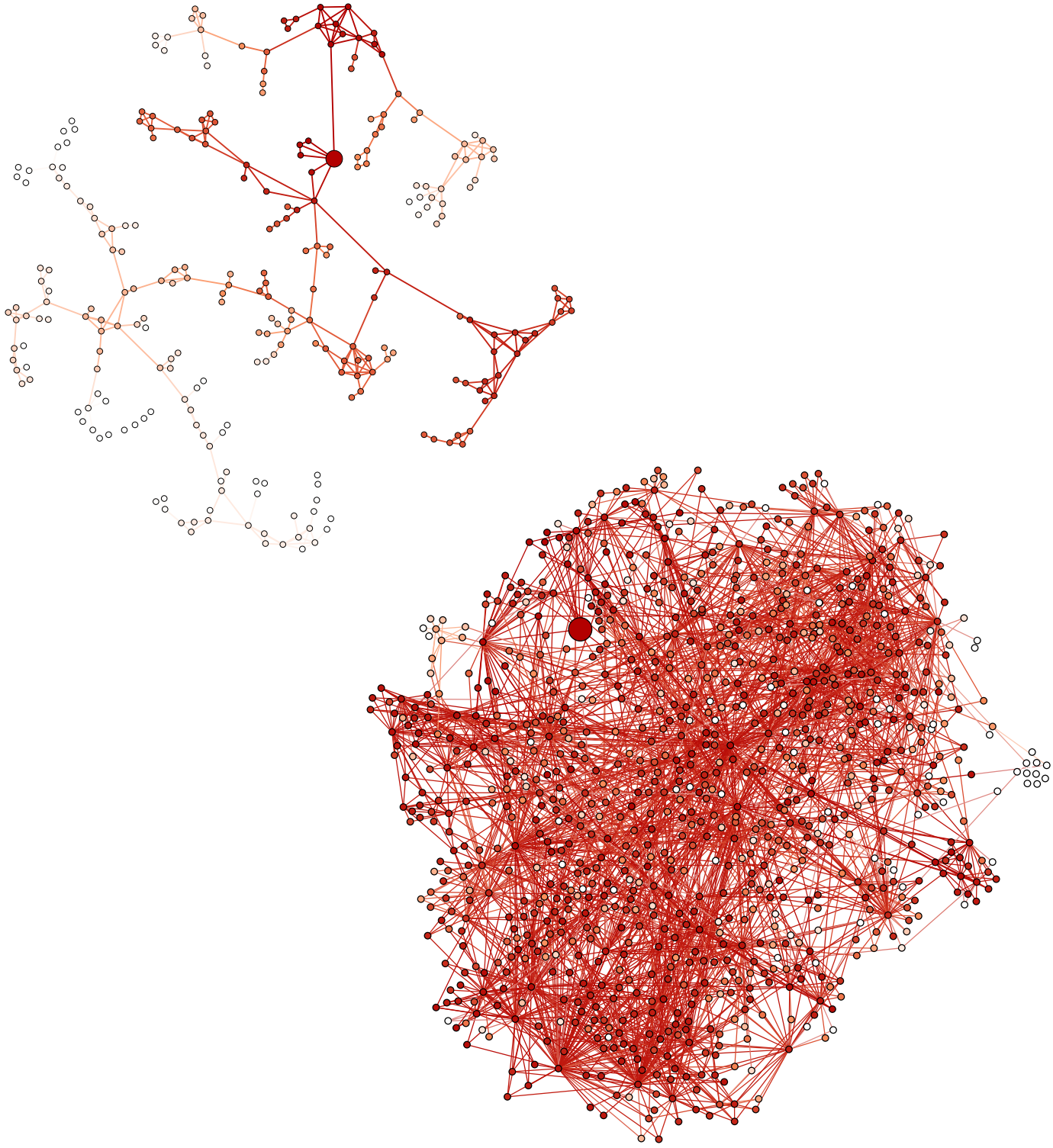
The role of the topology in non-Poissonian spreading dynamics on temporal networks



The role of the topology in non-Poissonian spreading dynamics on temporal networks

Alexey Medvedev

Supervisors: János Kertész

Center of Network Science, CEU; Technical University of Budapest (BME)

Gábor Pete

Alfréd Rényi Institute of Mathematics; Technical University of Budapest (BME)

Dissertation Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in Mathematics

Central European University

Budapest, Hungary

March, 2016

Alexey Medvedev: *The role of the topology in non-Poissonian spreading dynamics on temporal networks.*

Central European University, Budapest, Hungary

"The richest people in the world look for and build networks;
everyone else just looks for work."
Robert Kiyosaki

Acknowledgements

First and foremost I would like to express my sincere gratitude to my supervisors János Kertész and Gábor Pete for their support, patience and guidance through mysteries of the world of probability and complex networks. Their immense knowledge helped me in all the time of research and writing of this thesis. I could not have imagined having better advisers and mentors for my Ph.D study.

Besides my advisers, I would like to thank Júlia Komjáthy for her comments and suggestion to consider the model of a critical Erdős-Rényi graph, and Kimmo Kaski and Complex Networks research group in Aalto University, Finland, for providing the data and the kindest support during my research visit. This visit played an important role in writing the third chapter of this thesis.

I would like to heartfully thank the faculty of Central European University and Alfréd Rényi Institute of Mathematics for their support and mathematical lessons they taught me that I will never forget in my life.

My sincere thanks goes to my mom and my brother for their love, enormous support and encouragement during this four years of PhD studies. I would like to thank all of my friends around the globe for being there for me and making this trip to Europe an unforgettable life experience.

Last but not the least, I would like to say my warmest thanks to my beloved Yulia Chernova for her unconditional love, support and understanding, who provided me with an inspiration to discover and fight till 'the problems get solved'.

Abstract

Spreading is one of the most important dynamic processes on complex networks as it is the basis of a broad range of phenomena from epidemic contagion to diffusion of innovations. The speed and the behaviour of spreading models depend on number of factors. One of such factors, the topology of the underlying network, has been shown to influence the spreading process. In the current thesis we investigate both mathematically and numerically, how the changes in topological structure of the networks, such as appearance of new cycles by introducing extra edges or paths, which are so-called 'bridges', influences the model of non-Poissonian SI spreading on static, random and temporal networks.

The structure of the thesis is the following. In Chapter 1 we present the historical overview and main developments in the subject. In Chapter 2 we study the SI spreading with transmission times following power-law distribution with infinite mean. We derive the results both for general graphs and for specific models of random graphs. We study how the introduction of even one extra edge to a tree graph severely accelerates SI spreading on it. In Chapter 3 we study how the introduction of bridges influences the speed of the SI spreading on the real dataset of mobile phone calls. We derive that the introduction of topological bridges boosts the spreading and define a model which replicates this phenomena and compare results of this model on various theoretical network. We finish the thesis with the Discussion.

Contents

Contents

CHAPTER 1

Introduction

1.1 Networks as complex systems

Complex systems consist of interacting units on which processes take place. The constituents are represented by vertices and the interactions by edges of a graph. With reference to complex systems often the following names are used: graph → network, vertex → node and edge → link. Complex systems are abundant in nature and society and so are complex networks. Nodes in the networks represent people, vehicles, computers, braincells and links represent their connections or interactions. We can find examples of networks everywhere, starting from the macroscopic the networks of social interactions, computer networks and transportation networks around the globe, up to microscopic networks of protein-protein interactions in the living species or networks of neurons in the brain. These networks are dynamic in their nature and spreading is one of the examples of the dynamic processes that run in these networks. It can be the spreading of a computer virus in the computer networks [57], disease over a flight network [16] or innovation in the online social network [35].

Traditionally the study of networks has been the territory of graph theory. While initially small or regular graphs were in the focus, since the late 1950s large scale networks with no deterministic design principles have been described as random graphs. Large random graphs were first studied in detail by the Hungarian mathematicians Paul Erdős and Alfréd Rényi [23]. According to their model, one starts with N nodes and connects every pair of nodes with probability p, creating a graph with approximately $pN(N-1)/2$ edges distributed randomly. While the approximate character of this model was clear from the beginning, it dominated the thinking about complex networks for decades. The drastically increasing amount of data on complex systems had the effect that interest has shifted to more realistic

models. Our intuition clearly indicates that complex systems must display some organizing principles, which should be at some level encoded in their topology. As the topology of these networks indeed deviates from that of a perfectly random graph, tools and measures had to be developed to capture the underlying organizing principles in quantitative terms [4, 51]. Simple models were invented, like the Barabási-Albert or the Watts-Strogatz models, in order to capture important quantitative features of empirical networks [51]. Even those features that are vaguely defined, such as community structure, can be replicated in the network models [44].

Networks are not static in several sense. They usually result from dynamic, non-stationary processes, like growth in the number of nodes. Even in the stationary case, when the number of nodes remains constant, rewiring, death and birth processes may take place. Moreover, the function of a network is usually also dynamic. In all of the above mentioned examples the networks represent only the scaffold of the systems, where dynamic processes take place. These processes are random because of the topology and because of intrinsic stochasticity in their dynamics. Such processes include traffic (both internet and vehicular), chemical reactions, communication, and all kinds of spreading phenomena.

1.2 The development of spreading models

Spreading is one of the most important dynamic processes on complex networks [56,68] as it is the basis of a broad range of phenomena from epidemic contagion to diffusion of innovations. One of the original, and still primary, reasons for studying networks is to understand the mechanisms by which diseases, information, computer viruses, rumors, innovations spread over them [51].

The spreading problem initially came from epidemiology and had no relation to networks at all. Therefore we at the beginning use the language of 'people' for the vertices and 'diseases' that pass between them. Spreading processes can be described by different states a person can stay in. The simplest one is a two-state *SI model*, when a particular person is either in susceptible (S) or in infected (I) state, meaning that once a person caught a disease, there is no cure of it. More complicated case is three state *SIR model*, when we add a recovered (R) state, meaning that a person recovered from a disease is not any more susceptible. Among another well-studied models one can mention *SIS model*, when a recovered person can become susceptible again and the *SIRS model*, when a recovered person becomes susceptible only after

some time period.

The spreading problem initially came from epidemiology and had no relation to networks at all. Therefore, we at the beginning use the language of 'people' for the nodes in the network and 'infections' that pass between them. Spreading processes can be described by different states a person can stay in. The simplest one is a two-state *SI model*, when a particular person is either in susceptible (S) or in infected (I) state, meaning that once a person caught an infection, there is no cure of it. More complicated case is a three state *SIR model*, when we add a recovered (R) state, meaning that a person can be recovered from an infection is not any more susceptible. Among another well-studied models one can mention *SIS model*, when a recovered person can become susceptible again and the *SIRS model*, when a recovered person becomes susceptible only after some time period.

The simplest approach to the modelling of the spreading process makes no reference to the topology of peoples' interactions, considering a population, where every person could in principle have contact with anyone else with equal chance. This assumption is called the *fully mixed approximation*. In order to give a brief introduction to the mathematical models, consider the SI model under this approximation. Let $s(t)$ be a fraction of individuals who are susceptible at time t and let $x(t)$ be a fraction of people who are infected. Suppose β is the rate the infection spreads per unit time. Then this process can be described by system of differential equations:

$$\frac{ds}{dt} = -\beta sx,$$
$$\frac{dx}{dt} = \beta sx.$$

which gives an exact solution

$$x(t) = \frac{x_0 e^{\beta t}}{1 - x_0 + x_0 e^{\beta t}}.$$

In reality the structure of networks is far from the fully mixed, or complete graph assumption. Fortunately, the above models can be redefined in consideration of network structure. Consider again for simplicity the SI model on a connected network of n vertices. Let $s_i(t)$ be the probability for a vertex i to be susceptible and let $x_i(t)$ be the probability for a vertex i to be infected. Denote $\mathbf{A} = (A_{ij})_{i,j=1..n}$ the adjacency matrix and β the transmission rate. Then the describing system of differential equations is

$$\frac{ds_i}{dt} = -\beta s_i \sum_j A_{ij} x_j,$$
$$\frac{dx_i}{dt} = \beta s_i \sum_j A_{ij} x_j.$$

Despite the fact this model is the simplest one, it cannot be solved for arbitrary graphs in an exact way [50]. Therefore when studying the spreading processes one should use various kind of approximations, develop mean-field theories or operate with infinite graphs.

Spreading is a stochastic process. The usual approach, e.g. for simulations, is that one takes the static network and assumes that transmission of a disease is governed by a Poisson process. That means we take a network on N nodes and assume that transmission happens with at random times that are exponentially distributed. This approach heavily relies on the memoryless property of exponential distribution and thus the process is itself Markovian [66].

However, it has been shown empirically that spreading models with this Poissonian approach again produce results far from reality [67]. Complex systems are characterized by all kinds of inhomogeneities. The network itself is very inhomogeneous: the degrees, the activity of the constituents, the inter-event times have all broad distributions, which, together with various types of correlations do much impact on spreading in the network. The topological inhomogeneities are captured by the complex network models. [30, 55].

1.3 Studies of spreading on temporal networks

The underlying complex network is that of human interactions, however, this "Social Connectome" cannot be characterized by a static graph. The links may be active for some, often only short period and then inactive for the rest of the time [8]. An adequate framework to describe this situation is that of temporal networks [30].

A temporal network can be represented by a set of N nodes between which a complete trace of all interaction events $\mathcal{E}$ occurring within the time interval $[0, T]$ is known. Each such event can be represented by a quadruplet $e = (u, v, t, \delta t)$, where the event connecting nodes u and v begins at t and the interaction lasts until $t + \delta t$.

Recently large datasets on communication have been made available, where not only the participants are recorded but also the time stamps and durations of the communications, sometimes together with some additional information called metadata (like gender or age). The standard approach to analyze these data is to aggregate them over time and construct this way a weighted static network on which at most inhomogeneous Poisson processes are assumed. However, the empirical data shows that human interactions are bursty and dynamic of spreading differs from earlier expectations. The failure of Poissonian approximation was first shown in [67], where the empirical study of email activity patterns has been provided.

Recent empirical studies stimulate us to argue that inter-event communication times can be described more precisely with heavy-tailed or power law distributions [30].

The strategy of studying the effects of different kinds of inhomogeneities on spreading on temporal networks is to compare results obtained from empirical case with those reference networks [38]. The latters are obtained by randomizing the initial temporal one, preserving some properties of the original network. For example, network obtained by randomizing the times of each contact is such an example. This randomization preserves the number of contacts on each link and the structure of the aggregated network, whereas it destroys the temporal structure of the contact sequence on each link.

Thus, the main question is: how does the temporal structure affect the spreading on temporal networks [47]. The SI model was numerically simulated on the temporal network of phone calls between people, consisting of 4.5 million nodes and 9 million links. The authors showed that epidemic spreading slows down on temporal network compared to different randomized ones [38]. Another simulation for SIS model on artificial temporal network data led to similar results [49]. On the other hand, in a variant of the SI model, in which multiple infection attempts by infected individuals within a short time is necessary for a susceptible individual to be infected, epidemic spreading is more facilitated by empirical temporal structure compared to randomized reference data. The SIR model has been simulated on a temporal network of 20 million nodes with fixed recovery time [49]. The authors conclude, regarding to the bursty nature of contacts, that global outbreak in this model is suppressed on this type of temporal structure.

Empirical studies have shown that, in addition to the non-trivial distribution of events, there are dependencies between them [36,37]. It has been shown for SIR model that temporal correlation of events occurring on links tend to enhance the spreading process [49]. Correlations may occur in one time series on a link or there can be dependencies between events of different links resulting, e.g., in overrepresented patterns called temporal motifs [43]. The importance of correlations also has been noted in modeling of epidemic spreading on network of sexual contacts in Internet-mediated prostitution [30]. Recently such dependencies have become particularly interesting and simple queuing type models could shed light on how they influence the characteristic quantities of the process even on a single link [63]. Their effect on temporal networks remains to be shown.

1.4 Mathematical studies of spreading

It has been already mentioned that networks can be modeled as finite random graphs. The geometric structure of sparse random finite graphs (Erdős-Rényi graphs, d-regular random graphs, configuration models, preferential attachment models) is most often understood via passing to an infinite limiting random graph, e.g., a branching process tree. Taking this limit is done either via the Benjamini-Schramm local weak limit [6, 11, 12], or by exploration processes [17, 58, 64], where generating function methods and martingale techniques can be successfully applied.

On the other hand, the spreading phenomena can be modeled as a stochastic process. The limits of stochastic processes on finite graphs can be taken: the classical examples are thermodynamic limits from boxes of the $\mathbb{Z}^d$ lattice. Understanding a stochastic process is often simpler on the infinite graph, and the behavior of the finite graphs may be deduced from that. However, it is usually not clear without a careful study whether the behavior of a given stochastic process is determined by the local structure of the graph (such as degree distribution, local clustering effects, etc), or also influenced by some global structure that is lost in the above limiting procedures (such as bipartiteness). For instance, the maximum density of an independent subset of vertices (i.e., pairwise non-neighbors) is not local, while the maximum density of an independent subset of edges (i.e., a matching), is local.

An interesting example is bootstrap percolation, which is a dynamic percolation model where occupation probability depends on the number of already occupied neighbors. This is a spreading model similar to SI, more relevant to the spreading of innovation than to epidemics. On d-regular random graphs, despite the non-trivial geometry, a differential equation technique (analogous to the one we sketched above) was applied successfully in [7] to find that the initial critical density for complete occupation is basically the same as on d-regular trees [18].

First passage percolation is a mathematical analogue of the notion of SI spreading introduced on networks. It has been shown that the behavior of first passage percolation on sparse random graphs is universal across a lot of models [13]. The model of first passage percolation has also been studied on different types of static regular structures, such as hypercube and complete graph [22, 25]. In these models one is mainly interested in obtaining a limit law for the length of shortest weighted paths between two random vertices, which can be translated into spreading from one source. The more sophisticated approach is by using general type

branching processes [14]. This approach allows to study the spreading processes in full generality, such that there is no assumption on the transmission times of the disease. There the locally tree-like property of the network plays a crucial role in the result. In general most of the general mathematical methods of studying the spreading processes fail when the topology includes cycles.

In a number of works the influence of the topological properties on the spreading models is studied [20, 27, 65]. In these works authors are mainly providing the evidence to extend the usual phenomenon that "conductance determines mixing time of random walk" to SI, SIS, etc. models, and from Markovian to non-Markovian dynamics. The topological measure used by the authors is conductance, or equivalently, spectral gap. In our work adding a single edge does not significantly change conductance, but we show that the speed is significantly changed.

In the current thesis we investigate both mathematically and numerically, how the appearance of new cycles by introducing extra edges or paths, which are so-called 'bridges', influences the model of SI spreading on various networks. The structure of the thesis is the following. In Chapter 2 we study the SI spreading with transmission times following power-law distribution with infinite mean. We derive the results both for general graphs and for specific models of random graphs. We study how the introduction of even one extra edge to a tree graph severely accelerates SI spreading on it. In Chapter 3 we study how the introduction of bridges influences the speed of the SI spreading on the real dataset of mobile phone calls. We derive that the introduction of topological bridges boosts the spreading and define a model which replicates this phenomena and compare results of this model on various theoretical network.

CHAPTER 2

The role of extra edges in smoothing of the spreading curve

2.1 Motivation

In the current Chapter we analyse the behaviour of the SI spreading on tree-like networks with transmittion times having power-law distribution with infinite mean. The main motivation of this research is a question posed in the work of Kertesz and Horvath in [31]. Authors have found the presence of "jumps" in the SI spreading curves of computer simulated model of SI spreading. The *spreading curve* $N(t)$ is one of the main characteristics that identifies the SI spreading. The function $N(t)$ records pairs $(k/n, T_k)$, where $1 \leqslant k \leqslant n$ and T_k is the time to infect k vertices in the graph. In computer simulations M spreading curves are collected and averaged over M, thus obtaining an *average spreading curve* $\langle N(t) \rangle$, which is a set of pairs

$$\left(k/n, \frac{\sum_{i=1}^{M} T_k^{(i)}}{M} \right),$$

where $T_k^{(i)}$ denotes the time to infect k vertices in the i'th spreading curve. The authors have empirically studied the SI spreading process on several classes of trees and found that average spreading curves have "jumps" which are not smoothed out by increasing number of runs M. These jumps are explained by the presence of temporal "bottlenecks", which are large weights that occur on some edges. It is clear that if the theoretical expectation $\mathbb{E}(T_k)$ is infinite, then by the law of large numbers, for any function $f(k)$ there exists m such that for all $M > m$ we have

$$\frac{\sum_{i=1}^{M} T_k^{(i)}}{M} > f(k).$$

Therefore the jumps are expected for such k with $\mathbb{E}(T_k) = \infty$.

However, performing the same simulation on a cycle C_n with $n = 1000$ vertices with power-law weights $\alpha = 0.8$ we find no jumps after averaging over $M = 1000$ runs (see Figure 2.1). In Section 2.3.1 we find that in the case of a cycle $\mathbb{E}(T_k) \asymp k^{1/\alpha}$, where $\alpha \in (1/2, 1)$ and $1 \leqslant k \leqslant n$.

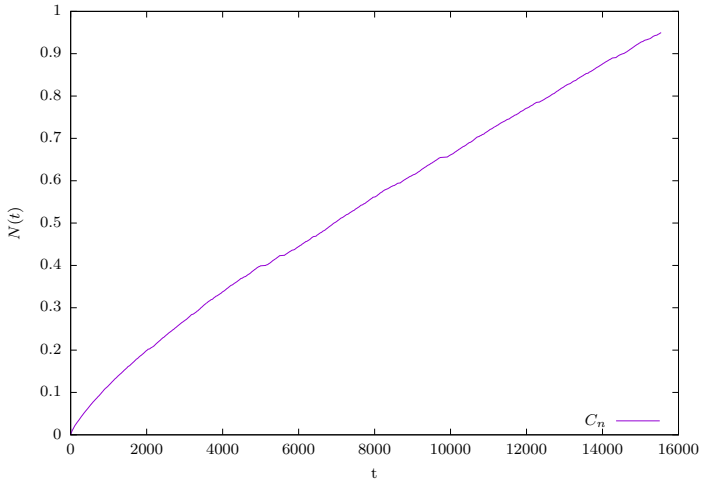


Figure 2.1. *Simulation of SI spreading with power-law weights with* $\alpha = 0.8$ *on the cycle* C_n *with* $n = 1000$ *vertices.*

At each time the cycle has two active edges that transmit the infection. In this case, for $\alpha \leqslant 1/2$ it is easy to see that the average time to infect the first vertex is a minimum of two i.i.d. weights and thus $\mathbb{E}(T_1) = \infty$, therefore $\alpha = 1/2$ is a threshold. On the contrary, when there are n active edges initially transmitting the infection, then for $\alpha > 1/2$ the average time to infect $n - 1$ vertices is also of order $n^{1/\alpha}$, which we establish in the Section 2.3.2. Motivated by these two "extreme" constructions we derive that for any graph G there exist a number of vertices $\kappa(G)$, such that there are no jumps before this $\kappa(G)$. The number $\kappa(G)$ identifies the place where the first bottleneck may appear. The following result is proven in Section 2.4.

Theorem 2.1.1. *Consider the graph* G *with root* s *on* n *vertices and the SI spreading process* $T = (T_j)_{j=1}^n$ *with power-law weights with* $\alpha \in (1/2, 1)$. *Then there exist* $\kappa(G)$ *such that for each* k, *where* $1 \leqslant k \leqslant \kappa(G)$, *the expected time to infect* k *vertices is bounded by*

$$\mathbb{E}(T_k) \leqslant Ck^{1/\alpha},$$

and for $k \geqslant \kappa(G)$, *the* $\mathbb{E}(T_k) = \infty$.

We have performed the simulation of the considered SI process with power-law weights with $\alpha = 0.8$ on the critical Galton-Watson tree with $N = 1000$ vertices and the average

spreading curve over $M = 1000$ runs and found the same jumps in the spreading curve (see Figure 2.2). However if we perform the same simulation with an extra edge attached to a root and a random vertex in the critical Galton-Watson tree, we observe there are no jumps up to fraction of the total size of the tree. This result is explained with mathematical rigour in Section 2.2.5.

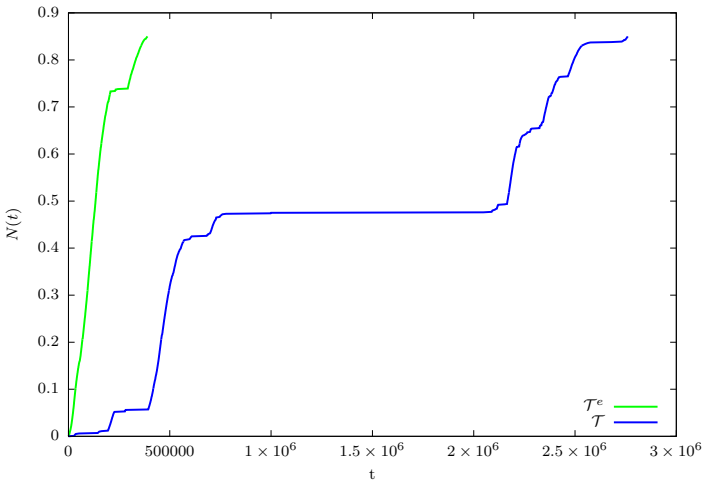


Figure 2.2. *Simulation of SI spreading with power-law weights with* $\alpha = 0.8$ *on the critical Galton-Watson tree* $\mathcal{T}$ *with* $n = 1000$ *vertices and the same tree with extra edge attached to the root and a uniform vertex denoted as* $\mathcal{T}^e$.

Theorem 2.1.2. *Consider the critical GW tree* $\mathcal{T}$ *conditioned on* $Z_N > 0$ *and the SI spreading process* $T = (T_j)_{j=1}^{|\mathcal{T}|}$ *with power-law weights with* $\alpha \in (1/2, 1)$, *where* $|\mathcal{T}|$ *denotes the size of* $\mathcal{T}$. *Then there exist* $\kappa(\mathcal{T}) < |\mathcal{T}|$ *with tight distribution such that for any* $k \geqslant \kappa(\mathcal{T})$,

$$\mathbb{E}(T_k) = \infty.$$

Denote as $\mathcal{T}_{+e}$ *the tree* $\mathcal{T}$ *with an extra edge attached to a root and one of the vertices of* $\mathcal{T}$ *uniformly at random. Then as* $N \to \infty$ *for any* $\varepsilon > 0$ *there exists* $\delta > 0$, *such that*

$$\mathbb{P}\Big(\frac{\kappa(\mathcal{T}_{+e})}{|\mathcal{T}^e|} > \delta\Big) > 1 - \varepsilon.$$

The first statement of the Theorem 2.1.2 tells us that the jumps start to appear after some short time in the spreading curve of the critical GW tree, but if we introduce one more edge, with high probability we infect a positive fraction of the vertices. This phenomena we call *smoothing of the spreading curve*, when a small topological change decreases jumps in large part of the spreading curve.

2.2 Basic definitions and technical lemmas

2.2.1 Graph-theoretic notions

Denote a simple graph $G = (V, E)$ with n vertices and m edges. We call a root of the graph G a distinguished vertex from V. We denote the set of vertices of graph G as V or $V(G)$ and the set of edges as E or $E(G)$. Let $s, t \in V$, then define a *(simple) path* (s, t) between vertices s and t of length l as the sequence of vertices $(v_0, v_1, \ldots, v_l)$, such that $v_0 = s$, $v_l = t$ and each $v_i \neq v_j$, for $0 \leqslant i \neq j \leqslant l$. We denote a path between s and t as (s, t)-*path*. Define the *graph distance* $d(s, t)$ between vertices s and t as the length of the shortest (s, t)-path.

Suppose there is a function $\xi : E \to \mathbb{R}^+$, then define the *weight* of an edge $e \in E$ as the value of this function $\xi(e)$. The function ξ is called a *weight function* or, simply, *weights*. Then $G_\xi = (V, E, \xi)$ denotes a *weighted simple graph* G with weights ξ.

The graph $H = (V', E')$ obtained by selecting the vertex subset $V' \subset V$ with the same root s along with the edges $E' \subset E$ connecting them, is called an induced subgraph of G and denoted as $H \sqsubset G$. We call graph G a cycle if it is a path $(v_0, \ldots, v_l)$, such that $v_0 = v_l$. We call the graph T a *tree* if it is a connected graph without cycles. It is easy to see that if $T = (V, E)$ is a tree on n vertices, then $|E| = n - 1$. We call a root of the graph G a distinguished vertex $s \in V$. Denote $|G|$ to be the total size of the vertex set, or $|G| = |V|$.

2.2.2 Functional notions

Let f and g be positive real-valued functions. Denote $f \sim g$ as $x \to \infty$, if $f(x)/g(x) \to 1$; $f = o(g)$ as $x \to \infty$, if $f(x)/g(x) \to 0$, and $f = O(g)$ as $x \to \infty$, if $|f(x)| \leqslant M|g(x)|$ for all $x \geqslant x_0$ for some x_0. We say $f(x) \asymp g(x)$ for all x, if there exist independent non-zero constants c and C, such that $cg(x) \leqslant f(x) \leqslant Cg(x)$ for all x.

In our computations we use the following approximation of an exponential function.

Lemma 2.2.1. *Let $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ be real sequences, such that $a_n \to 0$, $b_n \to \infty$ and $a_n b_n \to 0$ when $n \to \infty$. Then,*

$$(1 - a_n)^{b_n} \sim \exp(-a_n b_n).$$

Since the product $a_n b_n \to 0$ we may write that for all $n > n_0$ by the Taylor expansion of exponential function we have the following:

$$(1 - a_n)^{b_n} \sim \exp(-a_n b_n) \sim 1 - a_n b_n + (a_n b_n)^2/2. \tag{2.1}$$

Let $\gamma(s,x)$ denote the lower incomplete Gamma function,

$$\gamma(s,x) = \int_0^x t^{s-1}e^{-t}dt,$$

and $\Gamma(s,x)$ the upper incomplete Gamma function,

$$\Gamma(s,x) = \int_x^\infty t^{s-1}e^{-t}dt.$$

The following Lemmas show the asymptotic behaviour of $\gamma(s,x)$ and $\Gamma(s,x)$ [53].

Lemma 2.2.2. *Let $\gamma(s,x)$ be a lower incomplete Gamma function. Then as $x \to 0$,*

$$\frac{\gamma(s,x)}{x^s} \to \frac{1}{s}.$$

Lemma 2.2.3. *Let $\Gamma(s,x)$ be an upper incomplete Gamma function. Then as $x \to \infty$,*

$$\frac{\Gamma(s,x)}{x^{s-1}e^{-x}} \to 1.$$

We also use a well-known Lemma from analysis.

Lemma 2.2.4. *Let $f(x)$ be a continuous monotonic concave function. Then,*

$$\sum_{k=1}^{n} f(k) \leqslant \int_1^{n+1} f(x)dx.$$

Denote as $(b_n)_{n=1}^\infty$ the positive sequence that satisfies the following recurrence relation for some $C > 0$ and $1/2 < \alpha < 1$:

$$b_{n+1} \leqslant b_n + Cb_n^{1-\alpha}, \tag{2.2}$$

with the initial condition

$$b_1 = d = (\alpha C)^{1/\alpha}. \tag{2.3}$$

The following Lemma presents an upper bound on the sequence $(b_n)_{n=1}^\infty$.

Lemma 2.2.5. *Suppose $(b_n)_{n=1}^\infty$ defined as above in (2.2), (2.3). Then,*

$$b_n \leqslant dn^{1/\alpha},$$

where $d = (\alpha C)^{1/\alpha}$.

Proof. We prove the statement by induction. By definition the statement holds for $b_1 := d$. Suppose the statement holds for $n \geqslant 1$ and for any k, where $1 \leqslant k \leqslant n$, we have $b_k \leqslant dk^{1/\alpha}$. Then prove the statement for $n+1$. We can rewrite (2.2) as

$$b_{n+1} - b_n \leqslant C b_n^{1-\alpha}.$$

Making a telescopic sum, we have

$$b_{n+1} - b_1 \leqslant \sum_{k=1}^{n} C b_k^{1-\alpha}.$$

Then by induction hypothesis,

$$b_{n+1} - b_1 \leqslant \sum_{k=1}^{n} C d^{1-\alpha} k^{1/\alpha - 1},$$

and by Lemma 2.2.4 we may bound the sum with an integral and obtain:

$$\begin{aligned} b_{n+1} - b_1 &\leqslant \int_1^{n+1} C d^{1-\alpha} x^{1/\alpha - 1} dx = \alpha C d^{1-\alpha} \left((n+1)^{1/\alpha} - 1 \right) \\ &= d \left((n+1)^{1/\alpha} - 1 \right). \end{aligned} \tag{2.4}$$

Since b_1 is equal to d, then we can add it to both parts of the Equation (2.4) and have

$$b_{n+1} \leqslant d(n+1)^{1/\alpha}.$$

This finishes proof of the Lemma. □

2.2.3 Probabilistic notions

Denote random variables with small greek letters $\xi, \eta, \ldots$ and denote the pdf of an absolutely continuous random variable ξ as $p(t)$ and cdf as $F(t) := \mathbb{P}(\xi < t)$. Let ξ be a random variable, then we use the letter μ to denote the expected value $\mathbb{E}(\xi)$ and σ^2 for variance $Var(\xi)$. We call a random variable ξ to have a power-law distribution $pow(t_{min}, \alpha)$, where $t_{min}, \alpha > 0$, if it has the following density function:

$$p(t) = \begin{cases} t_{min}^{\alpha} \alpha \frac{1}{t^{\alpha+1}}, & \text{if } t_{min} \leqslant t; \\ 0, & \text{otherwise.} \end{cases} \tag{2.5}$$

Further in this chapter we shall refer to a power-law distributed random variable as the one with the parameter $t_{min} = 1$, or $\mathbb{P}(\xi > t) = t^{-\alpha}$, when $t > 1$.

Two random variables X and Y are coupled when they are defined on the same probability space and they have the correct marginal distributions. Formally speaking, random variables

$(\hat{X}_1, \hat{X}_2)$ are a *coupling* of the random variables (X_1, X_2) when $(\hat{X}_1, \hat{X}_2)$ are defined on the same probability space, and are such that the marginal distribution of $\hat{X}_i$ is the same as the distribution of X_i for $i = 1, 2$, i.e., for all measurable subsets E of $\mathbb{R}$,

$$\mathbb{P}(\hat{X}_i \in E) = \mathbb{P}(X_i \in E).$$

Let X and Y be two discrete random variables with

$$\mathbb{P}(X = x) = p_x, \quad \mathbb{P}(Y = y) = q_y, \quad x \in \mathcal{X}, y \in \mathcal{Y},$$

where $(p_x)_{x\in\mathcal{X}}$ and $(q_y)_{y\in\mathcal{Y}}$ are any two probability mass functions on two subsets $\mathcal{X}$ and $\mathcal{Y}$ of the same space. Define the *total variation distance* $d_{TV}(p, q)$ between measures p and q in following way:

$$d_{TV}(p, q) = \frac{1}{2} \sum_x |p_x - q_x|.$$

The main result linking the total variation distance of two discrete random variables and a coupling of them is the following Theorem, named after Strassen (see [61] for the original version; we use the formulation as in [64], p.59).

Theorem 2.2.6. *For any two discrete random variables X and Y with measures p and q, there exists a coupling $(\hat{X}, \hat{Y})$ of X and Y, such that*

$$\mathbb{P}(\hat{X} \neq \hat{Y}) = d_{TV}(p, q),$$

while for any coupling $(\hat{X}, \hat{Y})$ of X and Y,

$$\mathbb{P}(\hat{X} \neq \hat{Y}) \geqslant d_{TV}(p, q).$$

Let X and Y be two random variables, not necessarily defined on the same probability space. The random variable Y *stochastically dominates* the random variable X, which is denoted as $X \preceq Y$, if for every $x \in \mathbb{R}$, the following inequality holds:

$$\mathbb{P}(X \geqslant x) \leqslant \mathbb{P}(Y \geqslant x).$$

By definition the stochastic domination $X \preceq Y$ implies $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$. We also make use of the following Theorem [46].

Theorem 2.2.7. *The real random variable X stochastically dominates Y if and only if there is a coupling $(\hat{X}, \hat{Y})$ of X and Y, such that*

$$\mathbb{P}(\hat{X} \geqslant \hat{Y}) = 1.$$

2.2.4 Levy stable laws

Let X be a power-law distributed random variable with $0 < \alpha < 1$. Then it is easy to see that $\mathbb{E}(X) = \infty$. The function L is called a *slowly varying* function, if

$$\lim_{x\to\infty} L(tx)/L(x) = 1 \quad \text{for all } t > 0.$$

Consider the sequence of power-law distributed random variables $X_1, X_2, \dots$. Then the following limit Theorem holds for their sum (see [21], p.138).

Theorem 2.2.8. *Suppose $X_1, X_2, \dots$ are i.i.d. with a distribution that satisfies*

(i) $\lim_{x\to\infty} \mathbb{P}(X_1 > x)/\mathbb{P}(|X_1| > x) = \theta \in [0, 1]$;

(ii) $\mathbb{P}(|X_1| > x) = x^{-\alpha}L(x)$,

where $\alpha < 2$ and L is slowly varying. Let $S_n = X_1 + \cdots + X_n$, and

$$a_n = \inf\{x : \mathbb{P}(|X_1| > x) \leqslant n^{-1}\} \text{ and } b_n = n\mathbb{E}(X_1\mathbf{1}_{(|X_1|>x)}).$$

Then as $n \to \infty$, $(S_n - b_n)/a_n \xrightarrow{d} Y$, where Y has a nondegenerate distribution.

It is known that if $\mathbb{P}(|X_1| > t^{-\alpha})$ with $\alpha < 1$, then the normalizing constants are the following (see [21], p.142):

$$a_n = n^{1/\alpha} \quad \text{and} \quad b_n = 0.$$

Along with the limit Theorem 2.2.8 the local limit analogue holds [60].

Theorem 2.2.9. *Let $X_1, X_2, \dots$ be the sequence of i.i.d. random variables such that $\mathbb{P}\big(X_1 > t\big) = t^{-\alpha}$ for $t > 1$, $0 < \alpha < 1$. Then for any $x \in supp(Y)$ and $h > 0$*

$$\mathbb{P}\Big(\frac{S_k}{k^{1/\alpha}} \in (x, x+h)\Big) = \mathbb{P}\big(Y \in (x, x+h)\big) + o(1)\left(h + \frac{1}{k^{1/\alpha}}\right). \tag{2.6}$$

The limit distribution Y is called a *stable law* with same parameter α as X_1 and is given via a characteristic function ([21], p.141)

$$\phi_Y(t) = \exp(itc - b|t|^\alpha(1 + i\kappa \mathrm{sgn}(t)\tan(\pi\alpha/2))),$$

where $\kappa = 2\theta - 1$. The random variable Y is supported on $(0, \infty)$ and by the following Lemma has continuous density (see [24], p.657).

Lemma 2.2.10. *The stable law Y is absolutely continuous on $(0, \infty)$, therefore has density $f_Y(t)$.*

The following Lemmas describe the approximations of the starting and tail behaviour of the distribution Y for any $0 < \alpha < 1$ [24].

Lemma 2.2.11. *Let Y be the stable law with parameter $0 < \alpha < 1$. Then as $t \to \infty$ the probability density function of Y is*

$$f_Y(t) \asymp t^{-\alpha-1}.$$

Lemma 2.2.12. *Let $X_1, X_2, \dots$ be a sequence of i.i.d. random variables with $\mathbb{P}\big(X_1 > t\big) = t^{-\alpha}$ and denote $S_k = \sum_{i=1}^{k} X_i$ and $\overline{S_k} = S_k/k^{1/\alpha}$. Then for any small $c > 0$*

$$\mathbb{P}\big(\overline{S}_k < c\big) \asymp e^{-c^{-\alpha}}. \tag{2.7}$$

Proof. We have $\max\{X_1, X_2, \dots, X_k\} < S_k$ for every $k > 0$, then for any $\varepsilon > 0$

$$\begin{aligned} \mathbb{P}\Big(\overline{S_k} < c\Big) < \mathbb{P}\Big(\frac{\max\{X_1, X_2, \dots, X_k\}}{k^{1/\alpha}} < c\Big) = \\ = \left(1 - \frac{c^{-\alpha}}{k}\right)^k. \end{aligned} \tag{2.8}$$

The value of c is small, thus by Lemma 2.2.1 we can write $\left(1 - \frac{c^{-\alpha}}{k}\right)^k \sim e^{-c^{-\alpha}}$, as $k \to \infty$, and since the interval $(0, c)$ is bounded, we can find uniform constants such that

$$\mathbb{P}\Big(\overline{S_k} < c\Big) \asymp e^{-c^{-\alpha}}.$$

□

The stable law Y with the exponent $\alpha = 1/2$ has an explicit form of the distribution with the following density (see [21], p.141, Eq.(3.7.12)):

$$p_Y(y) = \frac{1}{\sqrt{2\pi}} y^{-3/2} e^{-\frac{1}{2y}}. \tag{2.9}$$

2.2.5 Critical Galton-Watson trees

A *Galton-Watson process* is usually defined as a process $\langle Z_n : n \geqslant 0\rangle$ of evolution of a particle system. We start with one particle, $Z_0 = 1$. It splits into k offspring particles with probability p_k and they constitute generation Z_1. Then each of these offsprings (should there be any) also have children with the same offspring distribution $\langle p_k : k \geqslant 0\rangle$, independently of each other and of their parent. This continues forever or until there are no children born.

In order to give a formal definition, let $\xi_1^{(j)}, \xi_2^{(j)}, \dots$, where $j \in \mathbb{N}$ denote i.i.d. non-negative integer distributed random variables with distribution $\mathbb{P}(\xi = k) = p_k$. We define the

Galton-Watson process $\langle Z_n : n \geqslant 0\rangle$ as the Markov process on the non-negative integers with the following recursion:

$$Z_{n+1} := \sum_{i=1}^{Z_n} \xi_i^{(n+1)}.$$

The quantity Z_n, where $n \geqslant 0$, is called a n'th generation of the process and we assume $\langle p_k : k \geqslant 0\rangle$ is not degenerate to avoid trivial cases. The process is called *critical* if $\mathbb{E}(\xi) = 1$.

For each Galton-Watson process one can define a genealogical tree where vertices are associated with the particles at each generation and there is an edge between two particles if one is the parent of the other. We denote these trees as *Galton-Watson (GW) trees.* Denote the set of all GW trees as $\langle GW\rangle$ and a random GW tree as $\mathcal{T}$. Let the root of the tree $\mathcal{T}$ be the particle in generation Z_0. We define the size of $\mathcal{T}$ as the number of vertices it contains. It is well known that a critical Galton-Watson (CGW) tree is almost surely finite (e.g. Theorem 3.1, p.84, [64]) and the following Theorem holds for the size of a CGW tree [42].

Theorem 2.2.13. *Let $\mathcal{T}$ be a critical GW tree with offspring distribution ξ such that $Var(\xi) := \sigma^2 < \infty$. Then as $n \to \infty$,*

$$\mathbb{P}(|\mathcal{T}| = n) \sim \frac{1}{\sigma\sqrt{2\pi}} n^{-3/2}.$$

The *height* of a GW tree $\mathcal{T}$ is the length of the longest path from the root or the maximum N, such that $(Z_N > 0)$. The following limit Theorem holds about the height of the tree $\mathcal{T}$ [41].

Theorem 2.2.14. *Let $\mathcal{T}$ be a critical GW tree with offspring distribution ξ such that $Var(\xi) := \sigma^2 < \infty$. Then we have:*

$$\lim_{N\to\infty} N\mathbb{P}(Z_N > 0) = \frac{2}{\sigma^2}.$$

The following Theorem provides the estimation of the probability of having a tree of height at least N conditioned on the exact size of this tree [3] and [42].

Theorem 2.2.15. *Let $\mathcal{T}$ be a critical GW tree with offspring distribution ξ, such that $Var(\xi) := \sigma^2 < \infty$ and let $H(\mathcal{T})$ denote the height of the tree $\mathcal{T}$. Then there exist constants C and c, such that*

$$\mathbb{P}(H(\mathcal{T}) \geqslant x \big| |\mathcal{T}| = n) \leqslant Ce^{-cx^2/n}.$$

We consider the set of GW trees *conditioned on* $Z_N > 0$, where $N > 0$, as the subset of trees $\langle GW\rangle$ with height at least N. Denote this set of conditioned GW trees as $\langle GW \big| Z_N > 0\rangle$. The expected limit size of the k'th generation in such trees is given in the following Theorem [48].

Theorem 2.2.16. *Let $\mathcal{T}$ be a critical GW tree with offspring distribution ξ, such that $Var(\xi) := \sigma^2 < \infty$. Then we have:*

$$\lim_{N\to\infty} \mathbb{E}(Z_k | Z_N > 0) = 1 + k\sigma^2.$$

We also consider the set $\langle GW^\infty \rangle$ of *infinite* CGW trees which are the geneological trees of a critical Galton-Watson process conditioned on *non-extinction*. We present the following construction of an infinite CGW tree due to Kesten [40].

The infinite critical GW tree $\mathcal{T}^\infty$ has two types of vertices: *normal* and *special*, with root being special. Normal vertices have offsprings according to independent copies of ξ, while special nodes have offspring according to independent copies of size-biased distribution $\hat{\xi}$, such that

$$\mathbb{P}(\hat{\xi} = k) := kp_k,$$

where $k = 0, 1, 2, \ldots$. All offsprings of a normal vertex are normal; when a special vertex produces a number of offsprings, one of its children is selected uniformly at random and becomes special, while all other children are normal.

An alternative construction of the critical infinite tree $\mathcal{T}^\infty$ is to start with the *spine* γ (an infinite path of special vertices from the root) and then at each node in the spine attach independent critical GW trees. The number of branches at each special node in the spine is a copy of $\hat{\xi} - 1$. Since each CGW tree is a.s. finite, it follows that $\mathcal{T}^\infty$ a.s. has exactly one infinite path from the root, viz. the spine.

2.2.6 SI spreading process

We are going to define the SI spreading process $T = (T_k)_{k=1}^n$ as a stochastic process on the weighted graph $G = (V, E, \xi)$ on n vertices and a root s, where for each $e \in E$ the $\xi(e)$ are i.i.d. random variables having some non-trivial distribution ξ. By definition each vertex $v \in V$ may be in one of the following two states: *susceptible* (S) or *infected* (I). The edge $e \in E$ is called *available* if one of the end vertices is in the infected state I and the other is susceptible S; *occupied*, if both vertices are in the infected state I; and *unavailable* otherwise. An infection is transmitted along the available edges from infected vertices to susceptible. The weight $\xi(e)$ of an edge $e = (u, v)$ is a passage time of an infection and if the vertex u becomes infected at time t, then v by definition immediately turns into infected state at time $t + \xi(e)$.

Another view on the SI process is the following. Consider $\xi(e)$ as a length of an edge $e \in E$. Then we can think of the transmission of the infection as a flow from infected vertices to susceptible ones through edges of lengths ξ with constant rate 1.

The process starts at time $t = 0$ with all vertices being in the state S and the root s is turned into the infected state I and along with time t the infection is transmitted along the available edges at rate 1. An available edge that transmits the infection at time t is called an *active* edge. The process runs until all vertices turn into infected state.

We refer to this process as *an SI process on graph G with weights having distribution ξ.*

2.3 Examples

2.3.1 Spreading on a cycle

In the following Subsection we consider the example of SI spreading with power-law weights on the graph on a cycle C_n on n vertices. The spreading on a cycle can be well approximated by a graph of a doubly infinite line. The graph of doubly infinite line, denoted as $G = (V, E)$, is defined by the vertex set of integers $V = \{0, \pm 1, \pm 2, \dots\}$ with a root at 0 and for each $i, j \in G$ the pair $(i, j) \in E$ iff $|i - j| = 1$. We consider the SI spreading process $T = (T_k)_{k=1}^{\infty}$ with power-law distributed random weights with $\alpha \in (1/2, 1)$, denoted as X_i, where i is a label of the greater vertex, if one of the ends is positive, and is the label of the smaller otherwise.

Theorem 2.3.1. *Let G be the graph of a doubly infinite line with root 0. Then in the SI spreading process $(T_k)_{k=1}^{\infty}$ with power-law weights $\alpha \in (1/2, 1)$ the expected time to infect k vertices is bounded:*

$$\mathbb{E}(T_k) \asymp k^{1/\alpha}.$$

Proof. Let N_t denote the number of vertices infected by time t and denote as $S_k = \sum_{i=1}^{k} X_i$ and $S_k^* = \sum_{i=-1}^{-k} X_i$. Then the probability of infecting n vertices by time t is the following:

$$\mathbb{P}(N_t = n) = \sum_{k=0}^{n} \mathbb{P}\big(S_k < t < S_{k+1}\big)\, \mathbb{P}\big(S_{n-k}^* < t < S_{n-k+1}^*\big)\,. \tag{2.10}$$

Consider the probability $\mathbb{P}\big(S_k < t < S_{k+1}\big)$. By definition of X_i, $\mathbb{P}(X_i < 1) = 0$, therefore when $0 < t < k$:

$$\mathbb{P}\big(S_k < t < S_{k+1}\big) = 0.$$

Denote the $\overline{S}_k = \frac{S_k}{k^{1/\alpha}}$, $\overline{X}_{k+1} = \frac{X_{k+1}}{k^{1/\alpha}}$ and $c = \frac{t}{k^{1/\alpha}}$. We can rewrite for t, where $k < t < \infty$

$$\begin{aligned}\mathbb{P}\big(S_k < t < S_{k+1}\big) &= \mathbb{P}\big(S_k < t < S_k + X_{k+1}\big) = \mathbb{P}\Big(\overline{S}_k < c < \overline{S}_k + \overline{X}_{k+1}\Big)\\ &= \mathbb{E}\Big(\mathbf{1}_{\{\overline{S}_k < c < \overline{S}_k + \overline{X}_{k+1}\}}\Big) = \mathbb{E}\Big(\mathbb{E}\Big(\mathbf{1}_{\{\overline{S}_k < c < \overline{S}_k + \overline{X}_{k+1}\}} \;\Big|\; \overline{S}_k\Big)\Big)\\ &= \mathbb{E}\Big(\mathbb{E}\Big(\mathbf{1}_{\{\overline{S}_k < c\}}\mathbf{1}_{\{\overline{X}_{k+1} > c - \overline{S}_k\}} \;\Big|\; \overline{S}_k\Big)\Big)\\ &= \mathbb{E}\Big(\mathbb{P}\Big(\overline{X}_{k+1} > c - \overline{S}_k \;\Big|\; \overline{S}_k\Big)\mathbf{1}_{\{\overline{S}_k < c\}}\Big)\end{aligned} \tag{2.11}$$

It is clear that $\mathbb{P}\big(\overline{X}_{k+1} > c - \overline{S}_k \mid \overline{S}_k\big) = \mathbb{P}\big(X_{k+1} > t - S_k \mid S_k\big)$. Then, if we denote $\tau = \frac{1}{k^{1/\alpha}}$, we have

$$\begin{aligned}\mathbb{P}\big(X_{k+1} > t - S_k \mid S_k\big) &= \begin{cases} (t - S_k)^{-\alpha}, & \text{if} \quad t - S_k > 1;\\ 1, & \text{else.}\end{cases}\\ &= \begin{cases} \dfrac{1}{k}\,(c - \overline{S}_k)^{-\alpha}, & \text{if} \quad 0 < \overline{S}_k < c - \tau;\\ 1, & \text{if} \quad c - \tau < \overline{S}_k < c.\end{cases}\end{aligned} \tag{2.12}$$

Since $\mathbb{P}\big(S_k < t < S_{k+1}\big) = 0$, where $0 < t < k$, we denote $c_0 := k/k^{1/\alpha}$ and have $\mathbb{P}\big(\overline{S_k} < c < \overline{S_{k+1}}\big) = 0$, where $0 < c < c_0$. By Theorem 2.2.9 we have that for any k and h,

$$\mathbb{P}\Big(\frac{S_k}{k^{1/\alpha}} \in (x, x+h)\Big) = \mathbb{P}\big(Y \in (x, x+h)\big) + o(1)\left(h + \frac{1}{k^{1/\alpha}}\right),$$

where Y is a stable law with parameter α. Let A be the point of maximum of the pdf $f_Y(y)$. By Lemmas 2.2.11 and 2.2.12 the calculation splits into two cases for which we use the following approximations of the limit law Y or the law of $\overline{S_k}$ directly:

1. If $c_0 < c < A$, then $f_{\overline{S_k}}(y) \asymp e^{-y^{-\alpha}}$ on the interval $y \in (0, c)$;
2. If $A < c < \infty$, then $f_Y(y) \asymp y^{-\alpha-1}$ on the interval $y \in (c, \infty)$.

Case 1: $\mathbf{c_0 < c < A}$. Make a diadic split of the interval $(0, c) = \bigcup_{i=0}^{\infty} c\big(1 - \frac{1}{2^i}, 1 - \frac{1}{2^{i+1}}\big)$. Then we can rewrite the expectation in (2.11) in the following way

$$\begin{aligned}\mathbb{E}\Big(\mathbb{P}\Big(\overline{X}_{k+1} > c - \overline{S}_k \mid \overline{S}_k\Big)\mathbf{1}_{\{\overline{S}_k < c\}}\Big) &= \int_0^{c/2} \frac{1}{k}(c-z)^{-\alpha} d\mathbb{P}_{\overline{S}_k}(z) +\\ &+ \sum_{i=1}^{\log_2 ck^{1/\alpha}} \int_{c(1-\frac{1}{2^i})}^{c(1-\frac{1}{2^{i+1}})} \frac{1}{k}(c-z)^{-\alpha} d\mathbb{P}_{\overline{S}_k}(z)\\ &+ \mathbb{P}\big(\overline{S}_k \in (c-\tau, c)\big) = (I) + (II) + (III)\end{aligned} \tag{2.13}$$

The first integral (I) can be calculated using the Lemma 2.2.12 in the following way

$$(I) = \int_0^{c/2} \frac{1}{k}(c-z)^{-\alpha} d\mathbb{P}_{\overline{S}_k}(z) \asymp t^{-\alpha}\mathbb{P}\big(\overline{S}_k \in (0, c/2)\big) \asymp t^{-\alpha} e^{-c^{-\alpha}}. \tag{2.14}$$

The second integral (II) can be represented as follows. Denote $c(k) = \log_2 ck^{1/\alpha}$. Then we can rewrite the integrals as

$$(II) \asymp \sum_{i=1}^{c(k)} t^{-\alpha} \left(\frac{1}{2^{i+1}}\right)^{-\alpha} \mathbb{P}\big(\overline{S}_k \in c\left(1-\frac{1}{2^i}, 1-\frac{1}{2^{i+1}}\right)\big) \tag{2.15}$$

Applying the Local Limit Theorem 2.2.9 we obtain

$$\mathbb{P}\big(\overline{S}_k \in c\left(1-\frac{1}{2^i}, 1-\frac{1}{2^{i+1}}\right)\big) = \mathbb{P}\big(Y \in c\left(1-\frac{1}{2^i}, 1-\frac{1}{2^{i+1}}\right)\big) + o(1)\left(\frac{1}{2^{i+1}} + \frac{1}{k^{1/\alpha}}\right) \asymp$$
$$\asymp \frac{c}{2^{i+1}} e^{-c^{-\alpha}} + o(1)\left(\frac{1}{2^{i+1}} + \frac{1}{k^{1/\alpha}}\right),$$

Then substituting this into (2.15) we obtain

$$(II) \asymp \sum_{i=1}^{c(k)} t^{-\alpha} \left(\frac{1}{2^{i+1}}\right)^{-\alpha} \left(\frac{c}{2^{i+1}} e^{-c^{-\alpha}} + o(1)\left(\frac{1}{2^{i+1}} + \frac{1}{k^{1/\alpha}}\right)\right) =$$
$$= t^{-\alpha} c e^{-c^{-\alpha}} \sum_{i=1}^{c(k)} \left(\frac{1}{2^{i+1}}\right)^{1-\alpha} + o(1) t^{-\alpha} \sum_{i=1}^{c(k)} \left(\frac{1}{2^{i+1}}\right)^{1-\alpha} + o(1) t^{-\alpha} \frac{1}{k^{1/\alpha}} \sum_{i=1}^{c(k)} \left(\frac{1}{2^{i+1}}\right)^{-\alpha}$$

The first two sums converge to a constant since they are geometric series with the parameter less than one. The last sum is calculated as follows:

$$\sum_{i=1}^{c(k)} \left(\frac{1}{2^{i+1}}\right)^{-\alpha} = 2^2 \sum_{i=0}^{c(k)-2} (2^\alpha)^i = 2^2 \frac{2^{\alpha \log_2 ck^{1/\alpha} - 1} - 1}{2^\alpha - 1} \asymp$$
$$\asymp c^\alpha k = t^\alpha,$$

and hence we can finally write

$$(II) \asymp t^{-\alpha} c e^{-c^{-\alpha}} + o(1) t^{-\alpha} + o(1) \frac{1}{k^{1/\alpha}}. \tag{2.16}$$

The last integral (III) can be simply written by Theorem 2.2.9 as

$$(III) = \mathbb{P}\big(\overline{S}_k \in (c-\tau, c)\big) = \mathbb{P}\big(Y \in (c-\tau, c)\big) + o(1)\left(\frac{1}{k^{1/\alpha}} + \frac{1}{k^{1/\alpha}}\right) \asymp$$
$$\asymp e^{-c^{-\alpha}} \frac{1}{k^{1/\alpha}} + o(1) \frac{1}{k^{1/\alpha}}. \tag{2.17}$$

Hence we plug (2.14), (2.16) and (2.17) into (2.13) and obtain the final result of this case

$$\mathbb{P}\big(S_k < t < S_{k+1}\big) \asymp t^{-\alpha} e^{-c^{-\alpha}} + t^{-\alpha} c\, e^{-c^{-\alpha}} + e^{-c^{-\alpha}} \frac{1}{k^{1/\alpha}} + o(1) t^{-\alpha} + o(1) \frac{1}{k^{1/\alpha}} \tag{2.18}$$

Case 2: $\mathbf{A < c < \infty}$. Make the same diadic split of the interval $(0, c) = \bigcup_{i=0}^{\infty} c\left(1-\frac{1}{2^i}, 1-\frac{1}{2^{i+1}}\right)$. Then we can split the expectation into three similar integrals:

$$\mathbb{E}\Big(\mathbb{P}\Big(\overline{X}_{k+1} > c - \overline{S}_k \mid \overline{S}_k\Big)\mathbf{1}_{\overline{S}_k < c}\Big) = \int_0^{c/2} \frac{1}{k}(c-z)^{-\alpha} d\mathbb{P}_{\overline{S}_k}(z) + \sum_{i=1}^{c(k)} \int_{c(1-\frac{1}{2^i})}^{c(1-\frac{1}{2^{i+1}})} \frac{1}{k}(c-z)^{-\alpha} d\mathbb{P}_{\overline{S}_k}(z)$$
$$+ \mathbb{P}\big(\overline{S}_k \in (c-\tau, c)\big) = (I) + (II) + (III) \tag{2.19}$$

Then the first integral (I) can be rewritten in the following way

$$(I) = \int_0^{c/2} \frac{1}{k}(c-z)^{-\alpha} d\mathbb{P}_{\overline{S}_k}(z) \asymp t^{-\alpha}\mathbb{P}\big(\overline{S}_k \in (0, c/2)\big).$$

Since c is bounded away from zero, then the probability $\mathbb{P}(\overline{S}_k \in (0, c/2)) \to 1$ as $c \to \infty$ and is $O(1)$ on this interval. Therefore,

$$(I) \asymp t^{-\alpha}. \tag{2.20}$$

The second integral (II) can be presented in the following way

$$(II) = \sum_{i=1}^{c(k)} t^{-\alpha} \left(\frac{1}{2^{i+1}}\right)^{-\alpha} \mathbb{P}\big(\overline{S}_k \in c\left(1 - \frac{1}{2^i}, 1 - \frac{1}{2^{i+1}}\right)\big), \tag{2.21}$$

where $c(k) = \log_2 ck^{1/\alpha}$. Applying the Local Limit Theorem 2.2.9 and the tail approximation from Lemma 2.2.11 we obtain

$$\mathbb{P}\big(\overline{S}_k \in c\left(1 - \frac{1}{2^i}, 1 - \frac{1}{2^{i+1}}\right)\big) = \mathbb{P}\big(Y \in c\left(1 - \frac{1}{2^i}, 1 - \frac{1}{2^{i+1}}\right)\big) + o(1)\left(\frac{1}{2^{i+1}} + \frac{1}{k^{1/\alpha}}\right) \asymp$$
$$\asymp \frac{1}{2^{i+1}} c^{-\alpha-1} + o(1)\left(\frac{1}{2^{i+1}} + \frac{1}{k^{1/\alpha}}\right),$$

Then substituting this into (2.21) we obtain

$$(II) \asymp \sum_{i=1}^{c(k)} t^{-\alpha} \left(\frac{1}{2^{i+1}}\right)^{-\alpha} \left(\frac{1}{2^{i+1}} c^{-\alpha-1} + o(1)\left(\frac{1}{2^{i+1}} + \frac{1}{k^{1/\alpha}}\right)\right) =$$
$$= t^{-\alpha} c^{-\alpha-1} \sum_{i=1}^{c(k)} \left(\frac{1}{2^{i+1}}\right)^{1-\alpha} + o(1) t^{-\alpha} \sum_{i=1}^{c(k)} \left(\frac{1}{2^{i+1}}\right)^{1-\alpha} + o(1) t^{-\alpha} \frac{1}{k^{1/\alpha}} \sum_{i=1}^{c(k)} \left(\frac{1}{2^{i+1}}\right)^{-\alpha}$$

Using the same arguments about the sums as in the case 1, we obtain

$$(II) \asymp t^{-\alpha} c^{-\alpha-1} + o(1) t^{-\alpha} + o(1) \frac{1}{k^{1/\alpha}}. \tag{2.22}$$

The last integral (III) can be simply written by the Local Limit Theorem 2.2.9 as follows:

$$\begin{aligned}(III) &= \mathbb{P}\big(\overline{S}_k \in (c-\tau, c)\big) = \mathbb{P}\big(Y \in (c-\tau, c)\big) + o(1)\left(\frac{1}{k^{1/\alpha}} + \frac{1}{k^{1/\alpha}}\right) \asymp \\ &\asymp c^{-\alpha-1} \frac{1}{k^{1/\alpha}} + o(1) \frac{1}{k^{1/\alpha}}.\end{aligned} \tag{2.23}$$

Hence we plug (2.20), (2.22) and (2.23) into (2.19) and obtain the final result in case 2:

$$\mathbb{P}\big(S_k < t < S_{k+1}\big) \asymp t^{-\alpha} + c^{-\alpha-1}\left(1 + \frac{1}{k^{1/\alpha}}\right) + o(1) t^{-\alpha} + o(1) \frac{1}{k^{1/\alpha}}. \tag{2.24}$$

Let us summarize the results of the cases. Let $c = t/k^{1/\alpha} < A$ then we have:

$$\mathbb{P}\big(S_k < t < S_{k+1}\big) \asymp t^{-\alpha} e^{-c^{-\alpha}} + t^{-\alpha} c\, e^{-c^{-\alpha}} + e^{-c^{-\alpha}} \frac{1}{k^{1/\alpha}} + o(1)\left(t^{-\alpha} + \frac{1}{k^{1/\alpha}}\right), \tag{2.25}$$

and if $A < t/k^{1/\alpha}$ then

$$\mathbb{P}\big(S_k < t < S_{k+1}\big) \asymp t^{-\alpha} + t^{-\alpha}c^{-\alpha} + c^{-\alpha-1}\frac{1}{k^{1/\alpha}} + o(1)\left(^{-\alpha} + \frac{1}{k^{1/\alpha}}\right). \tag{2.26}$$

Next we study the behaviour of $\mathbb{P}\big(S_k < t < S_{k+1}\big)$ for different values of k and t. Consider the early-time behaviour, or $t \asymp k$. Then $t/k^{1/\alpha} << A$ and we have

$$\mathbb{P}\big(S_k < t < S_{k+1}\big) \asymp \left(\frac{1}{k^\alpha} + \frac{1}{k^{1/\alpha}} + o(\frac{1}{k^2})\right) e^{-k^{1-\alpha}} + o(1)\left(\frac{1}{k^\alpha} + \frac{1}{k^{1/\alpha}}\right) \sim \frac{1}{k}e^{-c^{-\alpha}}.$$

Consider the mid-time regime, or $t \asymp k^{1/\alpha}$. Then $c = t/k^{1/\alpha} \asymp A$ and $e^{-c^{-\alpha}} \asymp e^{-A}$ and we have from (2.25) the following bound:

$$\mathbb{P}\big(S_k < t < S_{k+1}\big) \asymp \frac{1}{k} + \frac{1}{k} + \frac{1}{k^{1/\alpha}} + o(1)\left(\frac{1}{k^\alpha} + \frac{1}{k^{1/\alpha}}\right) \asymp$$
$$\asymp \frac{1}{k}(1 + o(1)) + \frac{1}{k^{1/\alpha}}(1 + o(1)).$$

and from (2.26) we have the same bound:

$$\mathbb{P}\big(S_k < t < S_{k+1}\big) \asymp \frac{1}{k}(1 + o(1)) + \frac{1}{k^{1/\alpha}}(1 + o(1)).$$

In the late-time regime, or $t/k^{1/\alpha} > A$, we have from (2.26) the following bound:

$$\mathbb{P}\big(S_k < t < S_{k+1}\big) \asymp t^{-\alpha}. \tag{2.27}$$

Finally we summarize that for any fixed k,

$$\mathbb{P}\big(S_k < t < S_{k+1}\big) \asymp \begin{cases} 0, & \text{if} \quad 0 < t < k; \\ \frac{1}{k}e^{-t^{-\alpha}k}, & \text{if} \quad k < t < O(k^{1/\alpha}); \\ t^{-\alpha}, & \text{if} \quad O(k^{1/\alpha}) < t < \infty. \end{cases} \tag{2.28}$$

Using the above estimations in (2.10) we may write that

$$\mathbb{P}(N_t = n) \asymp \begin{cases} 0, & \text{if } 0 < t < n; \\ \delta(n,t)t^{-\alpha}, & \text{if } n < t < n^{1/\alpha}; \\ nt^{-\alpha}, & \text{if } n^{1/\alpha} < t, \end{cases} \tag{2.29}$$

where $\delta(n,t)$ is a monotonic function in n and t such that $\delta(n,t) \to 0$ as $t \to n$ and $\delta(n,t) \to n$ as $t \to n^{1/\alpha}$. We observe that $\mathbb{P}(T_k > t) = \mathbb{P}(N_t < k)$, therefore for any fixed k and $t := \tau k^{1/\alpha}$ we have for $\tau > 1$

$$\mathbb{P}(T_k > \tau k^{1/\alpha}) = \mathbb{P}(N_{\tau k^{1/\alpha}} < k) = \sum_{i=0}^{k-1} \mathbb{P}(N_{\tau k^{1/\alpha}} = i) \asymp \sum_{i=0}^{k-1} i\,(\tau k^{1/\alpha})^{-2\alpha} \asymp \tau^{-2\alpha}. \tag{2.30}$$

The integral of the tail is finite for $\alpha \in (1/2, 1)$, therefore the expectation can be bounded by

$$\mathbb{E}(T_k) \asymp k^{1/\alpha}.$$

This finishes the proof of the Theorem. □

2.3.2 Spreading on a star

In the following Subsection we consider the example of the SI spreading process with power-law weights an n-star. The graph ST_n of an n-star is defined as the distinguished root vertex 0 and vertices $[n] = \{1, 2, \ldots, n\}$ attached to it. We consider the SI spreading process $T = (T_k)_{k=1}^n$ with power-law distributed random weights with $\alpha \in (1/2, 1)$, denoted as $X_1, X_2, \ldots, X_n$. The bound on the expected value $\mathbb{E}(T_k)$, where $k \in \{1, \ldots, n-1\}$, is given by the following Theorem.

Theorem 2.3.2. *Let ST_n be the graph of n-star. Then in the SI spreading process $T = (T_k)_{k=1}^n$ the expected time to infect k vertices is bounded up until $k \leqslant n-1$, and for $n \geqslant 2$,*

$$\mathbb{E}(T_k) \leqslant Ck^{1/\alpha}.$$

Proof. The geometry of the graph lets us directly write

$$T_k = X_{(k)},$$

where $X_{(k)}$ denotes the k'th order statistic of $X_1, X_2, \ldots, X_n$.

In further calculation we use T_{n-k} for convenience. The density of the distribution of T_{n-k} by definition is given by the following formula:

$$\begin{aligned} p_{T_{n-k}}(t) &= \frac{n!}{(n-k-1)!k!}(F_X(t))^{n-k-1}(1-F_X(t))^k p_X(t) = \\ &= \alpha\frac{n!}{(n-k-1)!k!}(1-t^{-\alpha})^{n-k-1}t^{-(k+1)\alpha-1} \end{aligned}$$

In order to simplify our calculations, denote $C_{n,k} := \frac{n!}{(n-k-1)!k!}$. The expectation of T_{n-k} is then given by the following formula:

$$\begin{aligned} \mathbb{E}(T_{(n-k)}) &= \int_1^\infty t p_{T_{n-k}}(t)dt = \int_1^{(n-k-1)^{1/\alpha}} t p_{T_{n-k}}(t)dt + \int_{(n-k-1)^{1/\alpha}}^\infty t p_{T_{n-k}}(t)dt = \\ &= (I) + (II). \end{aligned}$$

Denote $c_{n,k} := (n-k-1)^{1/\alpha}$ and bound the integrals separately. In the first integral (I) we take the function $(1-t^{-\alpha})^{n-k-1} < 1$, and hence

$$\begin{aligned} (I) &= \int_1^{c_{n,k}} t p_{T_{n-k}}(t)dt < \int_1^{c_{n,k}} \alpha C_{n,k} t^{-(k+1)\alpha}dt = \\ &= C_{n,k}(n-k-1)^{k+1}\frac{(n-k-1)^{1/\alpha}}{(k+1)-1/\alpha}. \end{aligned}$$

For the second integral (II) by Lemma 2.2.1 and obtain,

$$(II) = \int_{c_{n,k}}^{\infty} t p_{T_{n-k}}(t)dt \leqslant \alpha C_{n,k} \left[\int_{c_{n,k}}^{\infty} \left(1 - c_{n,k} t^{-\alpha} + \frac{c_{n,k}^2 t^{-2\alpha}}{2} \right) t^{-\alpha(k+1)} dt \right] =$$
$$= C_{n,k}(n-k-1)^{k+1}(n-k-1)^{1/\alpha} \frac{1}{2(k+2)}.$$

Show now that for all k, where $2 \leqslant k \leqslant n-2$ the product

$$C_{n,k}(n-k-1)^{k+1} < 1,$$

uniformly in n. Using the Stirling's formula $n! \sim \sqrt{2\pi} n^n e^{-n}$, we obtain

$$C_{n,k}(n-k-1)^{k+1} \sim \frac{e^{-1}}{\sqrt{2\pi}} \sqrt{\frac{n}{(n-k-1)k}} \left(\frac{n}{(n-k-1)k} \right)^n \frac{1}{k^k} =$$
$$= \frac{e^{-1}}{\sqrt{2\pi}} \sqrt{\frac{n}{(n-k-1)k}} \left(1 - \frac{k+1}{n} \right)^{-n} \frac{1}{k^k} \sim$$
$$\sim \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{(n-k-1)k}} \frac{e^k}{k^k} < 1.$$

In case when $k = 1$ we immediately have that there exist $C_1 > 0$, such that

$$C_{n,k}(n-k-1)^{k+1} = \frac{n(n-1)}{(n-2)^2} < C_1,$$

when $n \geqslant 3$. When $k = n-1$ we directly calculate the expectation and there exist $C_2 > 0$, such that

$$\mathbb{E}(T_{(1)}) = \int_1^{\infty} t p_{T_1}(t)dt = \alpha n \int_1^{\infty} t^{\alpha n} dt =$$
$$= \frac{1}{1 - 1/\alpha n} < C_2,$$

when $n \geqslant 3$. Summarizing the equations we have for $1 \leqslant k \leqslant n-1$:

$$\mathbb{E}(T_{(n-k)}) \leqslant C(n-k)^{1/\alpha},$$

where $C := \max(1, C_1, C_2)$. This finishes proof of the Theorem. □

2.4 Smoothing on general graphs

Consider a connected graph $G = (V, E)$ with n vertices and a root s and m edges. Denote the vertex set as $[n] = \{0, 1, 2, \ldots, n\}$ and let $s = 0$. Denote the i.i.d. random weights on the edges as $X_1, X_2, \ldots, X_m$, and let X_1 be defined on the probability space $\Omega = (\mathbb{R}_+, \mathcal{F}, \mathbb{P})$.

Denote as $P(s,t)$ the shortest weighted (s,t)-path and as $|P(s,t)|$ the total weight of the (s,t)-path. Let G_k be an induced subgraph of G that has k vertices and the same root s, and denote it as $G_k \sqsubset G$,.

Then we can define probability space $\Omega = \prod_{i=1}^{m}(\mathbb{R}_+, \mathcal{F}, \mathbb{P}) = (\mathbb{R}_+^m, \mathcal{F}^m, \mathbb{P}^m)$ of all possible random assignments of weights to the edges of the graph G with the product measure $P^m(d\omega) = P(d\omega_1) \times \cdots \times P(d\omega_m)$, where $\omega \in \Omega$.

We view the SI spreading process on the graph G as a stochastic process $T = \{T_k : k \in \{1, \ldots, n\}\}$, where T_k is the random time to infect k vertices, defined as the minimum over k-vertex subgraphs $G_k \subset G$ of the maximum over vertices $t \in V(G_k)$ of total weights of (s,t)-paths:

$$T_k = \min_{G_k \subset G} \max_{t \in V(G_k)} |P(s,t)|,$$

The stochastic process T is defined on the space Ω, equipped with natural filtration $\mathcal{F} = \{\mathcal{F}_k : k \in \{1, \ldots, n\}\}$. Denote the sample sequence $T(\omega) = \{T_k(\omega) : k \in \{1, \ldots, n\}\}$, where $\omega \in \Omega$ and let the set of occupied edges $\mathcal{O}(T_k(\omega))$, where $k \in \{1, \ldots, n\}$, be the set of indices of random variables of the occupied edges by time t in the sample sequence $T(\omega)$. In other words,

$$T_k(\omega) = \sum_{i \in \mathcal{O}(T_k(\omega))} X_i.$$

Let the *front of the epidemic* $F(T_k(\omega))$ be the set of active edges at time $T_k(\omega)$, where $k \in \{1, \ldots, n\}$, in the sample sequence $T(\omega)$. Define $\kappa(G)$ to be the minimal number of vertices k such that for each sample sequence $T(\omega)$ and for each $j < k$, the front $F(T_j(\omega))$ has at least two active edges. We say the active edge e is *old* and has age $\tau > 0$ at time t, if the edge has become active at time $t - \tau$. If $\tau = 0$ then an edge is called *new*.

We prove that if we can find a sample sequence $T(\omega)$ such that the front of the epidemics $F(T_i(\overline{\omega}))$ has one active edge, then there will be an infinite jump at this point on the spreading curve.

Lemma 2.4.1. *Let T be the SI spreading process on a graph G with power-law weights where $\alpha < 1$ and let there exist a sample sequence $T(\overline{\omega})$ and a number $i \in \mathbb{N}$, such that the front $F(T_i(\overline{\omega}))$ has one active edge. Then for each j, where $i + 1 \leqslant j \leqslant n$, the expected value*

$$\mathbb{E}(T_j) = \infty.$$

Proof. Since T_k is a function of $X_1, \ldots, X_m$ and since the graph G, the min and max in the definition of T_k are taken over a finite number of combinations, therefore there exists $\varepsilon > 0$

and a measurable subset $\Omega^\varepsilon \subset \Omega$, such that for every $\omega \in \Omega^\varepsilon$, the front of the epidemics $F(T_i(\omega))$ has one active edge. Formally, the set Ω^ε is defined as $\Omega^\varepsilon = \{\omega^\varepsilon \in \Omega : \forall j \in \mathcal{O}(T_i(\omega^\varepsilon)),\ |\omega_j^\varepsilon - \overline{\omega}_j| \leqslant \varepsilon\}$ and contains $\overline{\omega}$ by definition. Then, by definition of the expected value:

$$\begin{aligned}\mathbb{E}(T_{i+1}) = \int\limits_{\Omega} T_{i+1}(\omega)P^m(d\omega) = \int\limits_{\Omega\setminus\Omega^\varepsilon} T_{i+1}(\omega)P^m(d\omega) + \int\limits_{\Omega^\varepsilon} T_{i+1}(\omega)P^m(d\omega) \geqslant \\ \geqslant \int\limits_{\Omega^\varepsilon} T_{i+1}(\omega)P^m(d\omega).\end{aligned} \tag{2.31}$$

Denote the edge that is active at time T_{i+1} as X, and suppose it has age $t \geqslant 0$. Then the last expected value in (2.31) can be rewritten as

$$\int\limits_{\Omega^\varepsilon} T_{i+1}(\omega)P^m(d\omega) \geqslant T_{i+1}(\overline{\omega}) + \mathbb{E}(X - t|X > t) \geqslant \mathbb{E}(X - t|X > t) \geqslant \mathbb{E}(X) = \infty,$$

when $\alpha \in (1/2, 1)$. Since for all later times T_j we have $T_{i+1} \leqslant T_j$, therefore $\mathbb{E}(T_j) = \infty$, for each j, where $i + 1 \leqslant j \leqslant n$.

This finishes the proof of the Lemma. □

Based on the above Lemma 2.4.1 we can equivalently define the number $\kappa(G)$ to be the maximal number of vertices k such that $\mathbb{E}(T_k) < \infty$.

Define the delayed process $\overline{T} = \{\overline{T}_k : k \in \{1, \ldots, m\}\}$ on the same space. Index the edge set of the graph G as $E = \{e_1, e_2, \ldots, e_m\}$ in any order $\mathcal{P}$. The process $\overline{T}$ is coupled with the process T and spreads the infection along the edges of E in the following manner. Start with all vertices in the susceptible state S and let the vertex s be infected (in the state I). Then choose the two available edges incident to s with smallest indices in order $\mathcal{P}$ and spread the infection through them. At the time when one of these edges becomes occupied, choose the next available edge from E with the smallest index in $\mathcal{P}$ and repeat the procedure. If both active edges share one susceptible vertex, then when one edge gets occupied, choose two new available edges with smallest indices in $\mathcal{P}$. The process runs until there are no more available edges.

The process $\overline{T}$ resembles the process T, but has a prescribed order of occupation of the edges and always has at most two active edges in the front of the epidemics. Let us prove the following Lemma.

Lemma 2.4.2. *Let $\overline{T}$ be the delayed process on the tree G. Then for any i, where $0 < i < \kappa(G)$ and any $\widetilde{\omega} \in \Omega$ at the time $\overline{T}_i(\widetilde{\omega})$ the front of epidemics $F(\overline{T}_i(\widetilde{\omega}))$ has at least two active edges.*

Proof. Suppose that for some $i < \kappa(G)$ there is only one active edge in $F(\overline{T}_i(\widetilde{\omega}))$. Since the processes T and $\overline{T}$ are coupled, then there exists $\omega \in \Omega$ such that the sample sequence $\overline{T}(\widetilde{\omega}) = T(\omega)$, and $F(\overline{T}_i(\widetilde{\omega})) = F(T_i(\omega))$. Hence, $F(T_i(\omega))$ has one active edge for $i < \kappa(G)$, which contradicts the definition of $\kappa(G)$. □

The process $\overline{T}$ stochastically dominates the process T, which is proved in the following Lemma.

Lemma 2.4.3. *For any tree G on n vertices, for k such that $1 \leqslant k < \kappa(G)$,*

$$T_k \preceq \overline{T}_k$$

Proof. Consider any sample sequence $T(\omega)$. Then by definition the time $T_k(\omega)$ can be written as follows:

$$T_k(\omega) = \min_{G_k \subset G} \max_{t \in V(G_k)} |P(s,t)|.$$

Since processes $\overline{T}$ and T are coupled, then there exist $\widetilde{\omega}$ such that $\overline{T}(\widetilde{\omega}) = T(\omega)$. Define the induced subgraph $\widetilde{G}_k \subset G$ on k vertices occupied by the process $\overline{T}$ by time $\overline{T}_k(\omega)$. Then the time to occupy k vertices in the process $\overline{T}$ can be defined as the maximum over vertices in $V(\widetilde{G}_k)$ of weighted lengths of (s,t)-paths:

$$\overline{T}_k = \max_{t \in V(\widetilde{G}_k)} |P(s,t)|.$$

Since $\widetilde{G}_k$ is an induced subgraph, then the time $T_k(\omega) \leqslant \overline{T}_k(\widetilde{\omega})$, and therefore by Theorem 2.2.7 we have $T_k \preceq \overline{T}_k$. This finishes the proof of the Lemma. □

By definition of the process $\overline{T}$, at each time $\overline{T}_k$ it infects a vertex and the front has exactly two active edges with i.i.d. power-law weights X and Y, where X is the newly available edge and Y is either a new or an old edge. Define the random variable $Y_\tau := (Y - \tau | \tau, Y > \tau)$, where $\tau \geqslant 0$, to have the following probability measure:

$$\mathbb{P}(Y_\tau > s) = \frac{\mathbb{P}(Y > \tau + s)}{\mathbb{P}(Y > \tau)}.$$

Therefore, the process $\overline{T}$ satisfies the following equation:

$$\overline{T}_{k+1} = \overline{T}_k + \min(X, Y_\tau),$$

for each $k \geqslant 1$. Since the process starts with two new edges, then in case $\overline{T}_1$ is defined as follows:

$$\overline{T}_1 = \min(X, Y_0) = \min(X, Y).$$

Define the process $Q = \{Q_k : k \in \{1, \dots, n\}\}$ to be the process in which at each time there are two active edges in the front, one of them is always old with the age of the process and one is always new at each time t. When a new edge gets occupied a new edge is placed instead. When an old edge gets occupied, both edges are replaced with an old edge with an age of the process and a new edge. The process Q qualitatively constitutes the worst scenario the infection $\overline{T}$ can spread, having an ever old edge Y and spreading only along new edges X. Thus, the process Q then is defined by the following recursion:

$$\begin{aligned} Q_{k+1} &= Q_k + \min(X, Y - Q_k | Q_k, Y > Q_k), \\ Q_1 &= \min(X, Y). \end{aligned} \tag{2.32}$$

It is easy to see that since for each τ in the process $\overline{T}$ it is true that $Q_k \geqslant \tau$, thus we have

$$\mathbb{P}(Y - Q_k > s | Q_k, Y > Q_k) > \mathbb{P}(Y - \tau > s | \tau, Y > \tau),$$

thus $Q_k \geqslant \overline{T}_k$ and hence by Theorem 2.2.7 the process Q stochastically dominates the process $\overline{T}$.

We now prove the following technical Lemma.

Lemma 2.4.4. *Let X and Y be i.i.d. power-law random variables with $\alpha \in (1/2, 1)$. Then there exists constant $C = C(\alpha)$, such that for any $t > 1$:*

$$\mathbb{E}(\min\{X, Y - t\} | Y > t) \leqslant C t^{1-\alpha}.$$

Proof. The tail distribution of the minimum of considered random variables is given as follows:

$$\begin{aligned} \mathbb{P}\big(\min\{X, Y - t\} > s | Y > t\big) &= \frac{\mathbb{P}(X > s, Y - t > s, Y > t)}{\mathbb{P}(Y > t)} = \\ &= \begin{cases} t^{-\alpha} \left(\frac{s}{t}\left(1 + \frac{s}{t}\right)\right)^{-\alpha}, & \text{if } s > 1; \\ \left(1 + \frac{s}{t}\right)^{-\alpha}, & \text{if } 0 < s < 1. \end{cases} \end{aligned}$$

Then using the substitution $u = \frac{s}{t}$ we write the expected value as follows:

$$\begin{aligned}\mathbb{E}(\min\{X, Y-t\}|Y>t) &= \int_0^\infty \mathbb{P}\big(\min\{X, Y-t\} > s|Y>t\big)ds = \\ &= \int_0^1 \left(1+\frac{s}{t}\right)^{-\alpha} ds + t^{-\alpha}\int_1^\infty \left(\frac{s}{t}\left(1+\frac{s}{t}\right)\right)^{-\alpha} ds = \\ &= t\int_0^{1/t} (1+u)^{-\alpha} du + t^{1-\alpha}\int_{1/t}^\infty \left(u\left(1+u\right)\right)^{-\alpha} du = \\ &= (I) + (II)\end{aligned}$$

Let us calculate both integrals separately. The first integral (I) is straightforward to compute:

$$(I) = t\frac{1}{1-\alpha}\left[\left(1+\frac{1}{t}\right)^{1-\alpha} - 1\right] := f(t)$$

Using the L'Hospital rule it is easy to show that

$$f(t) \asymp 1.$$

Calculate the second integral (II). Splitting the interval of integration into two parts we obtain

$$t^{1-\alpha}\int_{1/t}^\infty \left(u\left(1+u\right)\right)^{-\alpha} du = t^{1-\alpha}\left[\int_{1/t}^1 \left(u\left(1+u\right)\right)^{-\alpha} du + \int_1^\infty \left(u\left(1+u\right)\right)^{-\alpha} du\right] \tag{2.33}$$

The first integral on the r.h.s. of (2.33) can be bounded in the following way:

$$\frac{1}{\sqrt{2}}\int_{1/t}^1 u^{-\alpha} du \leqslant \int_{1/t}^1 \left(u\left(1+u\right)\right)^{-\alpha} du \leqslant \int_{1/t}^1 u^{-\alpha} du,$$

or, equivalently,

$$\int_{1/t}^1 \left(u\left(1+u\right)\right)^{-\alpha} du \asymp \int_{1/t}^1 u^{-\alpha} du = \frac{1}{1-\alpha}\left(1-t^{\alpha-1}\right).$$

The second integral on the r.h.s. of (2.33) can be bounded in the following way:

$$\int_1^\infty (1+u)^{-2\alpha} du \leqslant \int_1^\infty \left(u\left(1+u\right)\right)^{-\alpha} du \leqslant \int_1^\infty u^{-2\alpha} du,$$

or, equivalently,

$$\int_1^\infty \left(u\left(1+u\right)\right)^{-\alpha} du \asymp \int_1^\infty u^{-2\alpha} du = \frac{1}{2\alpha-1}.$$

Hence, the integral (II) is bounded in the following way:

$$t^{1-\alpha}\int\limits_{1/t}^{\infty}\left(u\left(1+u\right)\right)^{-\alpha}du \asymp t^{1-\alpha}\left(\frac{1}{2\alpha-1}+\frac{1}{1-\alpha}\left(1-t^{\alpha-1}\right)\right).$$

Summarizing the above calculations we obtain

$$\mathbb{E}(\min\{X,Y-t\}|Y>t) \asymp t\frac{1}{1-\alpha}\left[\left(1+\frac{1}{t}\right)^{1-\alpha}-1\right]+t^{1-\alpha}\frac{1}{2\alpha-1}+\frac{t^{1-\alpha}-1}{1-\alpha}. \tag{2.34}$$

Using $\alpha<1$ we write the following upper bound:

$$\begin{aligned}\mathbb{E}(\min\{X,Y-t\}|Y>t) &\asymp 1+t^{1-\alpha}\left(\frac{1}{2\alpha-1}+\frac{1}{1-\alpha}\right)-\frac{1}{1-\alpha}\\ &\leqslant t^{1-\alpha}\left(\frac{1}{2\alpha-1}+\frac{1}{1-\alpha}\right)\\ &\leqslant Ct^{1-\alpha}.\end{aligned}$$

This finishes proof of the Lemma. □

Remark: When $\alpha=1$ we have

$$\begin{aligned}\mathbb{E}(\min\{X,Y-t\}|Y>t) &= t\int\limits_{0}^{1/t}\frac{1}{1+u}du+\int\limits_{1/t}^{\infty}\frac{1}{u(u+1)}du=\\ &= t\ln\left(1+\frac{1}{t}\right)+\ln(t+1).\end{aligned}$$

It is easy to show that this is the limit of (2.34) when $\alpha\to 1$. □

Using the law of total expectation, Lemma 2.4.4 and Jensen's inequality we have that

$$\begin{aligned}\mathbb{E}(Q_{k+1}) &= \mathbb{E}(Q_k+\min\{X,Y-Q_k\}) = \mathbb{E}\big(\mathbb{E}(Q_k+\min\{X,Y-Q_k\})|Q_k\big)\\ &= \mathbb{E}\big(\mathbb{E}(Q_k|Q_k)+\mathbb{E}(\min\{X,Y-Q_k\}|Q_k)\big) = \mathbb{E}(Q_k)+\mathbb{E}\big(\mathbb{E}(\min\{X,Y-Q_k\}|Q_k)\big)\\ &\leqslant \mathbb{E}(Q_k)+C\mathbb{E}(Q_k^{1-\alpha})\\ &\leqslant \mathbb{E}(Q_k)+C\mathbb{E}(Q_k)^{1-\alpha}.\end{aligned}$$

Then immediately we have $\mathbb{E}(Q_k)\leqslant b_k$, where b_k is defined as

$$b_{k+1}=b_k+Cb_k^{1-\alpha},$$

$$b_1=\mathbb{E}(Q_1)=\frac{2\alpha}{2\alpha-1}\leqslant(\alpha C)^{1/\alpha}:=d.$$

Then by Lemma 2.2.5 we have the following Lemma about the modified process Q_k.

Lemma 2.4.5. *Consider the modified process Q_k defined above in (2.32). Then, for $\alpha \in (1/2, 1)$ and $k \geqslant 1$, we have*

$$\mathbb{E}(Q_k) \leqslant dk^{1/\alpha}.$$

where $d = (\alpha C)^{1/\alpha}$.

Now we are able to prove the main Theorem of the current Section.

Theorem 2.4.6. *Consider the graph G on n vertices and root s and m edges and the SI spreading process $T = (T_j)_{j=1}^m$ with power-law weights with $\alpha \in (1/2, 1)$. Then for each k, where $1 \leqslant k < \kappa(G)$, the expected time to infect k vertices is bounded by*

$$\mathbb{E}(T_k) \leqslant dk^{1/\alpha}.$$

Proof. By Lemma 2.4.1 the process T is stochastically dominated by the coupled process $\overline{T}$, thus the following inequality is valid:

$$\mathbb{E}(T_k) \leqslant \mathbb{E}(\overline{T}_k),$$

where $1 \leqslant k \leqslant \kappa(G)$. By Lemma 2.4.2 the process $\overline{T}$ is dominated by the process Q, thus

$$\mathbb{E}(\overline{T}_k) \leqslant \mathbb{E}(Q_k),$$

where $1 \leqslant k \leqslant \kappa(G)$. By Theorem 2.4.5 the last expectation is bounded:

$$\mathbb{E}(Q_k) \leqslant dk^{1/\alpha}.$$

Therefore, we have the final statement of the Lemma:

$$\mathbb{E}(T_k) \leqslant dk^{1/\alpha}.$$

This finishes the proof of the Theorem. □

2.5 Smoothing on critical Galton-Watson tree with an extra edge

The smoothing effect on the static graphs is described by the Theorem 2.4.6. It turns out that this Theorem can be applied to the random graph models with slight modifications. In the current Section we prove that w.h.p. the smoothing happens up to the positive fraction of a large CGW tree $\mathcal{T}$ with an extra edge e attached to a root and a random vertex in $\mathcal{T}$. First we establish some technical lemmas.

2.5.1 Technical lemmas

Let the graph G be a cycle C_n of length n with N rooted trees $t_1, t_2, \ldots, t_N$ attached to it by extra edges $e_1, e_2, \ldots, e_N$, where $n \geqslant 3$, $N \geqslant 1$ (see Figure2.3). Let vertex $s \in V(C_n)$ be a root. Consider the SI spreading process $T = (T_j)_{j=1}^{n+N}$ with power-law weights with $\alpha \in (1/2, 1)$. Then, it is a simple combinatorics argument to show the exact value of $\kappa(G)$.

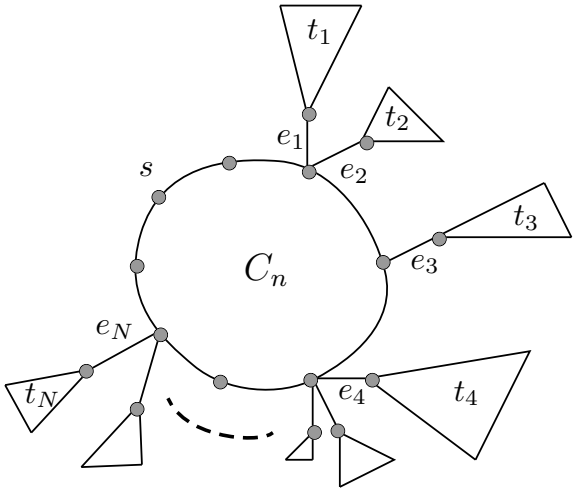


Figure 2.3. *Schematic structure of a cycle C_n with hanging trees $t_1, t_2, \ldots, t_N$.*

Lemma 2.5.1. *Let the graph G be a cycle C_n of length n with N rooted trees $t_1, t_2, \ldots, t_N$ attached to it by extra edges $e_1, e_2, \ldots, e_N$. Consider the SI spreading process T with power-law weights with $\alpha \in (1/2, 1)$, started from a vertex s on a cycle. Then $\kappa(G) = |G| - \max_i |t_i|$.*

Proof. Let $t_{(1)}, t_{(2)}, \ldots, t_{(N)}$ denote the labelling of trees in the descending order of their sizes and label each edge that attaches the tree $t_{(i)}$ to the cycle as $e_{(i)}$, where $i = 1, \ldots, N$. It is natural to couple T with the delayed process $\hat{T} = (\hat{T})_{j=1}^{n+N}$ constructed as follows. Start the delayed process $\overline{T}$, defined in the Section 2.4, from vertex s only on the cycle C_n until the last edge becomes active and let $\hat{T}_j = \overline{T}_j$, for $j = 1, 2, \ldots, n$. At that time we have all vertices incident to extra edges leading to attached trees in the state I. Then start the process $\overline{T}$ on two trees $t_{(1)} \cup e_{(1)}$ and $t_{(2)} \cup e_{(2)}$ and let $\hat{T}_j = \overline{T}$, where $j = n+1, \ldots, n+|t_{(2)}|$. When one of the trees become infected, start process $\overline{T}$ on the tree $t_{(3)}$ and proceed analogously up until there are no new trees to infect.

By Lemma 2.4.3 we see that the process $\hat{T}$ stochastically dominates the process T, and up to the same $\kappa(G)$.

It is obvious from Theorem 2.3.1 that $\kappa(G) \geqslant |C_n|$. Take a sample sequence $\hat{T}(\omega)$, such that $\xi(e_{(1)}) > \sum_j \xi_j$, where the sum is taken over all edges of trees, except $t_{(1)}$. Then up to a time when all trees get infected, the edge $e_{(1)}$ will be in the front of the epidemic $F(\hat{T}(\omega))$, resulting in smoothing up to $k_0 := |G| - |t_{(1)}|$ vertices. Take another sample sequence $\hat{T}(\overline{\omega})$, such that $\xi(e_{(i)}) > \sum_j \xi_j$, where $i \neq 1$ and the sum is taken over all edges of trees, except $t_{(i)}$, we obtain smoothing up to $k = |G| - |t_{(i)}| > k$ vertices. If there exist $e \in t_{(i)}$, where $i = 1, \ldots, N$, such that $\xi(e)$ is large, then following the same argument we obtain smoothing up to $k > k_0$ vertices. Hence, $\kappa(G) = k_0 = |G| - \max_i |t_i|$ and this finishes proof of the Lemma. □

Consider a critical Galton-Watson (CGW) tree $\mathcal{T}$ with offspring distribution ξ, such that $Var(\xi) = \sigma^2 < \infty$. Let $\mathcal{T}^\infty$ be an infinite CGW tree with the same offspring distribution ξ. We write as $(\mathcal{T}[k] = T)$ and $(\mathcal{T}^\infty[k] = T)$ the event that first k generations of the tree $\mathcal{T}$ and $\mathcal{T}^\infty$ respectively match the first k generations of a given tree T, where $1 \leqslant k \leqslant N$. Denote as $\#T_k$ the size of k'th generation in the tree T. The following Lemma holds for trees $\mathcal{T}$ and $\mathcal{T}^\infty$ [40].

Lemma 2.5.2. *Let $\mathcal{T}$ be a critical GW tree conditioned on $Z_N > 0$ with offspring distribution ξ. Then for any rooted labelled tree T of at least k generations:*

$$\lim_{N\to\infty} \mathbb{P}(\mathcal{T}[k] = T | Z_N > 0) = \#T_k \cdot \mathbb{P}(\mathcal{T}[k] = T).$$

Then

$$\mathbb{P}(\mathcal{T}^\infty[k] = T) = \#T_k \cdot \mathbb{P}(\mathcal{T}[k] = T).$$

Denote the conditioned measure $\mathbb{P}(\cdot \mid Z_N > 0) =: \mathbb{P}_N(\cdot)$ and the infinite measure $\mathbb{P}(\cdot \mid |\mathcal{T}| = \infty) =: \mathbb{P}_\infty(\cdot)$. It is natural that as $N \to \infty$ the trees $(\mathcal{T}|Z_N > 0)$ and $\mathcal{T}^\infty$ start to look same. The following Lemma shows it rigorously.

Lemma 2.5.3. *Let $\mathcal{T}$ be a critical GW tree conditioned on $Z_N > 0$ and $\mathcal{T}^\infty$ be an infinite critical GW tree. Then as $N \to \infty$ for any $\varepsilon > 0$ there exist $\delta > 0$, such that there exists a coupling between $\mathcal{T}$ and $\mathcal{T}^\infty$, for which*

$$\mathbb{P}(\mathcal{T}[\delta N] \neq \mathcal{T}^\infty[\delta N]) < \varepsilon.$$

Proof. Consider a rooted tree T with height $k \leqslant \delta N$. Then by Bayes' formula:

$$\begin{aligned} \mathbb{P}(\mathcal{T}[k] = T | Z_N > 0) &= \frac{\mathbb{P}(Z_N > 0 | \mathcal{T}[k] = T)}{\mathbb{P}(Z_N > 0)} \mathbb{P}(\mathcal{T}[k] = T) = \\ &= \frac{1}{\mathbb{P}(Z_N > 0)} \mathbb{P}(Z^{(1)}_{N-k} > 0 \cup \cdots \cup Z^{(\#T_k)}_{N-k} > 0) \mathbb{P}(\mathcal{T}[k] = T), \end{aligned} \tag{2.35}$$

where $Z^{(i)}_{N-k}$ denotes the $(N-k)$'th generation in the i'th copy of GW process Z started at level k. Thus, by Theorem 2.2.14 for a large N there exist $\varepsilon_1 > 0$ such that,

$$\mathbb{P}(Z_N > 0) = \frac{2}{\sigma^2 N}(1+\varepsilon_1). \tag{2.36}$$

Then using the union bound we have that when $N-k$ is large,

$$\mathbb{P}(\mathcal{T}[k] = T | Z_N > 0) \leqslant \frac{2\#T_k}{\sigma^2(N-k)}(1+\varepsilon_1).$$

In order to obtain a lower bound we further write the Equation (2.35) as follows:

$$\begin{aligned}\mathbb{P}(\mathcal{T}[k] = T | Z_N > 0) &= \frac{1}{\mathbb{P}(Z_N > 0)}\left(1 - \mathbb{P}(Z^{(1)}_{N-k} = 0 \cap \cdots \cap Z^{(\#T_k)}_{N-k} = 0)\right)\mathbb{P}(\mathcal{T}[k] = T) = \\ &= \frac{1}{\mathbb{P}(Z_N > 0)}\left(1 - (1 - \mathbb{P}(Z_{N-k} > 0))^{\#T_k}\right).\end{aligned} \tag{2.37}$$

Hence using (2.36) when $N-k$ is large there exist $\varepsilon_1' > 0$ such that the Equation (2.37) can be rewritten as follows:

$$\mathbb{P}(\mathcal{T}[k] = T | Z_N > 0) = \frac{\sigma^2 N}{2}\left(1 - \left(1 - \frac{2}{\sigma^2(N-k)}\right)^{\#T_k}\right)\mathbb{P}(\mathcal{T}[k] = T))(1+\varepsilon_1'). \tag{2.38}$$

We then rewrite (2.38) using the Taylor expansion of the logarithmic function $\ln(1-x)$ around $x = 0$:

$$\begin{aligned}\left(1 - \frac{2}{\sigma^2(N-k)}\right)^{\#T_k} &= \exp\left(\#T_k \ln\left(1 - \frac{2}{\sigma^2(N-k)}\right)\right) = \\ &= \exp\left(-\frac{2\#T_k}{\sigma^2(N-k)}\right)(1+o(1)).\end{aligned}$$

Since for any $x > 0$ we have $e^{-x} < 1 - x + x^2/2$ and when $N-k$ is large there exist $\varepsilon_2 > 0$, such that

$$\begin{aligned}\mathbb{P}(\mathcal{T}[k] = T | Z_N > 0) &\sim \frac{\sigma^2 N}{2}\left(1 - \exp\left(-\frac{2\#T_k}{\sigma^2(N-k)}\right)\right)\mathbb{P}(\mathcal{T}[k] = T) > \\ &> \frac{N}{N-k}\#T_k\mathbb{P}(\mathcal{T}[k] = T)(1-\varepsilon_2).\end{aligned}$$

Choosing ε_0 to be the maximum of ε_1 and ε_2 we have the following bounds on the probability $\mathbb{P}(\mathcal{T}[k] = T | Z_N > 0)$:

$$\frac{N}{N-k}\#T_k\mathbb{P}(\mathcal{T}[k] = T)(1-\varepsilon_0) \leqslant \mathbb{P}(\mathcal{T}[k] = T | Z_N > 0) \leqslant \frac{N}{N-k}\#T_k\mathbb{P}(\mathcal{T}[k] = T)(1+\varepsilon_0). \tag{2.39}$$

Next we prove that the probability that the tree $\mathcal{T}$ has a large number of vertices on the layer k is small. By Markov's inequality and Theorem 2.2.16 for any $C > 0$ it follows that for the infinite tree,

$$\mathbb{P}_\infty(Z_k \geqslant C(N-k)) \leqslant \frac{1+k\sigma^2}{C(N-k)}.$$

Since $k \leqslant \delta N$, we can rewrite the inequality as follows:

$$\mathbb{P}_\infty(Z_k \geqslant C(N-k)) < \frac{1+k\sigma^2}{C(N-k)} < \frac{1+\delta N\sigma^2}{CN(1-\delta)} = \\ = \frac{\delta\sigma^2}{C(1-\delta)} + \frac{1}{CN(1-\delta)}.$$

Then for any $\varepsilon_3 > 0$ and $\delta > 0$, when N is large enough, there exist $C > 0$ such that for all $k \leqslant \delta N$,

$$\mathbb{P}_\infty(Z_k \geqslant C(N-k)) < \varepsilon_3. \tag{2.40}$$

Denote the projection of measures $\mathbb{P}_N$ and $\mathbb{P}_\infty$ onto the trees with common first δN layers $\mathcal{T}[\delta N]$ as $\mathbb{P}_N\restriction_{\delta N}$ and $\mathbb{P}_\infty\restriction_{\delta N}$ respectively. We now prove that we can bound the total variation distance between the measures $\mathbb{P}_N\restriction_{\delta N}$ and $\mathbb{P}_\infty\restriction_{\delta N}$ is small. By definition,

$$d_{TV}\left(\mathbb{P}_N\restriction_{\delta N}, \mathbb{P}_\infty\restriction_{\delta N}\right) = \frac{1}{2}\sum_{t\in\{GW|Z_N>0\}} \left|\mathbb{P}_N\restriction_{\delta N}\left[\mathcal{T}[\delta N] = t\right] - \mathbb{P}_\infty\restriction_{\delta N}\left[\mathcal{T}^\infty[\delta N] = t\right]\right| = \\ = \frac{1}{2}\sum_{\substack{t\in\{GW|Z_N>0\}\\ \#t_{\delta N}<CN(1-\delta)}} |\dots| + \frac{1}{2}\sum_{\substack{t\in\{GW|Z_N>0\}\\ \#t_{\delta N}>CN(1-\delta)}} |\dots| = (I) + (II).$$

We bound separately both sums. It follows from the Equation (2.39) that

$$\left|\mathbb{P}_N\restriction_{\delta N}\left[\mathcal{T}[\delta N] = T\right] - \mathbb{P}_\infty\restriction_{\delta N}\left[\mathcal{T}^\infty[\delta N] = T\right]\right| \leqslant \varepsilon_0\frac{N}{N-\delta N}\mathbb{P}_\infty(\mathcal{T}[\delta N] = T).$$

Therefore, the first sum (I) can be bounded as follows:

$$\frac{1}{2}\sum_{\substack{t\in\{GW|Z_N>0\}\\ \#t_{\delta N}<CN(1-\delta)}} |\dots| \leqslant \frac{1}{2}\frac{\varepsilon_0}{1-\delta}\sum \#t_{\delta N}\mathbb{P}_\infty(\mathcal{T}[\delta N] = T) = \\ = \frac{1}{2}\frac{\varepsilon_0}{1-\delta}\mathbb{P}_\infty(Z_{\delta N} < CN(1-\delta)) < \\ < \frac{1}{2}\frac{\varepsilon_0}{1-\delta}.$$

Hence there exist $\varepsilon_4 > 0$ such that the sum $(I) < \varepsilon_4$. The second sum (II) can be bounded in the following way:

$$\frac{1}{2}\sum_{\substack{t\in\{GW|Z_N>0\}\\ \#t_{\delta N}>CN(1-\delta)}} |\dots| \leqslant \frac{1}{2}\sum_{\substack{t\in\{GW|Z_N>0\}\\ \#t_{\delta N}>CN(1-\delta)}} (\mathbb{P}_N(\mathcal{T}[\delta N] = t) + \mathbb{P}_\infty(\mathcal{T}^\infty[\delta N] = t)) \leqslant \\ \leqslant \frac{1}{2}\mathbb{P}_N(Z_{\delta N} \geqslant CN(1-\delta)) + \frac{1}{2}\mathbb{P}_\infty(Z_{\delta N} \geqslant CN(1-\delta)). \tag{2.41}$$

We have chosen δ and C in such way that the probability to have large Z_k in the infinite tree

is small. Therefore using Equations (2.39) and (2.40) we have the following:

$$\begin{aligned}\mathbb{P}_N(Z_{\delta N} \geqslant CN(1-\delta)) = \sum_{\substack{t\in\{GW|Z_N>0\}\\ \#t_{\delta N}\geqslant CN(1-\delta)}} \mathbb{P}_N(\mathcal{T}[\delta N]=t) \leqslant \sum (1+\varepsilon_0)\#t_{\delta N}\mathbb{P}(\mathcal{T}[\delta N]=t) = \\ = (1+\varepsilon_0)\sum \mathbb{P}_\infty(\mathcal{T}[\delta N]=t) = (1+\varepsilon_0)\mathbb{P}_\infty(Z_{\delta N}\geqslant CN(1-\delta)) < \\ < (1+\varepsilon_0)\varepsilon_3.\end{aligned} \tag{2.42}$$

Therefore there exist $\varepsilon_5 > 0$, such that

$$\mathbb{P}_N(Z_k \geqslant C(N-k)) \leqslant \varepsilon_5.$$

Hence putting (2.40) and (2.42) into (2.41), we obtain

$$\frac{1}{2}\sum_{\substack{t\in\{GW|Z_N>0\}\\ \#t_{\delta N}<CN(1-\delta)}} |\dots| < \frac{1}{2}(\varepsilon_5+\varepsilon_3).$$

Thus there exist $\varepsilon_6 > 0$, such that

$$\frac{1}{2}\sum_{\substack{t\in\{GW|Z_N>0\}\\ \#t_{\delta N}<CN(1-\delta)}} |\dots| < \varepsilon_6.$$

Summing everything up, we can write that for any $\varepsilon > 0$ and sufficiently large N we can find $\delta > 0$ such that the total variation distance between two measures is small:

$$d_{TV}\left(\mathbb{P}_N\restriction_{\delta N}, \mathbb{P}_\infty\restriction_{\delta N}\right) \leqslant \varepsilon.$$

Hence by Theorem 2.2.6 there exists a coupling of the random variables $\mathcal{T}[\delta N]$ and $\mathcal{T}^\infty[\delta N]$, such that the difference between them is small in total variation distance. This finishes the proof of the Lemma. □

2.5.2 Proof of Theorem 2.1.2

Let $\mathcal{T}_{+e}$ be the tree $\mathcal{T}$ with an extra edge attached to a root and one of the vertices of $\mathcal{T}$ uniformly at random. The following Theorem shows that on average there is a smoothing effect of an extra edge up to a size of a constant fraction of the tree $\mathcal{T}$.

By Lemma 2.5.3 for large N and any $\varepsilon' > 0$ we can find $\delta' > 0$ and a coupling of $\mathcal{T}$ and $\mathcal{T}^\infty$, such that the tree $\mathcal{T}[\delta' N]$ is not different from $\mathcal{T}^\infty[\delta' N]$ with probability $(1-\varepsilon')$.

Denote $n := \delta' N$. Let the part of the spine γ in the tree $\mathcal{T}[n]$ be denoted as γ_n. The number of offspings of each vertex i on γ_n is represented by shifted size-biased i.i.d. random variables

$$\eta_i = \hat{\xi}_i - 1,$$

where $0 \leqslant i \leqslant n$. Denote $\mu := \mathbb{E}\eta_1$, and it is easy to show that $\mu = \sigma^2 < \infty$. Let the total number of unconditional critical GW trees attached to the spine γ_n be denoted as $s_n := \sum_{j=0}^{n} \eta_j$ and label these trees as $\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_{s_n}$. Denote as $\mathcal{T}_0$ the tree $\mathcal{T}_0 = \mathcal{T} \setminus (\mathcal{T}_1 \cup \dots \cup \mathcal{T}_{s_n})$.

We now prove that the $\kappa(\mathcal{T})$ is tight. We can use Lemma 2.5.1 for the case when a cycle is just a vertex. Then the smoothing goes up to s_0 trees hanging from the root, and since $|\mathcal{T}_0|$ in this case is a.s. large it is true that

$$\kappa(\mathcal{T}) = \sum_{i=1}^{\eta} |\mathcal{T}_i|.$$

By Markov's inequality we have for $C > 0$

$$\mathbb{P}(\eta > C) \leqslant \frac{\sigma^2}{C}.$$

Therefore for any $\varepsilon_1' > 0$ there exist large enough $C > 0$, such that

$$\mathbb{P}(\eta < C) > 1 - \varepsilon_1'. \tag{2.43}$$

Hence with high probability the number of trees is bounded. Prove that with high probability also the total size of these trees is small. Then for $c > 0$ we have

$$\mathbb{P}\Big(\sum_{i=1}^{C} |\mathcal{T}_i| > cN^2\Big) < \mathbb{P}(|\mathcal{T}_1| > cN^2).$$

By Theorem 2.2.13 we can derive that for large N,

$$\begin{aligned}\mathbb{P}(|\mathcal{T}_1| > cN^2) &\sim \frac{1}{\sigma\sqrt{2\pi}} \sum_{n=cN^2} n^{-3/2} \sim \\ &\sim \frac{1}{\sigma\sqrt{2\pi}} \int_{cN^2}^{\infty} x^{-3/2} dx = \frac{1}{\sigma\sqrt{2N\pi c}}.\end{aligned}$$

Therefore we can write that for any $\varepsilon_2' > 0$ there exists $c > 0$ and large N such that

$$\mathbb{P}\Big(\sum_{i=1}^{C} |\mathcal{T}_i| < cN^2\Big) > 1 - \varepsilon_2'. \tag{2.44}$$

Thus the union of these two events hold with probability $1 - \varepsilon_1' - \varepsilon_2'$, which implies that $\kappa(\mathcal{T})$ is tight, which finishes the proof of the first statement of the Theorem.

Now we prove the second statement. Denote as $(e \in \mathcal{T}_0)$ the event that the extra edge e is attached to the tree $\mathcal{T}_0$. Then by Lemma 2.5.1, we have the following:

$$\kappa(\mathcal{T}_{+e}) = |\mathcal{T}| - \max_{i=0,\dots,s_n} |\mathcal{T}_i|.$$

The probability that the size of smoothing is proportional to the size of a tree $\mathcal{T}_{+e}$ can be bounded as follows:

$$\mathbb{P}\Big(\frac{\kappa(\mathcal{T}_{+e})}{|\mathcal{T}_{+e}|} > \delta\Big) = \mathbb{P}\Big(\frac{\kappa(\mathcal{T}_{+e})}{|\mathcal{T}_{+e}|} > \delta \Big| e \in \mathcal{T}_0\Big)\mathbb{P}(e \in \mathcal{T}_0) + \mathbb{P}\Big(\frac{\kappa(\mathcal{T}_{+e})}{|\mathcal{T}_{+e}|} > \delta \;\Big|\; e \notin \mathcal{T}_0\Big)\mathbb{P}(e \notin \mathcal{T}_0) \geqslant$$
$$\geqslant \mathbb{P}\Big(\frac{\kappa(\mathcal{T}_{+e})}{|\mathcal{T}_{+e}|} > \delta \;\Big|\; e \in \mathcal{T}_0\Big)\mathbb{P}(e \in \mathcal{T}_0),$$

for some $\delta > 0$. The probability of the edge e to be attached to the tree $\mathcal{T}_0$ can be bounded:

$$\mathbb{P}(e \in \mathcal{T}_0) \geqslant \mathbb{P}\Big(e \in \mathcal{T}_0 \;\Big|\; \frac{|\mathcal{T}[n]|}{|\mathcal{T}|} < \delta'\Big)\mathbb{P}\Big(\frac{|\mathcal{T}[n]|}{|\mathcal{T}|} < \delta'\Big). \tag{2.45}$$

Since the attachment is uniform, it would automatically follows that

$$\mathbb{P}\Big(e \in \mathcal{T}_0 \;\Big|\; \frac{|\mathcal{T}[n]|}{|\mathcal{T}|} < \delta'\Big) > 1 - \delta'.$$

We prove now that the size of the tree $\mathcal{T}[n]$ is small comparing to the total size of $\mathcal{T}$ with high probability. By the Law of Large Numbers for any $k > 0$ and large N we can find $\varepsilon_1 > 0$ such that:

$$\mathbb{P}(s_n \geqslant k\mu n) \leqslant \varepsilon_1.$$

Therefore the number of trees is bounded w.h.p. Consider $L > 0$, then using the law of total probability we have conditioned on $Z_N > 0$:

$$\begin{aligned}\mathbb{P}(|\mathcal{T}[n]| > Ln^2) &= \mathbb{P}(|\mathcal{T}[n]| > Ln^2 | s_n > k\mu n)\mathbb{P}(s_n > k\mu n) + \\ &+ \mathbb{P}(|\mathcal{T}[n]| > Ln^2 | s_n \leqslant k\mu n)\mathbb{P}(s_n \leqslant k\mu n) \leqslant \\ &\leqslant \mathbb{P}(s_n > k\mu n) + \mathbb{P}\Big(\sum_{i=1}^{k\mu n} |\mathcal{T}_i| > Ln^2\Big)\mathbb{P}(s_n \leqslant k\mu n) = \\ &= \varepsilon_1 + \mathbb{P}\Big(\sum_{i=1}^{k\mu n} |\mathcal{T}_i| > Ln^2\Big)\end{aligned} \tag{2.46}$$

Since the trees $\mathcal{T}_i$ are i.i.d. and by Theorem 2.2.13 there exist a slowly varying function ℓ, such that the tail distribution of $|\mathcal{T}_1|$ is written as follows:

$$\mathbb{P}(|\mathcal{T}_1| \geqslant x) = \ell(x)\frac{1}{\sigma\sqrt{2\pi}}\sum_{n=x}^{\infty} n^{-3/2} \sim$$
$$\sim \ell(x)c_1\int_x^{\infty} z^{-3/2}dz = \ell(x)c_2x^{-1/2}.$$

Hence, by Theorem 2.2.8 the size of the tree $|\mathcal{T}_1|$ is in the domain of attraction of the stable law Y with the following density:

$$p_Y(y) \sim \frac{1}{\sqrt{2\pi}}y^{-3/2}e^{-\frac{1}{2y}}.$$

Denote the sum of sizes of trees as $S_m = \sum_{i=1}^{m} |\mathcal{T}_i|$. Then by Lemma 2.2.2 we have that for large L:

$$\mathbb{P}\Big(\sum_{i=1}^{k\mu n} |T_i| > Ln^2\Big) = \mathbb{P}\Big(\frac{S_{k\mu n}}{(k\mu n^2)} > \frac{L}{(k\mu)^2}\Big) \sim \mathbb{P}\Big(Y > \frac{L}{(k\mu)^2}\Big) =$$
$$= \frac{1}{\sqrt{2\pi}} \int_{L/(k\mu)^2}^{\infty} y^{-3/2} e^{\frac{1}{2y}} dy = \frac{1}{\sqrt{2\pi}} \int_{0}^{(k\mu)^2/2L} z^{-1/2} e^{-z} dz =$$
$$= \frac{1}{\sqrt{2\pi}} \gamma\left(1/2, \frac{(k\mu)^2}{2L}\right) \sim \frac{k\mu}{2\sqrt{\pi}} L^{-1/2}.$$

Hence for any large L we can find $\varepsilon_2 > 0$ such that

$$\mathbb{P}\Big(\sum_{i=1}^{k\mu n} |T_i| > Ln^2\Big) < \varepsilon_2. \tag{2.47}$$

Using (2.47) in (2.46) we derive that for any $\varepsilon_3 > 0$ there exist large L, such that

$$\mathbb{P}(|\mathcal{T}[n]| < Ln^2) > 1 - \varepsilon_3. \tag{2.48}$$

Since the stable law has density around 0 being close to zero, we have for any $\varepsilon_4 > 0$ there exist small $\delta_2 > 0$ such that

$$\mathbb{P}(|\mathcal{T}[n]| > \delta_2 n^2) > 1 - \varepsilon_4. \tag{2.49}$$

We prove now that the total size of the tree $|\mathcal{T}|$ conditioned on $Z_N > 0$ is not large. By Bayes' formula:

$$\mathbb{P}(|\mathcal{T}| = M \mid Z_N > 0) = \frac{\mathbb{P}(Z_N > 0 \mid |\mathcal{T}| = M)}{\mathbb{P}(Z_N > 0)} \mathbb{P}(|\mathcal{T}| = M).$$

Since the condition $Z_N > 0$ implies that there exist at least one vertex at each distance k from the root, then for $k < N$,

$$\mathbb{P}(|\mathcal{T}| = k \mid Z_N > 0) = 0.$$

The event $(Z_N > 0 \big| |\mathcal{T}| = M)$ is equivalent to the event that $(H(\mathcal{T}) > N \big| |\mathcal{T}| = M)$, hence using Theorem 2.2.15, Theorem 2.2.14 and Theorem 2.2.13, for large N, we have that the following inequality holds for all possible M:

$$\mathbb{P}(|\mathcal{T}| = M \mid Z_N > 0) \leqslant C e^{-c_1 \frac{N^2}{M}} N M^{-3/2},$$

We now prove that the probability to have small $\mathcal{T}$ is small. Consider $\gamma > 0$, then we have

$$\mathbb{P}(|\mathcal{T}| < \gamma N^2 \mid Z_N > 0) \leqslant C \sum_{m < \gamma N^2} e^{-c_1 \frac{N^2}{m}} N m^{-3/2},$$

and we can bound the sum with an integral:

$$\mathbb{P}(|\mathcal{T}| < \gamma N^2 \mid Z_N > 0) \leqslant C \int_N^{\gamma N^2} e^{-c_1 \frac{N^2}{m}} N m^{-3/2} dm = \frac{C}{N^2} \int_{1/N}^{\gamma} e^{-\frac{c_1}{x}} x^{-3/2} dx =$$
$$= \frac{C}{\sqrt{c_1}} \int_{c_1 N}^{c_1/\gamma} e^{-y} y^{-1/2} dy < \frac{C}{\sqrt{c_1}} \int_{\infty}^{c_1/\gamma} e^{-y} y^{-1/2} dy <$$
$$< \frac{C}{\sqrt{c_1}} \Gamma\left(1/2, \frac{c_1}{\gamma}\right),$$

where $C, c_1 > 0$ are constants that depend on ξ. By Lemma 2.2.3 as $\gamma \to 0$,

$$\mathbb{P}(|\mathcal{T}| < \gamma N^2 \mid Z_N > 0) \leqslant C\gamma^{-1/2} e^{-c_1/\gamma}. \tag{2.50}$$

Then it follows from (2.50) that for any $\varepsilon_5 > 0$ there exist $\gamma > 0$, such that

$$\mathbb{P}(|\mathcal{T}| > \gamma N^2 \mid Z_N > 0) \geqslant 1 - \varepsilon_5. \tag{2.51}$$

Therefore from (2.48) and (2.51) it follows that

$$\mathbb{P}\Big(\frac{|\mathcal{T}[n]|}{|\mathcal{T}|} < \frac{Ln^2}{\gamma N^2}\Big) = \mathbb{P}\Big(\frac{|\mathcal{T}[n]|}{|\mathcal{T}|} < \frac{L\delta'^2}{\gamma}\Big) > 1 - \varepsilon_3 - \varepsilon_5,$$

and we can choose δ', such that $L\delta' = \gamma$. Then there exist $\varepsilon_6 > 0$, such that

$$\mathbb{P}\Big(\frac{|\mathcal{T}[n]|}{|\mathcal{T}|} < \delta'\Big) > 1 - \varepsilon_6. \tag{2.52}$$

Plugging (2.52) into the Equation (2.45) we obtain that there exist $\varepsilon_7 > 0$ such that

$$\mathbb{P}(e \in \mathcal{T}_0) > 1 - \varepsilon_7.$$

We now prove that the probability to have a large tree $\mathcal{T}$ is also small. Consider $K > 0$, then analogously we have:

$$\mathbb{P}(|\mathcal{T}| > KN^2 \mid Z_N > 0) \leqslant C \sum_{m > KN^2} e^{-c_1 \frac{N^2}{m}} N m^{-3/2}$$
$$\leqslant C \int_{KN^2}^{\infty} e^{-c_1 \frac{N^2}{m}} N m^{-3/2} dm = C \int_K^{\infty} e^{-\frac{c_1}{x}} x^{-3/2} dx =$$
$$= \frac{C}{\sqrt{c_1}} \int_0^{c_1/K} e^{-y} y^{-1/2} dy = \frac{C}{\sqrt{c_1}} \gamma\left(1/2, \frac{c_1}{K}\right)$$

and by Lemma 2.2.2 for K large,

$$\mathbb{P}(|\mathcal{T}| > KN^2 \mid Z_N > 0) \leqslant CK^{-1/2}.$$

Hence, for any $\varepsilon_8 > 0$ we can find $K > 0$, such that

$$\mathbb{P}(|\mathcal{T}| < KN^2 \mid Z_N > 0) > 1 - \varepsilon_8. \tag{2.53}$$

Therefore from (2.49) and (2.53) it follows that

$$\mathbb{P}\Big(\frac{|\mathcal{T}[n]|}{|\mathcal{T}|} > \frac{\delta_2 n^2}{KN^2}\Big) > 1 - \varepsilon_4 - \varepsilon_8.$$

Conditioned on the edge e to be attached to $\mathcal{T}_0$ we have by Lemma 2.5.1,

$$\kappa(\mathcal{T}_{+e}) > |\mathcal{T}[n]|.$$

Hence, we finally obtain that there exist $\varepsilon_9 > 0$, such that

$$\mathbb{P}\Big(\frac{\kappa(\mathcal{T}_{+e})}{|\mathcal{T}_{+e}|} > \frac{\delta_2 \delta'^2}{K}\Big| e \in \mathcal{T}_0\Big) > 1 - \varepsilon_9.$$

Hence finally we obtain that for any $\varepsilon > 0$ there exist $\delta > 0$, such that

$$\mathbb{P}\Big(\frac{\kappa(\mathcal{T}_{+e})}{|\mathcal{T}_{+e}|} > \delta\Big) > 1 - \varepsilon.$$

This finishes the proof of the Theorem. □

2.6 Critical versus near-critical Erdős-Rényi graphs and heuristics

Another model of random graphs where our results may be applied and the same smoothing phenomenon may be observed is the largest cluster of critical versus near-critical Erdős-Rényi graphs. We are presenting here the main ideas on a slightly heuristic level, and are planning to work out the details in the forthcoming paper version of this chapter of the thesis.

Let us consider the Erdős-Rényi graph $G(n,p)$ in its critical window for the emergence of a giant cluster, at $p = p_n(\lambda) = 1/n + \lambda/n^{4/3}$, with $\lambda \in (-\infty, \infty)$. We will use the standard monotone coupling of these random graphs.

The cluster of a typical vertex in $G(n, p_n(\lambda))$ in the critical case $\lambda = 0$ is locally a GW tree with Poisson offspring distribution with mean 1. As we are raising λ, extra edges appear in the standard coupling. Since the sizes of the largest clusters above the critical point are of order $n^{2/3}$, the number of extra edges in each such component is approximately Poisson $(\Theta(\lambda))$, while, of course, the extra edges also merge some of these components. That is, the large scale structure of large critical versus near-critical clusters resembles but does not

exactly coincide with our first example: a critical random tree conditioned to be large, plus a constant number of random edges.

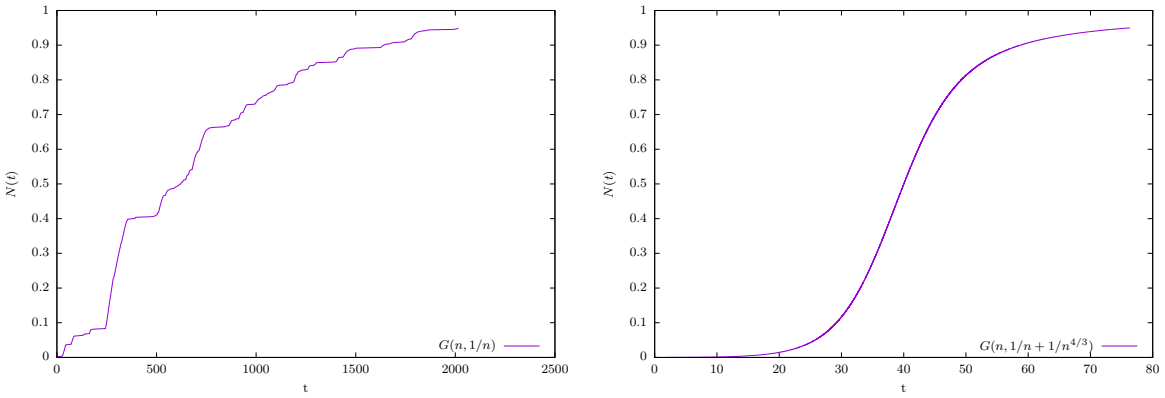
To make this picture more precise, the probability that the largest cluster of $G(n, p_n(\lambda))$ is a tree converges to some $t(\lambda) > 0$, which decays rapidly as $\lambda \to \infty$ (see [33]). This means that, for $\lambda = 0$, with a decent positive probability the SI spreading will encounter bottlenecks everywhere during the process, and, because of these bottlenecks at random locations, the averaged spreading curve will not converge and will produce jumps. However, at large $\lambda > 0$, a typical largest cluster will have one or more extra edges, hence, on a typical realization of the cluster, we expect to see the smoothing effect.

In order to apply our Theorem 2.4.6, we need to prove that $\kappa(G_p(n, p_n(\lambda)))$ is typically positive for these clusters. As before, this will be clear once we know that the cluster, having a volume of order $n^{2/3}$,

- has a diameter of order $n^{1/3}$,
- the extra edges are located quite randomly, so that the cycles created have length of order $n^{1/3}$, and
- the subgraphs hanging off from such a large cycle are random enough, with a critical GW-like structure, so that at least two of them have volume of order $n^{2/3}$.

Such a structure result is essentially provided by [1, 2], where the scaling limit of large near-critical clusters as metric spaces is described, building on classical work of Aldous [5]. The properties we are aiming at are probably most apparent from Procedure 1 in [2], p.9, which constructs the scaling limit of a large cluster conditioned to have volume $cn^{2/3}$ and $k \geq 0$ extra edges roughly as follows; the structure is basically independent of c. For $k \geq 2$, first one takes a 3-regular graph on $2(k-1)$ vertices, from a certain distribution; for $k = 0$, the graph is just a single vertex; for $k = 1$, the graph is a "lollipop": two vertices with an edge between them and a loop at one of the vertices. Then one replaces each edge of the graph by a copy of Aldous' Brownian Continuum Random Tree, independent apart from being conditioned to have total scaled volume c. Finally, one chooses the points where the neighbouring trees are glued to each other in an appropriate way. Since a positive distance in the scaling limit corresponds to a discrete distance of order $n^{1/3}$, for $k \geq 1$ our requirements above are easily seen to be satisfied.

The following results from our numerical simulations show that, even at a moderately off-critical value $\lambda = 1$, we practically always see the smoothing effect (see Figure 2.4).



(a) *(b)*

Figure 2.4. *Simulation of SI spreading with power-law inter-event times with* $\alpha = 0.8$ *on the Erdős-Rényi graph (ER) with* $n = 20000$ *vertices. (a) The critical ER graph with* $\lambda = 0$*. (b) The above critical ER graph with* $\lambda = 1$*.*

CHAPTER 3

The role of bridges in the SI spreading on temporal networks

3.1 Introduction

Recently much effort has been made to understand the main governing factors of spreading on networks [19,31,38]. Topological aspects like community structure [26] and temporal activity patterns like burstiness [9] severly influence the speed of spreading [32, 38, 49].

Different approaches have been applied to tackle this problem. One direction is to treat related models analytically and by numerical simulations [19, 31, 32, 34, 45, 67]. Another possibility is to use empirical data about the temporal networks and define an SI spreading process with those data. Such data include email logs [52,67] records of face to face encounters [59] or mobile call billing information [38, 49].

Mobile call records provide detailed insight into the dynamics of human interactions and can especially well be applied to study the different aspects of communications processes [15]. It was shown [54] that in mobile call networks subsribers from a city can be considered as a kind of community. We decided therefore to study how the communication and the spreading of information are structured in cities. In this Chapter we use time stamped mobile call data from a service provider of an European country for simulating the SI process on the real communication networks.

3.2 Basic notions

Throughout this Chapter we use the notion of networks, which is in a sense another name of graphs, used in computer science. Instead of the word "vertex" we use the word "node" and

edges are called sometimes "links". We use the notion of temporal networks as introduced in [55]. A *temporal network* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is defined as a set of nodes $\mathcal{V}$ between which a complete trace of all interaction events $\mathcal{E}$ occurring within the time interval $[0, T]$ is known. Each such event $\epsilon \in \mathcal{E}$ can be represented by a quadruplet $\epsilon = (u, v, t, \delta t)$, where the event connecting nodes u and v begins at t and the interaction is completed in time δt. As an example, δt may correspond to the duration of a phone conversation or flight time in an air transport network. Broadly, we define δt such that if an event e transmits something from u to v, the recipient can "make use" of it only after the time δt has elapsed. However, in some cases, events can be approximated as instantaneous so that $\delta t = 0$ and they can be represented with triplets $\epsilon = (u, v, t)$, as in [29]. In the current study we further use the latter notion.

The underlying or aggregate network $\Gamma = (V, E)$ of a temporal network is defined as the set of nodes $V(\Gamma) = \mathcal{V}$ and there exist a link between two nodes $u, v \in V(\Gamma)$ if there is at least one interaction $(u, v, t) \in \mathcal{E}$. Thus for each link $e = (u, v) \in E(\Gamma)$ one can project all interactions $(u, v, t) \in \mathcal{E}$ onto e and have all the information about the communication between nodes u and v directly on e. The collection of interactions on each link is called *interaction pattern*. The interaction pattern is *homogeneous* if the interaction events are uncorrelated in time and a *bursty* pattern is the pattern with significantly enhanced levels of activity followed by long periods of inactivity [28].

For any edge $(u, v) \in E(\Gamma)$, the time $\Delta t = t_2 - t_1$ between two consecutive interactions (u, v, t_1) and (u, v, t_2), where $t_1 < t_2$, is called an *inter-event time*. Once we have all inter-event times for edge e, we can define the inter-event time distribution $\mathbb{P}(\Delta t)$ for edge e. It is commonly used (see e.g. [28, 38]) that homogeneous interaction patterns are modelled by the exponential inter-event time distribution $\mathbb{P}(\xi > t) \sim e^{-\lambda t}$, where $\lambda > 0$, and bursty patterns are modelled by power-law tailed inter-event time distributions $\mathbb{P}(\xi > t) \sim t^{-\alpha}$, where $\alpha > 0$. The *average inter-event time* is the average of the inter-event times over the whole network. The average inter-event time gives an approximation of the common pattern of interactions in the network. We denote the mean of the inter-event distribution as μ and standard deviation as σ. The *burstiness coefficient* B measures how bursty the interaction is in the network and it can be in terms of inter-event time distribution:

$$B = \frac{\sigma - \mu}{\sigma + \mu}.$$

When $B \sim 1$ the interaction follows a bursty pattern, when $B \sim 0$ the pattern is homogeneous and resembles the one produced by the Poisson process and when $B \sim -1$ the interaction is

regular [28].

The underlying network Γ may be disconnected and consist of several subnetworks, or *components*, without links between them. We call the *largest connected component* $LCC(\Gamma)$ the component with the largest size. The quantity P_∞ denotes the proportion of nodes in $LCC(\Gamma)$ over the total size of the network.

We define a deterministic SI spreading process on the temporal network $\mathcal{G}$ as a dynamic process on the underlying network Γ. Let us remind its definition. Each node is labelled by one of two states: susceptible (S, non-informed) or infected (I, informed). Suppose that at time $t = 0$ all nodes are in the susceptible state and one node, the initiator, chosen at random, turns into the infected state. The infection is passed with time t through links from infected nodes to susceptible ones *instantaneously* at the first time when an interaction happens between these two nodes. The process runs until the connected component of the initiator becomes infected. Thus, in order to avoid uncertainties in the results we always consider spreading on the $LCC(\Gamma)$. One can define a spreading curve $N(t)$ as the ratio of nodes infected in the process by time t to the size of the $LCC(\Gamma)$. To obtain statistics, we perform in our simulations M runs of SI spreading process with random initiators from the $LCC(\Gamma)$. The average spreading curve is then defined as the average curve $\langle N(t)\rangle$:

$$\langle N(t)\rangle = \frac{\sum_{i=1}^{M} N_i(t)}{M}.$$

The *characteristic time* of the SI spreading on $\mathcal{G}$ is defined as the time τ, such that $\langle N(\tau)\rangle = 1/2$.

In the modelling Section we consider several network models. Three of them are basic network models: the Erdős-Rényi, the Barabási-Albert and the regular lattice model with next nearest neighbour interactions (NNN) and one is a model designed to reflect the community structure of the society. Let us describe them briefly here.

Two-dimensional regular lattice with NNN (Z) is widely used in theoretical studies of various physical models (see e.g. [62, 71]). The infinite lattice with NNN is defined as the infinite regular two-dimensional grid of nodes $[i, j]$, where $i, j \in \mathbb{Z}$, and nodes $[i_1, j_1]$ and $[i_2, j_2]$ are connected with an edge, if $|i_1 - i_2| + |j_1 - j_2| = 1$ or $|i_1 - i_2| + |j_1 - j_2| = 2$. The network Z is defined as an induced sub-network of an infinite on n^2 nodes $V = \{[i, j] : 1 \leqslant i, j \leqslant n\}$.

Erdős-Rényi random graph model (ER) is considered to be the first and mathematically simplest model of random graphs [23]. The model ER is constructed from an empty graph on n vertices by adding each possible edge with probability p.

Barabási-Albert network (BA) is a network model designed to model the preferential attachment of nodes [10]. The network BA starts with an initial connected network of m_0 nodes, where $m_0 > 0$. New nodes are added to the network one at a time. Each new node is connected to $m \leqslant m_0$ existing nodes with a probability that is proportional to the number of links that the existing nodes already have. Formally, the probability p_i that the new node is connected to node i is

$$p_i = \frac{k_i}{\sum_j k_j},$$

where k_i is the degree of node i and the sum is made over all existing nodes j.

The model of Kumpula et.al. (K) was proposed in [44] as a model of social networks with the community structure. Denote ξ to be the weight of links. The model starts with an empty network on n nodes and at each step the links are generated in the following two ways. If a node u has non-zero degree, we choose one of its neighbours v with probability proportional to the weight of the outgoing link $\xi(u, v)$. Then, if v has neighbours apart from u, say w, we choose it with probability proportional to the weight of its link $\xi(v, w)$. If there is no link between u and w we create it with probability p_Δ and attach initial weight δ_0, otherwise we add δ to the weights of all these three links. This mechanism is called local attachment (LA). Under the second mechanism, the node u creates a link with a random node v with weight δ_0 with some constant probability p_r, or with probability 1 if u has zero degree. This mechanism is called global attachment (GA). Finally any node can be replaced by a zero degree node with constant probability p_d, which is called node deletion (ND). These three mechanisms, i.e., LA, GA, and ND, are applied to all nodes at each step i and the network reaches a statistically stationary state after a sufficient number of steps D. We consider an unweighted version of this network, meaning that when using the network for the SI process we disregard the weights. The probability p_Δ is directly related to the average degree of the network.

The *average degree* of Γ is the average of node degrees over the node set. The network Γ on n nodes has a *small-world* property, if the diameter of network is at most of size $O(\log n)$ [69]. The network Γ is said to have *community structure*, if there is a partition of a node set into subsets with a high concentration of edges within subsets and low concentration between them [26]. We say that network model Γ with n nodes and m links has *non-zero percolation threshold* if deleting a sufficient proportion p of links, where $p < 1$, leaves the size of $LCC(\Gamma)$ to be of size $o(n)$ in the limit of $n \to \infty$ [39]. The important properties of all four above models are summarized in Table 3.1. The degree distributions of networks Z, ER and BA are

	Degree distribution $\mathbb{P}(k)$	Small-world	Non-zero percolation threshold	Communities
Z	~ 12	-	+	-
ER	$\sim Poi(np)$	+	+	-
BA	$\sim k^3$	+	-	-
K	$\ln\mathcal{N}(\mu,\sigma)$	+	+	+

Table 3.1. *Table of the main properties of the considered network models.*

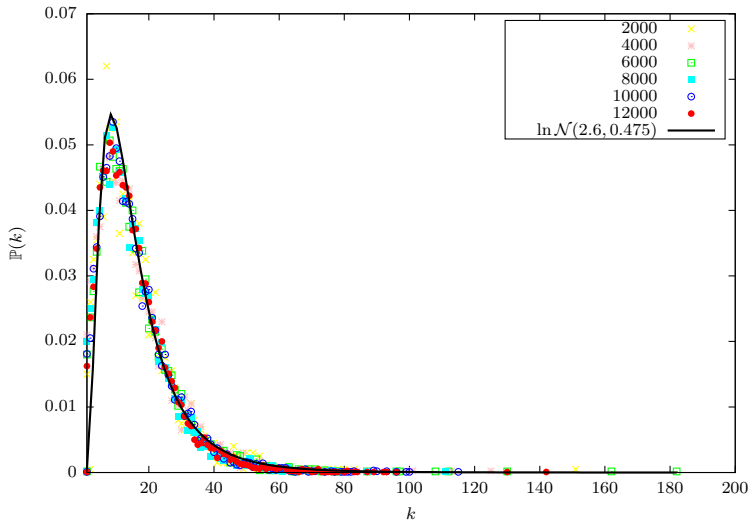


Figure 3.1. *The mass function of the degree distribution in networks K with all fixed parameters except the number of vertices. The fit function follows log-normal distribution with mean* 2.6 *and standard deviation* 0.475*.*

known, when the degree distribution of network K was measured empirically (see Figure 3.1).

3.3 Data and analysis

The dataset contains records of mobile phone calls (MPC) of one service provider of one European country. The call records contain the caller and callee's hashed codes, the starting time of the call and it's duration. The dataset spans over a five-month period from August to December of 2007 and contains $2,271,071,378$ call records between $53,955,465$ users. Among these users $9,769,376$ users have active contracts with the service provider during the observed period and we call them as *company users*. The rest of the users are called *non-company users*. Company users may have additional information about their location,

given by the ZIP code stated in the contract. The number of company users with the ZIP code is $6,272,586$, others have no declared ZIP code. The data contains two types of calls: 1) between two company users and 2) the calls between a company user and a non-company user. Since the data comes from one service provider, we have no information about the intitiated calls or geographical information for non-company users.

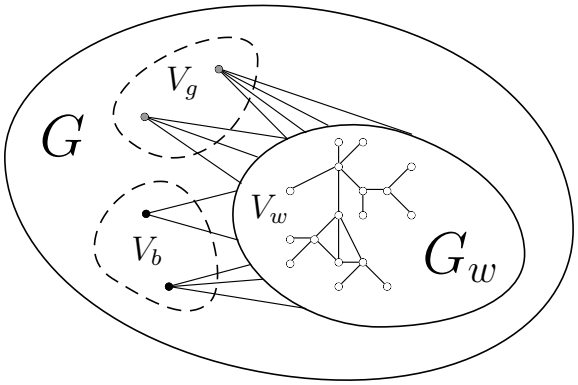


Figure 3.2. *Schematic structure of networks G_w and G. White nodes are denoted as V_w, grey nodes as V_g and black nodes as V_b.*

The MPC data generates a temporal network $\mathcal{G}$ with the node set of users and the set of events $\mathcal{E}$ of calls users. We call the nodes representing company users with known ZIP code as *white* nodes, the nodes representing company users without ZIP code as *grey* nodes and the nodes of non-company users as *black* nodes. We consider an SI spreading process on this network. It may happen that the data span is not enough to infect the whole network, thus periodic boundary conditions in time are applied for the data. We assume that when a call happens, spreading can go in both directions, from a callee to a caller or vice versa, thus, for the purpose of current research, the network is considered undirected. Since there is no data about calls between non-company users, there are no links between nodes representing these users.

In the current research we consider sub-networks of cities defined as follows. For each city the network $G_w = (V_w, E_w)$ is defined as the set of white nodes V_w with ZIP code matching one of ZIP codes of the city with the edge set E_w of calls between those users. The network G_w is an induced sub-network of Γ. The network $G = (V, E)$ is defined as an extension of network G_w with the set of grey nodes V_g and black nodes V_b, that have links to the nodes V_w. Along with the grey and black vertices we add edges connecting V_g and V_b to V_w. Clearly $G_w \subset G$ and $V(G) = V_w \cup V_g \cup V_b$ and by construction there are no edges within and between

vertices from V_g and V_b in G. Schematically the networks G_w and G can be represented in the Figure 3.2.

In the network G the grey and black nodes of degree $k \geqslant 2$ can be considered as *bridges* between white nodes. We compare the spreading on the set of white nodes without bridges (in G_w) and with bridges (in G). Therefore we perform the modification of network G by removing all grey and black vertices of degree 1. Then we consider the largest connected component $LCC(G_w)$ in network G_w and $LCC(G)$ in network G. We denote as $\langle k_w \rangle$ the average degree of white nodes.

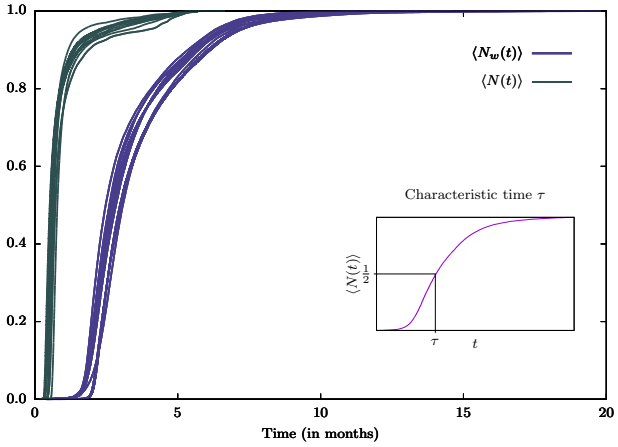


Figure 3.3. *Average spreading curves $\langle N_w(t) \rangle$ (dark blue) for 10 typical networks of cities G_w with average degree $\langle k_w \rangle = 4 \pm 0.2$. Respective average spreading curves $\langle N(t) \rangle$ for networks of same cities G are also given in dark grey. The inset illustrates the definition of the characteristic time τ.*

We study the behaviour of SI spreading on each network G_w and G using the average spreading curves $\langle N_w(t) \rangle$ and $\langle N(t) \rangle$. At each run i, where $i = 1, \ldots, M$, the initiator is chosen from the set of white nodes of $LCC(G_w)$ with degree at least 2 to avoid slowdown related to possible bottlenecks at the very beginning of the process. The characteristic time of SI spreading on networks G_w and G are denoted as τ_w and τ correspondingly. The sample of average spreading curves for cities with $\langle k_w \rangle = 4 \pm 0.2$ is given on Figure 3.3 along with the illustration of the notion of characteristic time.

3.4 Results

We consider the sample of all N cities with population between $10,000$ to $300,000$ people. Cities of smaller population are not considered because of small size of produced networks, and cities with larger population are underrepresented in the data. For each city i, where $i = 1, \ldots, N$, we construct networks $G_w(i)$ and $G(i)$. The size of networks G is proportional to the number of white nodes, however the dependence of the size on the population of city is vague, which is explained by differences in coverage of the company in different cities (see Figure 3.5 (a), (b)). We also observe the linear dependence between the number of white nodes $|V_w|$ and the total number of bridges $|V_g| + |V_b|$ in networks G (see Figure 3.6). Remember that P_∞ is the relative size of the $LCC(G_w)$ to the total size of G_w. Then on Figure 3.7, (b), we see that the dependence of $\langle k_w \rangle$ over the P_∞ shows the typical percolation behavior.

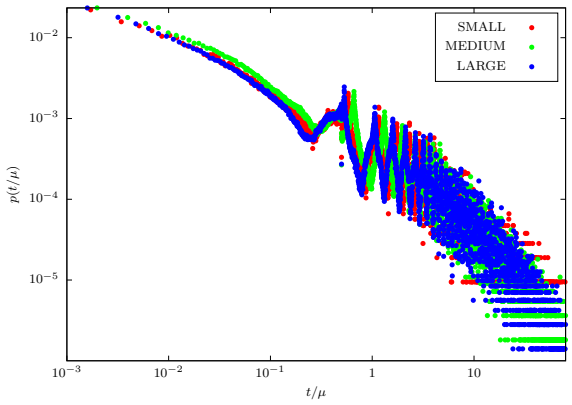


Figure 3.4. *The inter-event time distributions of three cities of different sizes of V_w. The small sized city is considered to have $|LCC(G_w)| \sim 10^3$, the medium sized city has $|LCC(G_w)| \sim 10^4$ and the large is with $|LCC(G_w)| \sim 10^5$.*

We denote the average inter-event time for each city i, where $1 \leqslant i \leqslant N$, as μ_i and the corresponding standard deviation is denoted as σ_i. These values are very similar for all cities and $\mu_i \in (0.04, 0.08)$ and $\sigma_i \in (0.1, 0.2)$ for almost all cities i. The burstiness coefficients B_i belong to the interval $(0.41, 0.46)$, indicating the bursty interaction patterns in all cities. The distribution of inter-event times in the cities is found to be similar for cities of different size (see Figure 3.4), which is explained as that people follow same patterns of calls irrespective of the size of their community.

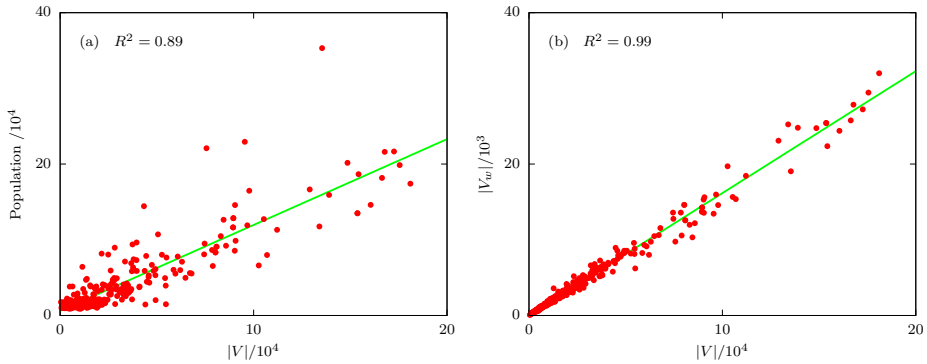


Figure 3.5. *(a) The relationship between scaled sizes of the networks G and the population of each city. The linear trend is vague, which represents the difference in coverage proportion of the company in each city. (b) Linear dependence between the size of network G and the number of white nodes $|V_w|$ for each city.*

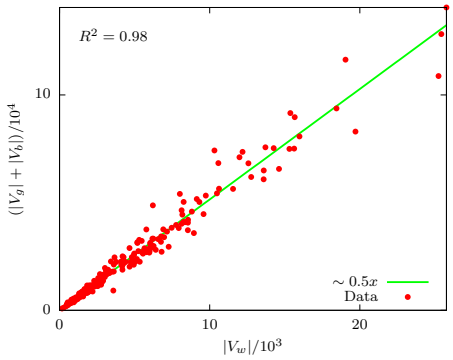


Figure 3.6. *Linear dependence between the total number of bridges $(|V_g| + |V_b|)/10^4$ and the number of white nodes $|V_w|/10^3$.*

We find the largest connected components $LCC(G_w)$ and $LCC(G)$ in each network G_w and G correspondingly. It is observed that there is no proper relation between population and the number of company users and cities with large population may produce sparse networks regardless of their population size.

We perform $M = 100$ runs of SI spreading process on each of the networks G_w and G. In agreement with our expectations we see a radical decrease in the characteristic times, when we switch from G_w to G (Figure 3.7). We observe when $P_\infty < 20\%$, the networks show scattered results in characteristic time. Thus we perform truncation of those cities and present the results on Figure 3.7 (b).

We find a clear power-law dependence of the characteristic time τ_w on the average degree

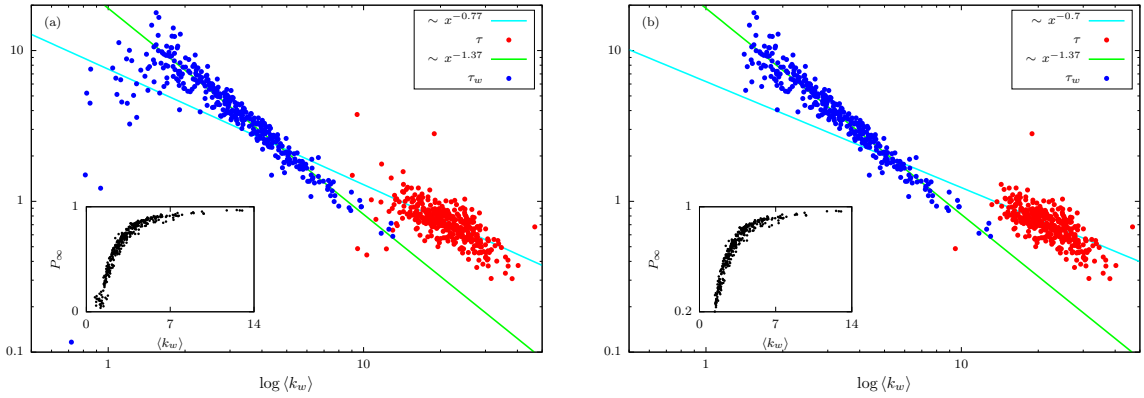


Figure 3.7. *The characteristic times τ_w (blue dots) and τ (red dots) relative to average degree of white nodes $\langle k_w \rangle$ in the networks G_w and G. (a) The whole sample of cities is plotted, with large scattering around $\langle k_w \rangle \sim 1$. In this case the networks become very sparse and the size of LCC is decreasing. (b) The same plot is obtained by deleting those cities with $P_\infty < 20\%$, which gives clearer results.*

$\langle k_w \rangle$ in networks G_w (see Figure 3.7). At the first sight it is surprising that $\langle k_w \rangle$ and not the size of the $LCC(G_w)$ controls the process. This may be explained by the small size effect of networks G_w, since most of the sizes of $LCC(G_w)$ belong to the interval $[10^2, 10^4]$.

The introduction of bridges of grey and black nodes drastically decreases the characteristic time. However, the dependence of the characteristic time τ on the average degree $\langle k_w \rangle$ in networks G is less clear (see Figure 3.7). The reason the characteristic time is so much decreased ($\tau \in [0.5; 1.5]$) is that spreading curves start to look almost the same for all values of $\langle k_w \rangle$, thus smearing the topology of the underlying network. In the next Section we model this effect on empirical networks.

3.5 Modeling

3.5.1 Model

We construct a simple model of networks G_w and G that replicates the observed behaviour of the SI spreading process. Our model is based on the observation that considered networks G_w only give a partial insight into the whole network of calls in the city, thus not all connections between nodes are revealed. We also observe from the data that the number of white nodes is proportional to the total number of bridges (see Figure 3.6).

Let us describe the model. Consider a network $\overline{G} = (\overline{V}, \overline{E})$ on n nodes represented by a model with a parameter p_d responsible for the average degree. We define a *diluted* network $\overline{G}_w^p$ as the initial network $\overline{G}$ with each link being deleted with probability p, where $0 \leqslant p \leqslant 1$. Let the nodes of the network $\overline{G}_w^p$ be white nodes. For each $\overline{G}_w^p$ we define network $\overline{G}^p$ by adding the set of bridge nodes B to $\overline{G}_w^p$ and connecting each bridge node with exactly two white nodes at random. We model the interaction patterns by attaching i.i.d. random variables ξ_i to the links, which represent the inter-event times. We consider two types of inter-event time distributions: the power-law distribution $pow(t_{min}, \alpha)$ with the following density function:

$$p_{pow}(t) = \begin{cases} t_{min}^{\alpha} \alpha \frac{1}{t^{\alpha+1}}, & \text{if } t_{min} \leqslant t; \\ 0, & \text{otherwise,} \end{cases} \tag{3.1}$$

and the shifted exponential distribution $exp(t_{min}, \lambda)$ with the following density function:

$$p_{exp}(t) = \begin{cases} \frac{1}{\lambda - t_{min}} \exp\left(-\frac{t - t_{min}}{\lambda - t_{min}}\right), & \text{if } t_{min} \leqslant t; \\ 0, & \text{otherwise,} \end{cases} \tag{3.2}$$

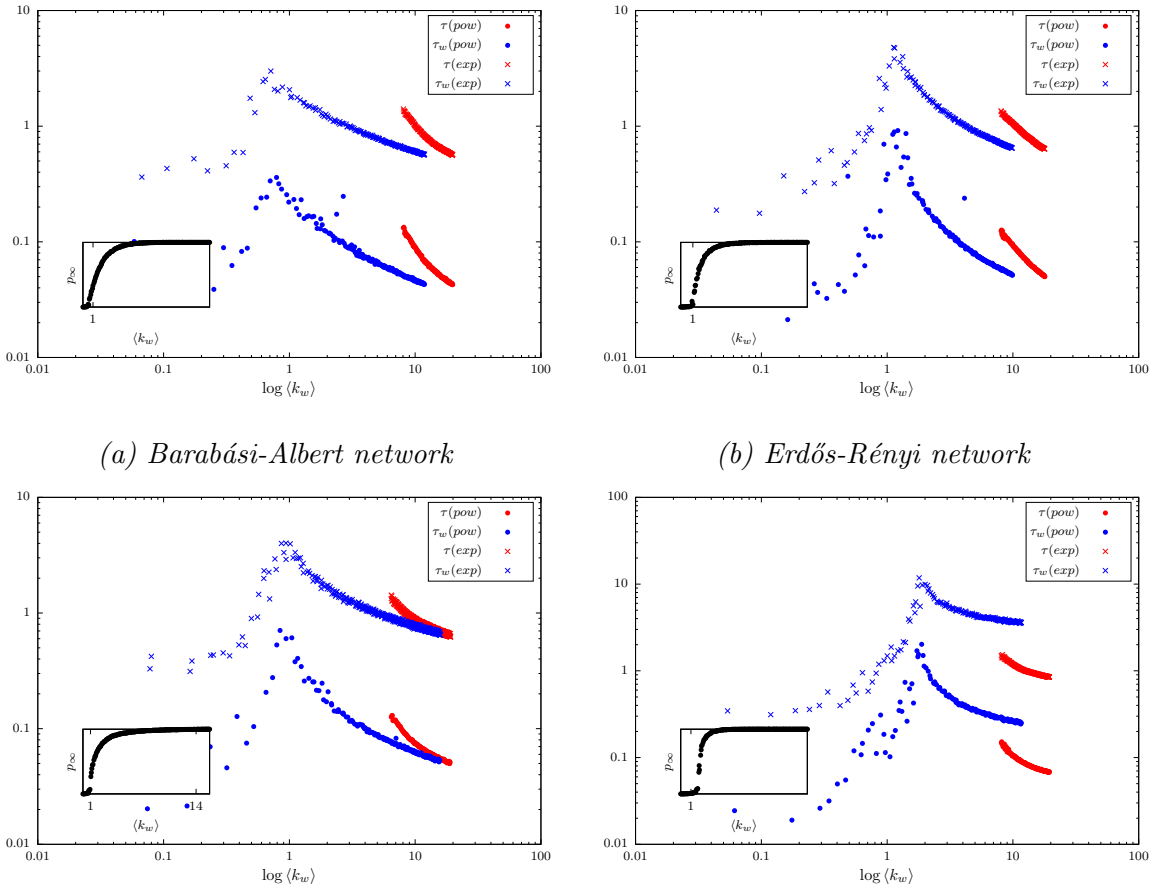
The distribution $\exp(t_{min}, \lambda)$ is defined such that it has the same support as $pow(t_{min}, \alpha)$.

3.5.2 Simulation

The simulation is performed on networks $\overline{G}_w^p$ and $\overline{G}^p$ constructed from the network $\overline{G}$, given by four network models defined earlier: 1) Regular lattice with NNN (Z); 2) Erdős-Rényi (ER); 3) Barabási-Albert (BA); 4) the model of Kumpula et.al. (K). Each considered network model has $n = 5000$ nodes (network Z has 4900 nodes) and the parameters of these models are tuned in such way that the average degree in the original $\overline{G}$ is around 12. The link deletion probability p goes from 0 to 1 with increment 0.05. The number of bridges added to each network is $5|V_w|$, which is obtained from the data.

We perform two sets of spreading runs, with $M = 100$ runs in each set: 1) with power-law inter-event times distribution with $t_{min} = 0.008$ and $\alpha = 1.2$ and 2) with exponential distribution with $t_{min} = 0.008$ and the parameter λ is chosen in the way that two distributions have same mean $\mu = 0.05$ obtained from the data. The results are presented on the Figure 3.8. All four network models show faster spreading with bursty interaction patterns modelled by the power-law inter-event times ($\tau_w(pow)$ and $\tau(pow)$) compared to the Poisson process ($\tau_w(exp)$ and $\tau(exp)$). Introduction of bridges decreases characteristic time in all models, which agrees with the empirical results in Section 3.4.

On Figure 3.9 we have the comparison between four models with power-law inter-event times. We observe that BA model is the fastest in terms of characteristic time, which is explained by the presence of large degree nodes and since the percolation threshold is zero, these nodes remain in the diluted network as well. It is known that community structure slows down spreading (see e.g. [70]). Therefore in the model K the spreading is slower than ER when $\langle k_w \rangle$ is large. In the diluted network with small average degree $\langle k_w \rangle$ the LCC has no apparent community structure and since it has more large degree nodes, it overtakes ER in spreading speed. The model Z is slow due to its regular structure. The introduction of bridges smears the topology of initial networks and thus the spreading results in the almost the same behaviour for all four models.



(a) Barabási-Albert network

(b) Erdős-Rényi network

(c) Network of Kumpula et.al.

(d) Regular lattice with NNN

Figure 3.8. *Characteristic times τ_w and τ relative to average degree of white nodes $\langle k_w \rangle$ in the simulated networks $\overline{G}_w^p$ and $\overline{G}^p$ with four different initial networks $\overline{G}$. Each network has 5000 nodes and 25000 bridges are added with both power-law (•) and exponential inter-event times (×) are considered. The following models are considered: (a) The Barabási-Albert network; (b) Erdős-Rényi network; (c) Network of Kumpula et.al; (d) Regular lattice with NNN interactions. The results show that Poisson process is slower on each model. Introduction of bridges decreases characteristic time and the results depend on the initial topology of the network.*

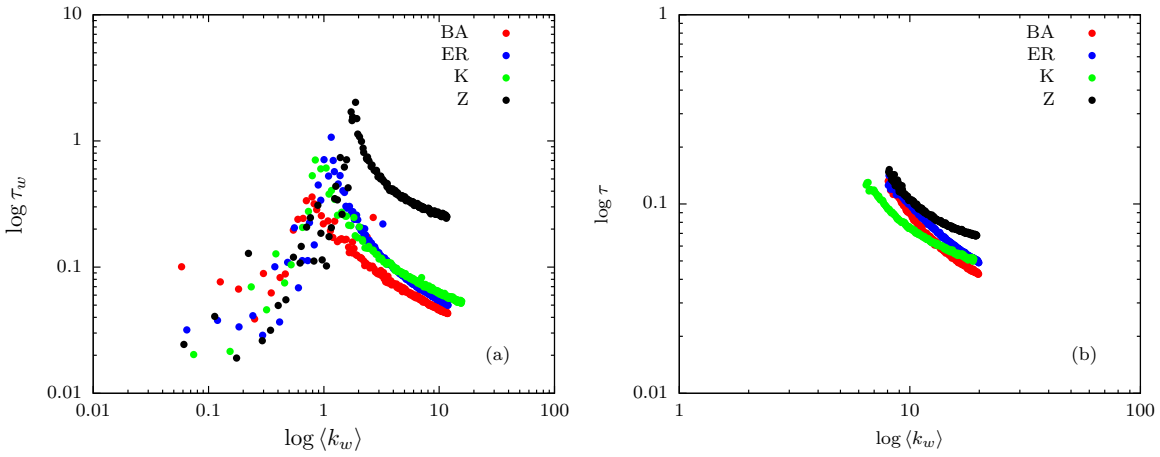


Figure 3.9. *Comparison of characteristic times τ_w and τ relative to average degree of white nodes $\langle k_w \rangle$ in the simulated networks $\overline{G}^p_w$ and $\overline{G}^p$ with four different initial networks $\overline{G}$. Each network has 5000 nodes and 25000 bridges are added with power-law distributed inter-event times. The results suppose that the degree distribution, or mainly the presence of large degree nodes, is a key factor governing the speed of spreading in these models. However, if these nodes belong to communities, a slowdown happens. Introduction of bridges decrease characteristic time and smears topological differences.*

CHAPTER 4

Discussion

In the last Chapter we have studied the role of bridges in the SI spreading on real and simulated networks. We have confirmed that the introduction of sufficient number of bridges decreases the characteristic time of the process, thus making spreading faster. The effect of bridges is more vivid when the network has low average degree. In this case the LCC has low number of cycles and the network locally has structure of a tree. In Chapter 2 we have derived that addition of one bridge to a tree drastically increases the average time to infect k nodes in the SI process. The real networks in our dataset show similar effect after introducing a sufficient number of bridges.

We have constructed a simple model that replicates the observed effect of bridges. Based on this model, we have found evidence that topology of the underlying network has effect in spreading. According to our expectations the presence of large degree nodes and absence of communities increases speed of SI spreading. When the average degree of the network is less than one, it falls apart into the number of disconnected components, each of which having a tree-like structure. Introduction of bridges makes this network connected, but the topology becomes homogeneous irrespective of the original topology and spreading curves look almost the same for different network models.

In the Chapter 2 we have shown how the addition of one edge can drastically increase average speed of spreading. Moreover, the more extra edges one adds, the more speed-up one can obtain. However, the speed-up can be achieved if these extra edges introduce large cycles into the structure of network., because it is more probable that the short cycles be contained inside the communities, which can cause a slowdown of spreading. From the proof of the Theorem 2.3.2 we can derive that large degree nodes are also important in speeding up the process, since on an n-star the first $n/2$ get occupied in $O(\log(n))$ time.

In the Chapter 2 we have studied the case of smoothing when $\alpha \in (1/2, 1)$. We acknowledge that the analogous result can be derived for any $\alpha \in (1/(d+1), 1/d]$, where $d > 1$. It is straightforward to show the analogue of Lemma 2.4.3 for one new and $d-1$ old edges and obtain

$$\mathbb{E}(\min(X, Y_1 - t, \ldots, Y_{d-1} - t) | Y_1, \ldots, Y_{d-1} > t) \leqslant C_d t^{1-\alpha}.$$

Then stochastically dominating the original process by the delayed process with d active edges every time in the front of epidemic and then constructing the appropriate process Q we can prove that the expected time to infect k vertices is bounded for $k < \kappa_d$. However, this means we would approximate the original process by spreading on the d infinite rays with common root. This construction is more complex and it is hard to find an example of tree-like graphs where such κ_d is not tight.

Bibliography

[1] L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. *Probability Theory and Related Fields*, 152(3):367–406, 2010.

[2] L. Addario-Berry, N. Broutin, and C. Goldschmidt. Critical random graphs: Limiting constructions and distributional properties. *Electron. J. Probab.*, 15:no. 25, 741–775, 2010.

[3] L. Addario-Berry, L. Devroye, and S. Janson. Sub-gaussian tail bounds for the width and height of conditioned galton–watson trees. *Ann. Probab.*, 41(2):1072–1087, 2013.

[4] R. Albert and A. lászló Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys*, page 2002.

[5] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *The Annals of Probability*, 25(2):812–854, 1997.

[6] D. Aldous and R. Lyons. Processes on unimodular random networks. *Electron. J. Probab.*, 12:no. 54, 1454–1508, 2007.

[7] J. Balogh and B. G. Pittel. Bootstrap percolation on the random regular graph. *Random Structures & Algorithms*, 30(1-2):257–286, 2007.

[8] S. Bansal, J. Read, B. Pourbohloul, and L. Meyers. The dynamic nature of contact networks in infectious disease epidemiology. *Journal of Biological Dynamics*, 4(5):478–489, 9 2010.

[9] A.-L. Barabasi. The origin of bursts and heavy tails in human dynamics. *Nature*, 435:207, 2005.

[10] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[11] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:no. 23, 1–13, 2001.

[12] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi. Asymptotic behavior and distributional limits of preferential attachment graphs. *Ann. Probab.*, 42(1):1–40, 01 2014.

[13] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra. First passage percolation on random graphs with finite mean degrees. *Ann. Appl. Probab.*, 20(5):1907–1965, 10 2010.

[14] S. Bhamidi, R. Van Der Hofstad, and J. Komjáthy. The front of the epidemic spread and first passage percolation. *Journal of Applied Probability*, 51A:101–121, 12 2014.

[15] V. D. Blondel, A. Decuyper, and G. Krings. A survey of results on mobile phone datasets analysis. *EPJ Data Science*, 4(1), Aug. 2015.

[16] G. Bobashev, R. J. Morris, and D. M. Goedecke. Sampling for global epidemic models and the topology of an international airport network. *PLoS ONE*, 3(9):1–8, 09 2008.

[17] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Struct. Algorithms*, 31(1):3–122, Aug. 2007.

[18] B. Bollobás, K. Gunderson, C. Holmgren, S. Janson, and M. Przykucki. Bootstrap percolation on galton-watson trees. *Electron. J. Probab.*, 19:no. 13, 1–27, 2014.

[19] J. Delvenne, R. Lambiotte, and L. Correa da Rocha. Diffusion on networked systems is a question of time or structure. *Nature Communications*, 6, 6 2015.

[20] J.-C. Delvenne, R. Lambiotte, and L. E. C. Rocha. Diffusion on networked systems is a question of time or structure. *Nature Communications*, 6:7366+, June 2015.

[21] R. Durrett. *Probability: Theory and examples.* CAMBRIDGE U PRESS, 2011.

[22] M. Eckhoff, J. Goodman, R. Hofstad, and F. R. Nardi. Short paths for first passage percolation on the complete graph. *Journal of Statistical Physics*, 151(6):1056–1088, 2013.

[23] P. Erdös and A. Rényi. On random graphs i. *Publicationes Mathematicae Debrecen*, 6:290, 1959.

[24] W. Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York, 1971.

[25] J. A. Fill and R. Pemantle. Percolation, first-passage percolation and covering times for richardson's model on the n-cube. *Ann. Appl. Probab.*, 3(2):593–629, 05 1993.

[26] S. Fortunato. Community detection in graphs. *Physics Reports*, 486(3–5):75 – 174, 2010.

[27] A. Ganesh, L. Massoulie, and D. Towsley. The effect of network topology on the spread of epidemics. In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, volume 2, pages 1455–1466 vol. 2, March 2005.

[28] K.-I. Goh and A.-L. Barabási. Burstiness and memory in complex systems. *EPL (Europhysics Letters)*, 81(4):48002, 2008.

[29] P. Holme. Network reachability of real-world contact sequences. *Phys. Rev. E*, 71:046119, Apr 2005.

[30] P. Holme and J. Saramäki. Temporal networks. *Physics Reports*, 519(3):97 – 125, 2012. Temporal Networks.

[31] D. X. Horváth and J. Kertész. Spreading dynamics on networks: the role of burstiness, topology and non-stationarity. *New Journal of Physics*, 16(7):073037, 2014.

[32] J. L. Iribarren and E. Moro. Impact of human activity patterns on the dynamics of information diffusion. *Phys. Rev. Lett.*, 103:038702, Jul 2009.

[33] S. Janson, D. E. Knuth, T. Łuczak, and B. Pittel. The birth of the giant component. *Random Structures & Algorithms*, 4(3):233–358, 1993.

[34] H.-H. Jo, J. I. Perotti, K. Kaski, and J. Kertész. Analytically solvable model of spreading dynamics with non-poissonian processes. *Phys. Rev. X*, 4:011041, Mar 2014.

[35] M. Karsai, G. Iñiguez, K. Kaski, and J. Kertész. Complex contagion process in spreading of online innovation. *Journal of The Royal Society Interface*, 11(101), 2014.

[36] M. Karsai, K. Kaski, A.-L. Barabási, and J. Kertész. Universal features of correlated bursty behaviours. *Scientific Reports*, 2:397, 2012.

[37] M. Karsai, K. Kaski, and J. Kertész. Correlated dynamics in egocentric communication networks. *PLoS ONE*, 7(7):1–9, 07 2012.

[38] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, and J. Saramäki. Small but slow world: How network topology and burstiness slow down spreading. *Phys. Rev. E*, 83:025102, Feb 2011.

[39] H. Kesten. *Percolation theory for mathematicians*, volume 2 of *Progress in Probability and Statistics*. Birkhäuser Boston, Mass., 1982.

[40] H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques*, 22(4):425–487, 1986.

[41] H. Kesten, P. Ney, and F. Spitzer. The galton-watson process with mean one and finite variance. *Theory of Probability & Its Applications*, 11(4):513–540, 1966.

[42] V. F. Kolchin. *Random mappings*. Optimization Software Inc., 1986.

[43] L. Kovanen, M. Karsai, K. Kaski, J. Kertész, and J. Saramäki. Temporal motifs in time-dependent networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(11):P11005, 2011.

[44] J. M. Kumpula, J.-P. Onnela, J. Saramäki, K. Kaski, and J. Kertész. Emergence of communities in weighted networks. *Phys. Rev. Lett.*, 99:228701, Nov 2007.

[45] D. Liben-Nowell and J. Kleinberg. Tracing information flow on a global scale using internet chain-letter data. *Proceedings of the National Academy of Sciences*, 105(12):4633–4638, 2008.

[46] T. Lindvall. On strassen's theorem on stochastic domination. *Electron. Commun. Probab.*, 4:no. 7, 51–59, 1999.

[47] N. Masuda and P. Holme. Predicting and controlling infectious disease epidemics using temporal networks - f1000prime biology reports - f1000. *F1000Prime Biology Reports*, (5), 2013.

[48] A. Meir and J. W. Moon. On the altitude of nodes in random trees. *Canadian Journal of Mathematics*, 30:997–1015, 1978.

[49] G. Miritello, E. Moro, and R. Lara. Dynamical strength of social ties in information spreading. *Phys. Rev. E*, 83:045102, Apr 2011.

[50] M. Newman. *Networks: An Introduction*. Oxford University Press, Inc., New York, NY, USA, 2010.

[51] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[52] M. E. J. Newman, S. Forrest, and J. Balthrop. Email networks and the spread of computer viruses. *Phys. Rev. E*, 66:035101, Sep 2002.

[53] F. W. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark. *NIST Handbook of Mathematical Functions*. Cambridge University Press, New York, NY, USA, 1st edition, 2010.

[54] R. K. Pan, M. Kivelä, J. Saramäki, K. Kaski, and J. Kertész. Using explosive percolation in analysis of real-world networks. *Phys. Rev. E*, 83:046112, Apr 2011.

[55] R. K. Pan and J. Saramäki. Path lengths, correlations, and centrality in temporal networks. *Phys. Rev. E*, 84:016105, Jul 2011.

[56] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87:925–979, Aug 2015.

[57] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 86:3200–3203, Apr 2001.

[58] A. Rudas, B. Tóth, and B. Valkó. Random trees and general branching processes. *Random Structures & Algorithms*, 31(2):186–202, 2007.

[59] J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, V. Colizza, L. Isella, C. Régis, J.-F. Pinton, N. Khanafer, W. Van den Broeck, and P. Vanhems. Simulation of an seir infectious disease model on the dynamic contact network of conference attendees. *BMC Medicine*, 9(1):1–15, 2011.

[60] C. Stone. A local limit theorem for nonlattice multi-dimensional distribution functions. *Ann. Math. Statist.*, 36(2):546–551, 04 1965.

[61] V. Strassen. The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, 36(2):423–439, 1965.

[62] V. E. Tarasov. Lattice model with nearest-neighbor and next-nearest-neighbor interactions for gradient elasticity. *Discontinuity, Nonlinearity and Complexity*, 4(1):11–23, 2015.

[63] S. Vajna, B. Tóth, and J. Kertész. Modelling power-law distributed interevent times. *New J. Phys.*, 15(arXiv:1211.1175):103023, Nov 2012. Comments: 5 figures.

[64] R. van der Hofstad. *Lecture Notes Random Graphs and Complex Networks*. in preparation (accessed 18.03.2016), 2016.

[65] P. Van Mieghem. The viral conductance of a network. *Comput. Commun.*, 35(12):1494–1506, July 2012.

[66] P. Van Mieghem, J. Omic, and R. Kooij. Virus spread in networks. *Networking, IEEE/ACM Transactions on*, 17(1):1–14, 2009.

[67] A. Vazquez, B. Rácz, A. Lukács, and A.-L. Barabási. Impact of non-poissonian activity patterns on spreading processes. *Phys. Rev. Lett.*, 98:158702, Apr 2007.

[68] A. Vespignani. Modelling dynamical processes in complex socio-technical systems. *Nature Physics*, 8:32–39, 2012.

[69] D. J. Watts and S. H. Strogatz. Collective dynamics of'small-world'networks. *Nature*, 393(6684):409–10, 1998.

[70] X. Wu and Z. Liu. How community structure influences epidemic spread in social networks. *Physica A: Statistical Mechanics and its Applications*, 387(2–3):623 – 630, 2008.

[71] H. J. W. Zandvliet. The 2d ising square lattice with nearest- and next-nearest-neighbor interactions. *EPL (Europhysics Letters)*, 73(5):747, 2006.