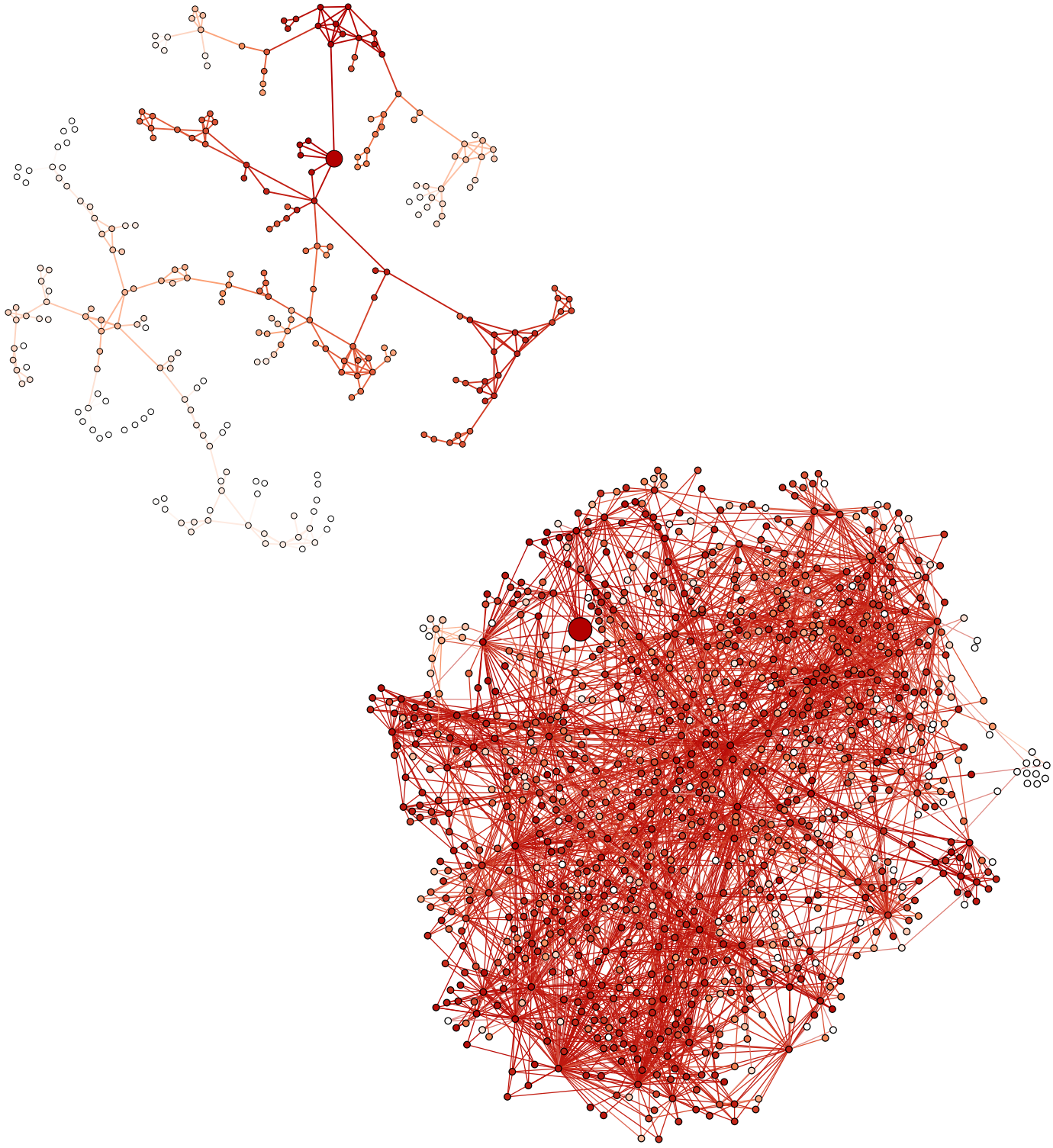
The role of the topology in non-Poissonian spreading dynamics on temporal networks



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

October, 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 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 go to my mom and my brother for their love, enormous support and encouragement during these four years of PhD studies. I would like to thank all 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, she 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 a 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', influence the model of non-Poissonian SI spreading on deterministic, 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 expectation. We derive the results both for general graphs and for specific models of random graphs. We study how the introduction of just one extra edge to a tree 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 spreading and define a model which replicates this phenomenon and compare results of this model on various theoretical networks. We finish the thesis with the discussion.

Contents

Contents i

1 Introduction 1

1.1 Networks as complex systems . . . 1

1.2 The development of spreading models . . . 2

1.3 Studies of spreading on temporal networks . . . 4

1.4 Mathematical studies of spreading . . . 5

2 The role of extra edges in speeding up spreading on graphs 9

2.1 Motivation . . . 9

2.2 Basic definitions and technical lemmas . . . 13

2.2.1 Graph-theoretic notions . . . 13

2.2.2 Functional notions . . . 14

2.2.3 Probabilistic notions . . . 16

2.2.4 Critical Galton-Watson trees . . . 20

2.2.5 SI spreading process . . . 22

2.3 Examples . . . 22

2.3.1 Spreading on a cycle . . . 23

2.3.2 Spreading on a star . . . 24

2.4 General deterministic graphs . . . 25

2.5 Critical Galton-Watson tree with an extra edge . . . 31

2.5.1 Technical lemmas . . . 32

2.5.2 Main Theorem . . . 36

2.6 Critical versus near-critical Erdős-Rényi graphs and heuristics . . . 43

3 The role of bridges in spreading on temporal networks 47

3.1 Introduction . . . 47

3.2 Basic notions 47
3.3 Data and analysis 51
3.4 Empirical results 53
3.5 Modeling 56
3.5.1 Model 56
3.5.2 Simulation 57

4 Discussion **61**

Bibliography **68**

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 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 [58], disease over a flight network [16] or innovation in the online social network [36].

Traditionally the study of networks has been the territory of graph theory, that has emerged from the works of Leonhard Euler around 200 years ago. 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 approximation character of this model was clear from the beginning, it dominated the representation of complex networks for decades. The drastically increasing amount of

data on complex systems had such an effect that interest 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 describe the underlying organizing principles in quantitative terms [4, 52]. Simple models were invented, like the Barabási-Albert or the Watts-Strogatz models, in order to capture important quantitative features of empirical networks [52]. Even those features that are vaguely defined, such as community structure, can be replicated in the network models [45].

Networks are not static in several senses. 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 [57,70] 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 [52].

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 recover from an infection and becomes insusceptible to the infection. Among other 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 *mean-field 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 of the infection spread of an individual per unit time. The infection can be passed from an infected to a susceptible person, and since there are at each time $s(t)x(t)$ such pairs of people, then the total rate of spreading in the system is $\beta s(t)x(t)$. Since the system has finite size, then we have at each time $s(t) = 1 - x(t)$ and the process is described by the following system of equations:

$$\frac{dx}{dt} = \beta x(1-x), \tag{1.1}$$
$$s + x = 1$$

The differential equation (1.1) occurs in many places in biology, physics, and elsewhere, and is called the *logistic growth equation*. It can be solved using standard methods to give

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

In reality the structure of networks is far from the mean-field, or the 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 at time t and let $x_i(t) = 1 - s_i(t)$ 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. In order to catch an infection between times t and $t + dt$, the vertex i must be susceptible in the first place and must have an infected neighbor that transmits the infection with probability βdt. Then the describing system of differential equations is

$$\frac{dx_i}{dt} = \beta s_i \sum_j A_{ij} x_j,$$
$$s_i(t) + x_i(t) = 1.$$

Despite the fact this model is the simplest one, it cannot be solved for arbitrary graphs in an exact way [51]. Therefore, when studying the spreading processes one should use various kinds 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 the Poisson process. That means we take a network on N nodes and assume that transmission happens 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 [68].

However, it has been shown empirically that spreading models with this Poissonian approach again produce results far from reality [69]. 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, 56].

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). One standard approach to analyze these data is to aggregate time stamps over time and construct 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 [69], 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 the results obtained from empirical case with those reference networks [39]. The latter are obtained by randomizing the initial temporal one, preserving some properties of the original network. For example, the 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 [48]? 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 [39]. Another simulation for SIS model on artificial temporal network data led to similar results [50]. On the other hand, in a version of the SI model, in which multiple infection attempts within a short time is necessary for a susceptible individual to be infected, the epidemic spreading is 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 [50]. 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 [37,38]. It has been shown for SIR model that temporal correlation of events occurring on links tends to enhance the spreading process [50]. 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 [44]. 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 [65]. 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, 60, 66], 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).

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].

The first passage percolation problem is a mathematical analogue of the notion of the SI epidemic spreading on networks. It has been shown that the behavior of the first passage percolation on sparse random graphs is universal across a lot of models [13]. The model of the first passage percolation has also been studied on different types of static regular structures, such as the hypercube and the 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 to use the general type branching process [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 general, this approach to studying the spreading processes fail when topology includes cycles.

In a number of works the influence of the topological properties on the spreading models is studied [20, 27, 67]. In these works the authors mainly provide 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 expectation. We derive the results both for deterministic graphs and for specific models of random graphs. We study how the introduction of just one extra edge to a tree 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 phenomenon and compare results of this model on various theoretical networks.

CHAPTER 2

The role of extra edges in speeding up spreading on graphs

2.1 Motivation

In the current Chapter we analyze the behavior of SI spreading on tree-like networks, where transmission times have power-law distribution with infinite expectation, which means the assumption that the transmission time of an infection from an infected vertex to a susceptible one through an edge connecting them is a random variable ξ with tail distribution $\mathbb{P}(\xi > t) \sim t^{-\alpha}$, where $0 < \alpha < 1$.

The main motivation of this research comes from the question, posed in the work of J. Kertesz and D.X. Horvath [31]. The authors considered computer simulations of SI spreading model on a graph with n vertices and the distinguished root vertex s, which acts as the initiator. In computer simulations, the behavior of the SI spreading with one initiator s is analyzed by assigning i.i.d. random passage times to each edge of the graph and performing a realization of these times, then calculating shortest weighted paths between the root vertex s and each of other vertices in the graph. Such procedure is also called a *run* and the outcome of the run is recorded as the *spreading curve*, which is the collection of points $(T_k, k/n)$ joined by lines in consecutive order, where T_k is the time when k vertices are infected in the process and where $1 \leqslant k \leqslant n$. After performing M runs of simulations, the statistical average of all collected spreading curves is taken over the first coordinate, thus obtaining the *average spreading curve*, which is the set of points

$$\left(\langle T_k \rangle, k/n\right),$$

joined by lines in consecutive order, where $\langle T_k \rangle = \frac{1}{M} \sum_{i=1}^{M} T_k^{(i)}$ and $T_k^{(i)}$ denotes the time to

infect k vertices in the i'th spreading curve. The curve represents the increasing function and reflects the typical spreading behavior in the simulation.

In the paper [31] the authors have found the presence of "uncontrolled (horizontal) jumps" in the average spreading curves of the computer simulated model of SI spreading on the networks of trees – large distances between two consecutive points of the average spreading curve that do not decrease with increasing the number of runs. The jump is represented by the following difference

$$\frac{1}{M}\left(\sum_{i=1}^{M} T_{k+1}^{(i)} - \sum_{i=1}^{M} T_k^{(i)}\right),$$

for some $k < n$. The uncontrolled jumps are explained by the presence of so-called temporal "bottlenecks", which are large passage times that occur on some particular edges. In the current work we investigate the time when the first jump appears. Indeed, it is clear that if the theoretical expectation $\mathbb{E}(T_k)$ is infinite, then by the Law of Large Numbers (LLN), for any $K > 0$ there exists $m > 0$, such that for all $M > m$ we have

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

Therefore, the first jump on the curve is expected for such k for which $\mathbb{E}(T_k) < \infty$ and $\mathbb{E}(T_{k+1}) = \infty$. The next jump occurs as if we started the process anew after the first bottleneck, without the already infected part of the graph. This (and all the later jumps) can also be investigated with the current method.

J. Kertesz and D.X. Horvath have made a similar hypothesis, based on the calculation of the expected first infection time, which is the time when a first neighbor of the root s is infected (notionally it is the second infection in the system, but we let the infection of the root happen at the beginning of the observation, i.e. at $t = 0$). Denote the degree of s as d_s, then the expected time of the first infection is defined as

$$\mathbb{E}(T_2) = \mathbb{E}(\min\{X_1, \ldots, X_{d_s}\}),$$

where $X_1, \ldots, X_{d_s}$ are the random passage times attached to the edges, incident to s. The authors have considered the passage times having power-law distribution ξ and have shown, that when the parameter $\alpha < 1/d_s$, the $\mathbb{E}(T_2)$ is infinite, and, thus, there should be a jump in the beginning of the curve.

In the current work we study the case of SI spreading with i.i.d. passage times having power-law distribution with $\alpha \in (1/2, 1)$, and show that the spreading curve has no uncontrolled jumps whenever the process has two edges to pass the infection. Consider the

simulation of the SI process on a cycle C_n with $n = 1000$ vertices and power-law passage times ξ having tail $\mathbb{P}(\xi > t) \sim t^{-\alpha}$ with $\alpha = 0.8$. We observe 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 $1 \leqslant k \leqslant n$.

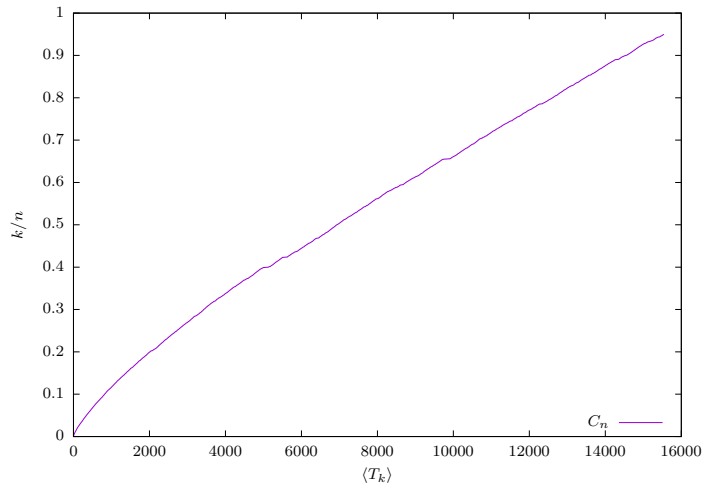


Figure 2.1: *Simulation of SI spreading with power-law weights ξ with tail $\mathbb{P}(\xi > t) \sim t^{-\alpha}$ with $\alpha = 0.8$ on the cycle C_n with $n = 1000$ vertices.*

The reason is that at each time the process on the cycle has two active edges that transmit an infection and for $\alpha \leqslant 1/2$ it is easy to see that the expected first infection time is the minimum of two i.i.d. weights and, thus, $\mathbb{E}(T_2) = \infty$, therefore, $\alpha = 1/2$ is a threshold.

A very different graph that exhibits a similar behavior is the star graph: an infected root with n neighbors. Here, for for $\alpha > 1/2$ the expected time to infect $n - 1$ neighbors of the root is also of order $n^{1/\alpha}$, which we establish in the Section 2.3.2. In this case the slow-down comes from the old edges, since for ξ having power-law distribution we have

$$\mathbb{P}(\xi > t + s \mid \xi > s) > \mathbb{P}(\xi > t),$$

for almost all s.

Motivated by these two "extreme" constructions, in Section 2.4 we derive that for any finite connected graph G with the root s there exists a number $\kappa(G, s)$, such that the average spreading curve of the SI process with power-law passage times with $\alpha \in (1/2, 1)$ has no uncontrolled jumps before $\kappa(G, s)$ vertices are infected. The number $\kappa(G, s)$ identifies the place

where the first temporal bottleneck appears and we show it has graph-theoretical meaning:

$$\kappa(G,s) = \min_{e \in E(G)} |\mathcal{C}(s, G\backslash e)|,$$

where $|\mathcal{C}(s, G\backslash e)|$ is the size of the connected component of the root s after deleting the edge e. The following general result is proven in Section 2.4.

Theorem 2.1.1. *Consider the graph G with the 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 exists the number $\kappa(G,s)$ such that for each k, where $1 \leqslant k \leqslant \kappa(G,s)$, the expected time to infect k vertices is bounded by*

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

and for $k > \kappa(G,s)$, the expectation $\mathbb{E}(T_k) = \infty$.

We have performed the simulation of the same SI spreading process with power-law passage times with $\alpha = 0.8$ on the critical Galton-Watson (CGW) tree with $N = 1000$ vertices and one initially infected root s and found that the jumps start to occur almost from the beginning of the process (see Figure 2.2). However, if we perform the same simulation on the same tree with an extra edge attached to the root and a randomly picked vertex, we observe there appear no jumps up to substantial fraction of the total size of the tree. This phenomenon is explained with mathematical rigor in the following theorem in Section 2.2.4.

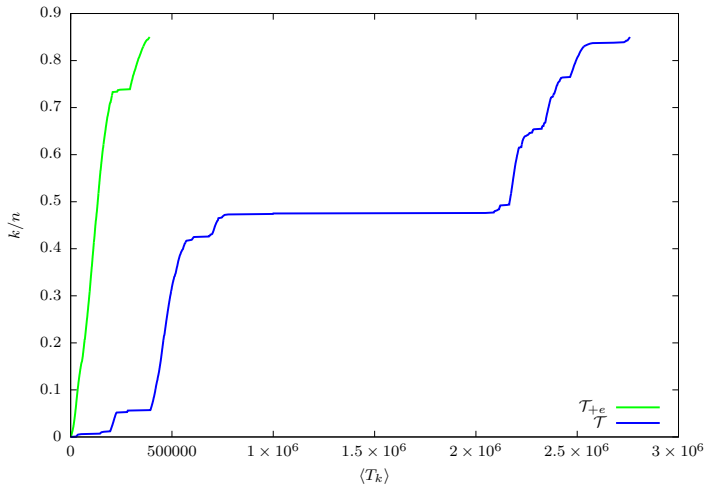


Figure 2.2: *Simulation of SI spreading with power-law weights with $\alpha = 0.8$ on the critical Galton-Watson tree $\mathcal{T}$ with $n = 1000$ vertices and the tree, denoted as $\mathcal{T}^e$, which is the same tree $\mathcal{T}$ with extra edge attached to the root and a uniformly chosen vertex.*

Theorem 2.1.2. *Consider the CGW tree $\mathcal{T}^N$ with one root s conditioned to have height at least $N > 0$ and the SI spreading process with power-law weights with $\alpha \in (1/2, 1)$. Denote as $\mathcal{T}^N_{+e}$ the tree $\mathcal{T}^N$ with an extra edge attached to the root and one of the vertices of $\mathcal{T}^N$ chosen uniformly at random, and $|\mathcal{T}^N|$ denotes the number of vertices in $\mathcal{T}^N$. Then as $N \to \infty$,*

- *the sequence of r.v. $\kappa(\mathcal{T}^N, s)$ is tight;*
- *for any $\varepsilon > 0$ there exists $\delta > 0$, such that*

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

The first statement of theorem says that w.h.p. there is an uncontrolled jump on the average spreading curve of the SI spreading on the CGW tree slightly after the beginning of the process. The second statement says that addition of just one more edge eliminates uncontrolled jumps on the average spreading curve up to a positive fraction of the graph and this fraction is occupied in finite expected time. This phenomenon we call the *smoothing of the spreading curve*, meaning that the small change in the graph structure eliminates uncontrolled jumps up to a comparatively large part of the average spreading curve.

2.2 Basic definitions and technical lemmas

In the following Section we establish some theoretical notions that are further used in the Chapter.

2.2.1 Graph-theoretic notions

In the following Chapter we always consider simple rooted graphs $G = (V, E, s)$ on n vertices with m edges and call the root of the graph G the distinguished vertex $s \in V$. We denote the set of vertices of graph G as V or $V(G)$ and the set of edges as E or $E(G)$. Denote $|G|$ to be the total size of the vertex set, or $|G| = |V|$. Let $s, t \in V$, then define a *(simple) path* between vertices s and t of length ℓ as the sequence of vertices $(v_0, v_1, \ldots, v_\ell)$, such that $v_0 = s$, $v_\ell = t$ and each $v_i \neq v_j$, for $0 \leqslant i \neq j \leqslant \ell$. We denote the 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.

Define the *weight* of an edge $e \in E$ as the value of function $\xi(e)$, where $\xi : E \to \mathbb{R}^+$. The function ξ is called a *weight function* or, simply, *weights*. In the following Chapter we consider weighted rooted graphs, where the weights are represented by i.i.d. random variables.

The graph $H = (V', E', s)$ obtained by selecting the vertex subset $V' \subset V$ along with the edges $E' \subset E$ connecting them and the same root s is called an *induced rooted subgraph* of G and denoted as $H \sqsubset G$. We call graph G a *cycle* if it is a path $(v_0, \ldots, v_\ell)$, such that $v_0 = v_\ell$. 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 $|E| = n - 1$.

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.

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

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

and $\Gamma(s, x)$ denote 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 $\Gamma(s, x)$ [54].

Lemma 2.2.1. *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 2.2.2. *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 2.2.3. *Let $f(x)$ be a positive continuous monotonically increasing function. Then,*

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

When $f(x)$ is positive continuous monotonically decreasing, then

$$\sum_{k=1}^{n} f(k) \leqslant \int_0^n f(x) dx.$$

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

$$b_{n+1} \leqslant b_n + Cb_n^{1-\alpha}, \tag{2.1}$$

with the initial condition

$$b_1 = d = (\alpha C)^{1/\alpha}. \tag{2.2}$$

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

Lemma 2.2.4. *Consider the sequence $(b_n)_{n=1}^{\infty}$ defined in (2.1), (2.2). 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 some $n > 1$ 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 (2.1) as

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

Making a telescopic sum, we have

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

Then by the induction hypothesis,

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

and by Lemma 2.2.3 we may bound the sum with an integral and obtain:

$$\begin{aligned} b_{n+1} - b_1 &\leqslant \int_1^{n+1} Cd^{1-\alpha}x^{1/\alpha-1}dx = \alpha Cd^{1-\alpha}\left((n+1)^{1/\alpha} - 1\right) \\ &= d\left((n+1)^{1/\alpha} - 1\right). \end{aligned} \tag{2.3}$$

Since b_1 is equal to d, then we can add it to both parts of (2.3) 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 a cdf of an absolutely continuous random variable ξ as $F(t) := \mathbb{P}(\xi < t)$ and a pdf as $p(t)$. We use the letter μ to denote the expected value of ξ and σ^2 for the variance $Var(\xi)$. We say a random variable ξ has power-law distribution $pow(t_{min}, \alpha)$, where $t_{min}, \alpha > 0$, if the density function of ξ is the following:

$$p(t) = \begin{cases} t_{min}^{\alpha} \alpha \frac{1}{t^{\alpha+1}}, & \text{if } t_{min} \leqslant t; \\ 0, & \text{otherwise.} \end{cases} \tag{2.4}$$

Further in the chapter we refer to a power-law distributed random variable as the one with $t_{min} = 1$, or having tail distribution $\mathbb{P}(\xi > t) = t^{-\alpha}$, 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 the 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 [63] for the original version; we use the formulation as in [66], p.59).

Theorem 2.2.5. *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 [47].

Theorem 2.2.6. *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*

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

The following lemma is technical and is used in the proof of Theorem 2.4.5.

Lemma 2.2.7. *Let X and Y be i.i.d. power-law distributed random variables with $\alpha \in (1/2, 1)$. Then, for any $t > 1$:*

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

with the constant factors depending on α.

Proof. The conditional tail distribution of the minimum of considered random variables is the following:

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

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

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

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 straightforward to show that there exists $C_1 > 0$, such that

$$f(t) \asymp C_1 \tag{2.5}$$

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

$$t^{1-\alpha}\int_{1/t}^{\infty}(u\,(1+u))^{-\alpha}\,du = t^{1-\alpha}\left[\int_{1/t}^{1}(u\,(1+u))^{-\alpha}\,du + \int_{1}^{\infty}(u\,(1+u))^{-\alpha}\,du\right]. \tag{2.6}$$

The first integral on the r.h.s. of (2.6) can be bounded in the following way:

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

or, equivalently,

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

The second integral on the r.h.s. of (2.6) can be bounded in the following way:

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

or, equivalently,

$$\int_{1}^{\infty}(u\,(1+u))^{-\alpha}\,du \asymp \int_{1}^{\infty}u^{-2\alpha}du = \frac{1}{2\alpha-1}.$$

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

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

Summarizing the above calculations we obtain

$$\mathbb{E}(\min\{X, Y-t\}|Y>t) \asymp t\frac{1}{1-\alpha}\left[\left(1+\frac{1}{t}\right)^{1-\alpha}-1\right] + t^{1-\alpha}\frac{1}{2\alpha-1} + \frac{t^{1-\alpha}-1}{1-\alpha}. \tag{2.7}$$

Using (2.5) we write the following upper bound:

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

with the constant factors depending on α. This finishes proof of the lemma. □

Consider the sequence of i.i.d. power-law distributed random variables $X_1, X_2, \dots$ with $0 < \alpha < 1$. Then the following limit theorem holds for their sum (see [21], p.138-142).

Theorem 2.2.8. *Let $X_1, X_2, \dots, X_n$ are i.i.d. random variables such that $\mathbb{P}(X_1 > x) \sim cx^{-\alpha}$ with $0 < \alpha < 2$ and $c > 0$. Denote $S_n = X_1 + \cdots + X_k$ and let*

$$a_k = k^{1/\alpha}, b_k = ck, \quad \text{if } 1 < \alpha < 2;$$
$$a_k = k^{1/\alpha}, b_k = 0, \quad \text{if } 0 < \alpha < 1.$$

Then as $k \to \infty$:

$$\frac{S_k - b_k}{a_k} \xrightarrow{d} Y,$$

where Y has a nondegenerate distribution.

In the framework of the thesis we do not consider the case $\alpha = 1$, thus it was omited from the theorem. Along with the Theorem 2.2.8 the local limit analogue holds [62].

Theorem 2.2.9. *Let $X_1, X_2, \dots$ be the sequence of i.i.d. random variables such that $\mathbb{P}(X_1 > x) \sim x^{-\alpha}$ with $0 < \alpha < 2$ and $c > 0$. Denote $S_n = X_1 + \cdots + X_k$ and Y be the limiting distribution of $(S_k - b_k)/a_k$. Then for any $x \in supp(Y)$ and $h > 0$*

$$\mathbb{P}\Big(\frac{S_k - b_k}{a_k} \in (x, x+h)\Big) = \mathbb{P}\big(Y \in (x, x+h)\big) + o(1)\left(h + \frac{1}{a_k}\right), \tag{2.8}$$

where $o(1)$ is given for $k \to \infty$ and uniform in h.

The limiting distribution Y is called *stable law* and it has the same parameter α as X_1. The stable law is only given via a characteristic function ([21], p.141) and it is impossible to write its density in most cases. However, it is known that 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 $(0, \infty)$, and therefore, has density $f_Y(t)$.*

The following lemma describes the approximation of the tail behaviour of the distribution Y for any $0 < \alpha < 2$ [24].

Lemma 2.2.11. *Let Y be the stable law with parameter $0 < \alpha < 2$. Then as $t \to \infty$ the probability density function of Y is*

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

The particular example of the stable law Y with the exponent $\alpha = 1/2$ can be expressed by the following density (see [21], p.141, Eq.(3.7.12)):

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

2.2.4 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 system of particles. The process starts with one particle, or $Z_0 = 1$. Then this particle splits into k offspring particles with probability p_k and they constitute the generation Z_1. Then each of these offsprings (should there be any) also has children with the same offspring distribution $\langle p_k : k \geqslant 0\rangle$, independently of each other and of their parent. The process continues forever or until there are no children born.

In order to give a formal definition, let $\xi_1^{(j)}, \xi_2^{(j)}, \ldots$, where $j \in \mathbb{N}$ denote i.i.d. non-negative integer distributed random variables with distribution ξ, where $\mathbb{P}(\xi = k) = p_k$. 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. These geneological trees are called as *Galton-Watson (GW) trees*. These trees can be viewed as rooted labeled trees. The root of the tree $\mathcal{T}$ corresponds to the particle in k'th generation, and it is denoted by $< 0 >$. A generic particle of the generation Z_k is indexed as $< 0, l_1, \ldots, l_k >$, where $l_r \geqslant 1$, $1 \leqslant r \leqslant k$. The particles $< 0, l_1, \ldots, l_{k-1}, j >$, where $j = 1, 2, \ldots$, denote the children of the particle $< 0, l_1, \ldots, l_{k-1} >$ in generation $k-1$. Of course not for all j does $< 0, l_1, \ldots, l_{k-1}, j >$ correspond to an actual vertex of $\mathcal{T}$. Let $N(0, l_1, \ldots, l_{k-1})$ be the number of children of $< 0, l_1, \ldots, l_{k-1} >$ in the branching process. Then $< 0, l_1, \ldots, l_{k-1}, j >$ is a vertex of $\mathcal{T}$ for $1 \leqslant j \leqslant N(0, l_1, \ldots, l_{k-1})$.

Denote the set of all GW trees as $\langle GW\rangle$ and a randomly chosen GW tree as $\mathcal{T}$. Let the root of the tree $\mathcal{T}$ be the particle in generation Z_0. A *critical GW (CGW) tree* is the genealogical tree of a critical GW process. The size of $\mathcal{T}$ is defined as the number of vertices it contains and is denoted as $|\mathcal{T}|$. It is well known that a CGW tree is almost surely finite (e.g. Theorem 3.1, p.84, [66]) and the following theorem provides a bound on the size of a CGW tree [43].

Theorem 2.2.12. *Let $\mathcal{T}$ be a CGW tree with integer offspring distribution ξ, such that*

$Var(\xi) := \sigma^2 < \infty$. *Then for* $n \to \infty$,

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

The *height* $H(\mathcal{T})$ of a GW tree $\mathcal{T}$ is the length of the longest path from the root in the tree or the maximum N, such that $(Z_N > 0)$. The following limit theorem about the height of the tree $\mathcal{T}$ holds [42].

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

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

The following theorem provides an upper bound on the probability of having a tree of height at least N conditioned on the exact size of this tree [3], [43].

Theorem 2.2.14. *Let* $\mathcal{T}$ *be a CGW tree with offspring distribution* ξ*, such that* $Var(\xi) := \sigma^2 < \infty$. *Then there exist positive 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 [49].

Theorem 2.2.15. *Let* $\mathcal{T}$ *be a critical GW tree with offspring distribution* ξ*, such that* $Var(\xi) := \sigma^2 < \infty$. *Then we have:*

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

Kesten in [41] proved that the conditional distribution of trees, conditioned on $Z_N > 0$, converges in distribution to an infinite CGW tree $\mathcal{T}^\infty \in \langle GW^\infty \rangle$, where $\langle GW^\infty \rangle$ is the set of *infinite* CGW trees which are the geneological trees of a critical Galton-Watson process conditioned on *non-extinction*. The infinite tree has the following construction. The 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 a number of offsprings according to the size-biased distribution $\hat{\xi}$, where

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

and $k = 0, 1, 2, \ldots$. Every offspring of a normal vertex is 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 tree $\mathcal{T}^\infty$ is to start by taking an infinite path γ of special vertices from the root, which is called a *spine*, and then attach $\nu = \hat{\xi} - 1$ independent CGW trees at each node of the spine. 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.

2.2.5 SI spreading process

We define the SI spreading process $T = (T_k)_{k=1}^n$ as a stochastic process on the finite rooted graph $G = (V, E, s)$ on n vertices, where to each $e \in E$ we attach an i.i.d. random weight $\xi(e)$ with 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 *active* 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 *unoccupied* otherwise. An infection is transmitted along the active edges from infected vertices to susceptible ones. The weight $\xi(e)$ of an edge $e = (u, v)$ is a passage time of an infection.

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 transmission of an infection as a flow from infected vertices to susceptible ones through edges of lengths ξ. 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 active edges at rate 1. The process runs until all vertices turn into infected state.

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

2.3 Examples

In the following subsections we consider two major extreme constructions defined in Section 2.1 and present a bound on the expected time to infect k vertices in the SI process with weights having power-law distribution $pow(\alpha)$ with $\alpha \in (1/2, 1)$.

2.3.1 Spreading on a cycle

Consider the example of SI spreading with power-law weights on the graph of a cycle C_n with n vertices. The spreading on a cycle can be well approximated by the graph of the doubly infinite line, denoted as $G = (V, E, s)$, which is defined by the vertex set of integers $V = \{0, \pm 1, \pm 2, \dots\}$ with the root $s = 0$ 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 X_i with $\alpha \in (1/2, 1)$, where i is a label of the larger vertex if one of the endpoints of the edge is positive, and is the label of a smaller vertex otherwise.

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

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

where constant factors depend on α.

Proof. Let $S_k = \sum_{i=1}^{k} X_i$ and $S_k^* = \sum_{i=-1}^{-k} X_i$. Note that,

$$\min\{S_{k/2}, S_{k/2}^*\} \leqslant T_k \leqslant \min\{S_k, S_k^*\}.$$

Then it is enough to prove that $\mathbb{E}(\min\{S_k, S_k^*\}) \asymp k^{1/\alpha}$. By Theorem 2.2.8 the sum S_k as $k \to \infty$ is in the domain of attraction of the stable law Y with the same parameter α:

$$\mathbb{P}(S_k/k^{1/\alpha} > t) \xrightarrow[k\to\infty]{} \mathbb{P}(Y > t).$$

Denote $\overline{S}_k = S_k/k^{1/\alpha}$. The convergence is given via the convergence of characteristic functions, where the limit characteristic function is given by [59]:

$$\phi_Y(t) = \lim_{k\to\infty} \phi_{\overline{S}_k}(t) = \exp(-C_1|t|^{\alpha}), \tag{2.10}$$

where $C_1 > 0$ is constant that depends on α. Hence, in the bounded interval $|t| < 1$ the convergence in (2.10) is uniform in t, thus we can write

$$\phi_{\overline{S}_k}(t) = \exp(-C_1|t|^{\alpha}(1 + o(1))),$$

where $o(1) \to 0$ as $k \to \infty$ uniformly in $|t| < 1$. Using the relation between the tail distribution and the characteristic function, given by the following inequality ([21], Eq. (3.3.1)):

$$\mathbb{P}(|X| > 2/u) \leqslant \frac{1}{u}\int_{-u}^{u} (1 - \phi_X(t))dt,$$

where X is a random variable with characteristic function $\phi_X(t)$, we derive that when t is sufficiently large then for all k,

$$\mathbb{P}(\overline{S}_k > t) \leqslant t \int_{-2/t}^{2/t} 1 - \exp(-C_2|x|^\alpha)dx < t \int_{-2/t}^{2/t} C_2|x|^\alpha dx = C_3 t^{-\alpha}, \tag{2.11}$$

where $C_3 > 0$ is constant that depends on α. Thus we have for sufficiently large t:

$$\mathbb{P}(\min\{S_k, S_k^*\}/k^{1/\alpha} > t) \leqslant C_4 t^{-2\alpha},$$

where $C_4 > 0$ is constant that depends on α. Since S_k is positive then we can find a random variable Z with power-law tail with exponent 2α such that $|\min\{S_k, S_k^*\}/k^{1/\alpha}| < Z$ a.s. for all $k > 0$, and thus by Dominated Convergence Theorem for $\alpha > 1/2$ we have convergence of expectations

$$\mathbb{E}(\min\{S_k, S_k^*\}/k^{1/\alpha}) \xrightarrow[k\to\infty]{} \mathbb{E}(\min\{Y, Y^*\}).$$

where $Y, Y*$ are stable with parameter α. The minimum of Y, Y^* has power-law tail with exponent 2α thus has finite expectation and we have:

$$\mathbb{E}(\min\{S_k, S_k^*\}/k^{1/\alpha}) \asymp 1,$$

for all $k > 0$, which implies the statement of the theorem. □

2.3.2 Spreading on a star

Consider the example of the SI spreading process with power-law weights on the n-star. The graph ST_n of the n-star is defined as the distinguished root vertex 0 and vertices $\{1, 2, \ldots, n-1\}$ 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 $X_1, X_2, \ldots, X_{n-1}$. 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 the n-star, where $n \geqslant 2$. Then in the SI spreading process $T = (T_k)_{k=1}^n$ on ST_n with weights having power-law distribution with $\alpha \in (1/2, 1)$ the expected time to infect k vertices is bounded for $k < n-2$,*

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

where $C > 0$ is constant that depends on α, and for $k = n-2$,

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

Proof. Denote $X_{(k)}^{n-1}$ the k'th order statistic of $X_1, \dots, X_{n-1}$, then we have $T_k = X_{(k)}^{n-1}$ and it is obvious that

$$X_{(k)}^{n-1} \preceq X_{(k)}^{k+1},$$

for $k \leqslant n-2$. Then, it is straightforward to calculate the tail distribution of $X_{(k)}^{k+1}$:

$$\begin{aligned}\mathbb{P}(X_{(k)}^{k+1} > t) = 1 - \mathbb{P}(X_{(k)}^{k+1} < t) = 1 - (k+1)\mathbb{P}(X_1, \dots, X_k < t, X_{k+1} > t) - \mathbb{P}(X_1, \dots, X_{k+1} < t) = \\ = 1 - (k+1)(1-t^{-\alpha})^k t^{-\alpha} - (1-t^{-\alpha})^{k+1}.\end{aligned} \tag{2.12}$$

Using for $t > k^{1/\alpha}$ the bound $(1-t^{-\alpha})^k \sim \exp(-kt^{-\alpha}) > 1 - kt^{-\alpha}$, and for $t < k^{1/\alpha}$ the bound $\mathbb{P}(X_{(k)}^{k+1} > t) \leqslant 1$, and plugging them into (2.12), we obtain the following upper bound on the expectation:

$$\begin{aligned}\mathbb{E}(X_{(k)}^{k+1}) \leqslant \int_0^{k^{1/\alpha}} dt + C_1(k+1)k \int_{k^{1/\alpha}}^{\infty} t^{-2\alpha} dt = \\ = k^{1/\alpha} + \frac{1}{2\alpha - 1}\frac{k+1}{k}k^{1/\alpha} \leqslant C_2 k^{1/\alpha}.\end{aligned}$$

Hence, we have the first statement of the theorem. In order to prove the second, we need to obtain the lower bound using for $t > k^{1/\alpha}$ the bound $\exp(-kt^{-\alpha}) < 1 - kt^{-\alpha} + \frac{1}{2}k^2 t^{-2\alpha}$, and for $t < k^{1/\alpha}$ the bound $\mathbb{P}(X_{(k)}^{k+1} > t) > 0$:

$$\begin{aligned}\mathbb{E}(X_{(k)}^{k+1}) \geqslant (k^2-1)\int_{k^{1/\alpha}}^{\infty} t^{-2\alpha} dt - \frac{1}{2}(k+1)k^2 \int_{k^{1/\alpha}}^{\infty} t^{-3\alpha} dt = \\ = \frac{1}{2\alpha - 1}\frac{k^2-1}{k^2}k^{1/\alpha} - \frac{1}{3\alpha - 1}\frac{(k+1)k^2}{2k^3}k^{1/\alpha} \geqslant C_3 k^{1/\alpha},\end{aligned}$$

where $C_3 > 0$ is constant that depends on α. Thus,

$$\mathbb{E}(X_{(k)}^{k+1}) \asymp k^{1/\alpha},$$

which finishes proof of the theorem. □

2.4 General deterministic graphs

Consider the connected rooted graph $G = (V, E, s)$ with n vertices and m edges and the distinguished root vertex $s \in V$. Denote the i.i.d. random weight attached to the edge $e \in E$ as X_e, where X_e is 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 such (s,t)-path. Let $\mathcal{G}_k$ be the set of subtrees of G on k vertices with the same root s.

Then we can define the 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 edges of the graph G with the product measure $P^m(d\omega) = P(d\omega_1) \times \cdots \times P(d\omega_m)$.

We can define 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 $H_k \in \mathcal{G}_k$ of the maximum over vertices $t \in V(H_k)$ of the total weight of the shortest weighted (s, t)-path. In symbols,

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

The 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 $\omega \in \Omega$. Remember, an occupied edge is an edge with both ends in the infected state I. Each sample sequence $T(\omega)$, $\omega \in \Omega$, defines an order $\epsilon_{T(\omega)} = (e_1^\epsilon, e_2^\epsilon, \ldots, e_m^\epsilon)$ on the edge set, in which they are occupied by the process. It may happen that at some $T_k(\omega)$ two or more edges incident to a newly infected vertex become occupied at the same time. In this case we assume that first the edge on the shortest (weighted) path to the root is occupied, and the rest are occupied with respect to some fixed generic order $\bar{\epsilon}$ on the edge set to eliminate ambiguity. Given $T(\omega)$, one can also define the *FPP tree* $G(T_k(\omega))$ for each $k \leqslant n$ as the subtree from $\mathcal{G}_k$ that consists of those edges that successfully passed the infection, i.e., the occupied edges on the shortest paths between the root s and the infected vertices. It may happen that some vertex has two or more shortest paths with equal total weight, but since the system is finite, this event happens with zero probability.

In the current framework, for each $\omega \in \Omega$ the spreading curve is defined as the set of pairs $(T_k(\omega), k/n)$, where $1 \leqslant k \leqslant n$, and the average spreading curve is the set of pairs $\{(\mathbb{E}(T_k), k/n) : 1 \leqslant k \leqslant n\}$. Our goal is to mathematically define the position of the first temporal bottleneck, responsible for the jump on the average spreading curve. We restrict ourselves to consideration of power-law distributed weights with $\alpha \in (1/2, 1)$, however some lemmas consider general weights.

First temporal bottlenecks. Remember we call an edge active at time t if one of its incident vertices is in state S and the other is in state I. In other words, an active edge is an edge that currently transmits an infection. Let $\omega \in \Omega$ and the *front of the epidemic* $F(T_k(\omega))$ be the set of edges, that are active at time $T_k(\omega)$, where $k \in \{1, \ldots, n\}$, in the sample sequence $T(\omega)$. Define $\kappa(G, s)$ to be the *maximal* 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. In other words, it is the *minimal* k such that there exists $\omega \in \Omega$ with $F(T_k(\omega))$ having one or zero active edge. 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 now prove that if there is a sample sequence $T(\omega)$ and a number i, for which the front of the epidemics $F(T_i(\omega))$ has one active edge, then there is a big jump at this point on the average spreading curve.

Lemma 2.4.1. *Let G be a finite rooted graph and let T be the SI spreading process with weights having absolutely continuous distribution ξ, such that $\mathbb{E}(\xi) = \infty$. Let there exist $\omega_0 \in \Omega$, the sample sequence $T(\omega_0)$ and a number $i \in \mathbb{N}$, such that the front $F(T_i(\omega_0))$ has one active edge. Then for each j, where $i + 1 \leqslant j \leqslant n$, the expected passage time is*

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

Proof. The sample sequence $T(\omega_0)$ defines the order of occupation of the edge set $\epsilon_{T(\omega)} = (e_1^\epsilon, e_2^\epsilon, \ldots, e_m^\epsilon)$. Since all edge weights have absolutely continuous distribution and the number of edges is finite, there exists a subset $\mathcal{A}(\omega_0)$ with positive measure of sample sequences with the same order of occupation of edges as in $\epsilon_{T(\omega)}$. More precisely, there exists a small $\varepsilon > 0$ such that the set:

$$\mathcal{A}(\omega_0) = \{\omega : |X_{e_j^\epsilon}(\omega) - X_{e_j^\epsilon}(\omega_0)| < \varepsilon\},$$

which has positive measure for any $\varepsilon > 0$ since the state space is finite, has the property that the order of occupation of the edges is the same for all $\omega \in \mathcal{A}(\omega_0)$ as in $\epsilon_{T(\omega_0)}$. For this, one can take ε to be smaller than half of the minimum of all the absolute differences between the edge weights of different edges (which is almost surely positive).

Then, since the front $F(T_i(\omega_0))$ has one active edge, then for each $\omega \in \mathcal{A}(\omega_0)$, the front $F(T_i(\omega))$ also has one active edge. Therefore, we have

$$\mathbb{E}(T_{i+1} - T_i \mid \mathcal{A}(\omega_0)) = \mathbb{E}(X) = \infty,$$

and by the law of total expectation

$$\mathbb{E}(T_{i+1}) = \mathbb{E}(T_{i+1} - T_i) + \mathbb{E}(T_i) > \mathbb{E}(T_{i+1} - T_i \mid \mathcal{A}(\omega_0))\mathbb{P}(\mathcal{A}(\omega_0)) = \infty.$$

Since for all j, where $i + 1 < j < n$ we have $T_j \preceq T_{j+1}$, then we have $\mathbb{E}(T_{j+1}) = \infty$, which finishes the proof of the lemma. □

There exists a combinatorial description of $\kappa(G, s)$, which is given in the following Lemma.

Lemma 2.4.2. *Let G be a finite rooted graph with root s and let T be the SI spreading process on G with weights having absolutely continuous distribution ξ, such that $\mathbb{E}(\xi) = \infty$. Then,*

$$\kappa(G, s) = \min_{e \in E(G)} |\mathcal{C}(s, G \backslash e)|,$$

where $|\mathcal{C}(s, G\backslash e)|$ is the size of the connected component of vertex s in the graph G without edge e.

Proof. Suppose that there exists $\omega \in \Omega$ and the number k, where $0 < k < n$, such that $F(T_k(\omega))$ has one active edge $e \in E$. Then at time $T_k(\omega)$ we can divide vertices of G into two classes: infected (in state I) and susceptible (in state S). In the induced subgraph on infected vertices all edges are occupied, and in the subgraph on susceptible vertices all edges are unoccupied, and there exists only one edge e between these two subgraphs. Hence, the active edge e is a cut edge, and, by definition, the size of the infected subgraph equals to $k = |\mathcal{C}(s, G\backslash e)|$. Since $\kappa(G, s)$ is defined as the minimum of such k, hence

$$\kappa(G, s) = \min_{e \in E(G)} |\mathcal{C}(s, G \backslash e)|.$$

This finishes the proof of the Lemma. □

Delayed process $\overline{T}$. Fix an arbitrary order on the edge set $\epsilon = (e_1^\epsilon, e_2^\epsilon, \ldots, e_m^\epsilon)$. Define the process $\overline{T} = \{\overline{T}_k : k \in \{1, \ldots, n\}\}$, coupled with the original process T as follows. Start with all vertices in the susceptible state S and let the root s be infected (in the state I). Then choose the two active edges incident to s with smallest indices in order ϵ and spread the infection through them. At the time when one of these edges becomes occupied, choose the next active edge from E with the smallest index in ϵ and repeat the procedure. If both active edges share one susceptible vertex, then when one edge gets occupied, choose two new active edges with smallest indices in ϵ. The process runs until there are no more new active edges to take and the remaining times are assumed to be infinite.

Obviously, for each $k < \kappa(G, s)$ and each $\omega \in \Omega$ the front $F(\overline{T}_k(\omega))$ has two active edges, since the delayed process can be turned into a particular realization of an original process with ϵ order of edge occupation.

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

Lemma 2.4.3. *Let $G = (V, E, s)$ be a finite rooted graph with $|V| = n$ and $|E| = m$ and let T be the SI spreading process on G with positive weights having arbitrary distribution ξ. Then the delayed process $\overline{T}$ stochastically dominates the process T.*

Proof. Consider an $\omega \in \Omega$. Then the sample sequence of the delayed process $\overline{T}_k(\omega)$ induces the FPP tree $G(\overline{T}_k(\omega))$ and we have

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

On the other hand, the original process is given by the minimum over all possible subtrees on k vertices:

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

Therefore, we have $T(\omega) \leqslant \overline{T}(\omega)$ and, therefore, $T \preceq \overline{T}$. □

The delayed process runs slower than the original one, but next we define the Q process, which is even slower, but is necessary to achieve our final statement.

Process Q. Define the process $Q = \{Q_k : k \in \{1, \dots, n\}\}$ to be the process in which at each time Q_k there are two active edges with weights X and Y in the front: one of them is always old, with the age of the process, and an other is new. In symbols, let

$$\begin{aligned} Q_1 &= 0, \\ Q_2 &= \min\{X, Y\}, \\ Q_{k+1} &= Q_k + \min\{X, Y - Q_k | Q_k, Y > Q_k\}, \end{aligned} \tag{2.13}$$

where the unconditional X, Y are are i.i.d. edge weights. The process Q qualitatively constitutes the worst scenario the infection can spread on $\mathbb{Z}$, having an ever old edge Y and spreading only along new edges X in one direction. The following lemma provides a bound on the expected time to infect k vertices in the process Q.

Lemma 2.4.4. *Consider the process Q_k defined above in (2.13) with $X, Y \sim pow(\alpha)$. Then, for $\alpha \in (1/2, 1)$ and for each k, where $k \geqslant 1$, we have*

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

where $d > 0$ is a constant that depends on α.

Proof. Using the law of total expectation, Lemma 2.2.7 and Jensen's inequality we have that

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

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

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

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

By Lemma 2.2.4 this sequence is bounded and we have for any $k \geqslant 1$:

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

This finishes the proof of the lemma. □

Define the random variable X_s, where $s > 0$, with the following probability measure:

$$\mathbb{P}(X_s > t) := \mathbb{P}(X - s > t | X > s).$$

We call the random variable X to have a *shifted power-law* distribution $shiftpow(\alpha)$ with $\alpha > 0$, if $\mathbb{P}(X > t) = (t+1)^{-\alpha}$, when $t \geqslant 0$, and $\mathbb{P}(X > t) = 1$ otherwise. Note that for any $s > 0$, if the random variable $X \sim shiftpow(\alpha)$ with $\alpha > 0$, then $X_s \stackrel{d}{=} (s+1)X$ and therefore, for any $s_1 < s_2$:

$$X_{s_1} \preceq X_{s_2}. \tag{2.14}$$

In other words, if we consider the SI spreading with shifted power-law weights, then the older edges dominate the newer ones. We prove now the main theorem of this section.

Theorem 2.4.5. *Let $G = (V, E, s)$ be a connected rooted graph with root s and $|V| = n$. Let T be the SI spreading process on G with power-law weights with $\alpha \in (1/2, 1)$. Then for each k, where $1 \leqslant k \leqslant \kappa(G, s)$, the expected time to infect k vertices is bounded by*

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

where $C > 0$ is a constant, that depends on α.

Proof. Let a random variable $X \sim pow(\alpha)$, then $(X - 1) \sim shiftpow(\alpha)$. Define $T^{(X-1)}$ to be the coupled to T SI spreading process with shifted power-law weights with the same parameter α as in T. The shift in weights is deterministic, hence by time T_k, the process with shifted weights is faster than the original process by the cumulative shift of not more than $k - 1$:

$$T_k - T_k^{(X-1)} \leqslant k - 1,$$

and equality holds only if the spreading happens along a path of length $k-1$ in the graph. The cumulative shift depends on the shape of the FPP tree given by the process at time T_k and can therefore be non-deterministic.

Since for $X, Y \sim shiftpow(\alpha)$, then by Lemma 2.4.3 the delayed process $\overline{T}^{(X-1)}$ with the same weights dominates the process $T^{(X-1)}$, and we have for any $k \leqslant \kappa(G,s)$:

$$T_k \preceq \overline{T}_k^{(X-1)} + (k-1).$$

Now consider the process Q with shifted power-law weights denoted as $Q^{(X-1)}$. Since old edges dominate newer ones, the process $Q^{(X-1)}$ dominates the process $\overline{T}^{(X-1)}$, hence we have

$$T_k \preceq Q_k^{(X-1)} + (k-1). \tag{2.15}$$

Since the shift is negative, then we have $Q^{(X-1)} \preceq Q$. Hence, we have for each k, where $1 \leqslant k \leqslant \kappa(G,s)$,

$$T_k \preceq Q_k + (k-1).$$

By Lemma 2.4.4 we have

$$\mathbb{E}(T_k) \leqslant dk^{1/\alpha} + (k-1) \leqslant (d+1)k^{1/\alpha},$$

which finishes the proof of the theorem. □

Based on Theorem 2.4.5 we can equivalently define the number $\kappa(G,s)$ to be the maximal number of vertices k in the SI spreading process with power-law weights with $\alpha \in (1/2, 1)$, such that the expected time to infect k vertices is finite. From the statement of Lemma 2.4.2 we have the following corollary.

Corollary 2.4.6. *Let G be a finite 2-edge-connected rooted graph with root s and let T be the SI spreading process on G with power-law weights, where $1/2 < \alpha < 1$. Then for each $k \leqslant n$,*

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

2.5 Critical Galton-Watson tree with an extra edge

The appearance of jumps on the average spreading curve of the SI process on deterministic graphs is described by Theorem 2.4.5. It has been shown that any rooted graph has a specific number of vertices that can be infected in finite time in expectation. We apply this result

to random graph models. In the current section we prove that w.h.p. the average spreading curve of the SI process with power-law weights with $\alpha \in (1/2, 1)$ on a large critical Galton-Watson tree $\mathcal{T}$ has a jump shortly after the beginning of the process, but addition of just one extra edge e between the root and a random vertex in $\mathcal{T}$ eliminates jumps up to occupying a positive fraction of vertices.

We consider the case of a critical Galton-Watson (CGW) tree with integer offspring distribution ξ, such that $Var(\xi) = \sigma^2 < \infty$. First we establish some technical lemmas.

2.5.1 Technical lemmas

Denote as $\mathcal{T}$ a CGW tree with a respective process denoted as Z and 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 the first k generations of the tree $\mathcal{T}$ and $\mathcal{T}^\infty$ respectively match the first k generations of a given tree T. Denote as $\#T_k$ the size of k'th generation in the tree T. The following lemma holds for trees $\mathcal{T}$ and $\mathcal{T}^\infty$ [41].

Lemma 2.5.1. *Let $\mathcal{T}$ be a CGW tree with offspring distribution ξ. Then, for any rooted vertex-labeled 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

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

It is natural that as $N \to \infty$ the conditioned tree $\mathcal{T}^N := (\mathcal{T}|Z_N > 0)$ and $\mathcal{T}^\infty$ w.h.p. start to look similar. The question now is how large (as a function of N) that similar part is. The following lemma derives an exact result on this.

Lemma 2.5.2. *Let $\mathcal{T}^N$ be a CGW 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 $\delta > 0$ and a coupling between $\mathcal{T}^N$ and $\mathcal{T}^\infty$, such that*

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

Proof. In order to prove the statement of the theorem we show that the conditioned measure is close to the infinite measure in total variation distance. Consider a rooted tree T with height k, where $k \leqslant \delta N$ and $\delta > 0$ is small. Then by Bayes' formula we can write

$$\begin{aligned} \mathbb{P}(\mathcal{T}[k] = T | Z_N > 0) &= \frac{\mathbb{P}(Z_N > 0 | \mathcal{T}[k] = T)}{\mathbb{P}(Z_N > 0)} \mathbb{P}(\mathcal{T}[k] = T) = \\ &= \frac{\mathbb{P}(Z^{(1)}_{N-k} > 0 \cup \cdots \cup Z^{(\#T_k)}_{N-k} > 0)}{\mathbb{P}(Z_N > 0)} \mathbb{P}(\mathcal{T}[k] = T), \end{aligned} \tag{2.16}$$

where $Z^{(i)}_{N-k}$ denotes the $(N-k)$'th generation in the copy of the CGW process $Z^{(i)}$, started from a vertex at level k. By Theorem 2.2.13, for a large N there exists $\varepsilon_0 > 0$ such that,

$$\frac{2}{\sigma^2 N}(1-\varepsilon_0) < \mathbb{P}(Z_N > 0) < \frac{2}{\sigma^2 N}(1+\varepsilon_0). \tag{2.17}$$

We have that when $N-k$ is large enough, the analogue of (2.17) is valid for each $Z^{(i)}_{N-k}$ with another $\varepsilon_1 > 0$:

$$\frac{2}{\sigma^2 (N-k)}(1-\varepsilon_1) < \mathbb{P}(Z^{(i)}_{N-k} > 0) < \frac{2}{\sigma^2 (N-k)}(1+\varepsilon_1), \tag{2.18}$$

where $1 \leqslant i \leqslant \#T_k$. In order to simplify the further calculations we take common $\varepsilon_2 := \max(\varepsilon_0, \varepsilon_1)$ instead of ε_0 and ε_1 in (2.17) and (2.18). Now we obtain an upper bound on $\mathbb{P}(\mathcal{T}[k] = T | Z_N > 0)$ using the union bound on the right-hand side of (2.16) and together with (2.17) and (2.18):

$$\begin{aligned} \mathbb{P}(\mathcal{T}[k] = T | Z_N > 0) &\leqslant \frac{\#T_k \mathbb{P}(Z_{N-k} > 0)}{\mathbb{P}(Z_N > 0)} \mathbb{P}(\mathcal{T}[k] = T) \leqslant \\ &\leqslant \frac{N}{(N-k)} \#T_k \mathbb{P}(\mathcal{T}[k] = T) \frac{1+\varepsilon_2}{1-\varepsilon_2}. \end{aligned} \tag{2.19}$$

Therefore, we can write that for small enough k there exists $\varepsilon_3 > 0$, such that

$$\mathbb{P}(\mathcal{T}[k] = T | Z_N > 0) \leqslant \frac{N}{(N-k)} \#T_k \mathbb{P}(\mathcal{T}[k] = T)(1+\varepsilon_3). \tag{2.20}$$

In order to obtain a lower bound on $\mathbb{P}(\mathcal{T}[k] = T | Z_N > 0)$ we rewrite (2.16) using (2.20) as follows:

$$\begin{aligned} \mathbb{P}(\mathcal{T}[k] = T | Z_N > 0) &= \frac{1}{\mathbb{P}(Z_N > 0)} \left(1 - \mathbb{P}(Z^{(1)}_{N-k} = 0 \cap \cdots \cap Z^{(\#T_k)}_{N-k} = 0)\right) \mathbb{P}(\mathcal{T}[k] = T) = \\ &= \frac{1}{\mathbb{P}(Z_N > 0)} \left(1 - (1 - \mathbb{P}(Z_{N-k} > 0))^{\#T_k}\right) \mathbb{P}(\mathcal{T}[k] = T) > \\ &> \frac{1}{\mathbb{P}(Z_N > 0)} \left(1 - \left(1 - \frac{2(1-\varepsilon_1)}{\sigma^2 (N-k)}\right)^{\#T_k}\right) \mathbb{P}(\mathcal{T}[k] = T). \end{aligned} \tag{2.21}$$

Since for any x, where $0 < x$, we have:

$$1 - x < \exp(-x) < 1 - x + x^2/2,$$

therefore for $n \geqslant 1$:

$$1 - (1-x)^n > 1 - \exp(-nx) > nx - \frac{(nx)^2}{2}. \tag{2.22}$$

We rewrite (2.21) using (2.22) for $x = \mathbb{P}(Z_{N-k} > 0)$ and $n = \#T_k$ as follows:

$$\begin{aligned} \mathbb{P}(\mathcal{T}[k] = T | Z_N > 0) &> \frac{\mathbb{P}(\mathcal{T}[k] = T)}{\mathbb{P}(Z_N > 0)} \left(1 - \left(1 - \frac{2(1-\varepsilon_1)}{\sigma^2 (N-k)}\right)^{\#T_k}\right) \mathbb{P}(\mathcal{T}[k] = T) = \\ &= \frac{\mathbb{P}(\mathcal{T}[k] = T)}{\mathbb{P}(Z_N > 0)} \left(\#T_k \mathbb{P}(Z_{N-k} > 0) - \frac{1}{2} \left(\#T_k \mathbb{P}(Z_{N-k} > 0)\right)^2\right). \end{aligned} \tag{2.23}$$

Using (2.20) we can further write:

$$\mathbb{P}(\mathcal{T}[k]=T|Z_N>0) > \frac{\mathbb{P}(\mathcal{T}[k]=T)}{\mathbb{P}(Z_N>0)}\left(\frac{2\#T_k(1-\varepsilon_2)}{\sigma^2(N-k)} - \frac{1}{2}\left(\frac{2\#T_k(1-\varepsilon_2)}{\sigma^2(N-k)}\right)^2\right), \tag{2.24}$$

and then by (2.19) we obtain

$$\begin{aligned}\mathbb{P}(\mathcal{T}[k]=T|Z_N>0) &> \mathbb{P}(\mathcal{T}[k]=T)\frac{\sigma^2 N}{2(1+\varepsilon_2)}\left(\frac{2\#T_k(1-\varepsilon_2)}{\sigma^2(N-k)} - \frac{1}{2}\left(\frac{2\#T_k(1-\varepsilon_2)}{\sigma^2(N-k)}\right)^2\right) > \\ &> \mathbb{P}(\mathcal{T}[k]=T)\#T_k\left(\frac{N}{N-k}\frac{1-\varepsilon_2}{1+\varepsilon_2} - \#T_k\frac{N}{(N-k)^2}\right).\end{aligned} \tag{2.25}$$

Therefore, we can write that for small enough k there exists $\varepsilon_4>0$ and a bounded $C_\varepsilon>0$ which is uniquely defined by ε_4, such that

$$\mathbb{P}(\mathcal{T}[k]=T|Z_N>0) > \mathbb{P}(\mathcal{T}[k]=T)\#T_k\left(\frac{N}{N-k} - \#T_k\frac{C_\varepsilon N}{(N-k)^2}\right)(1-\varepsilon_4). \tag{2.26}$$

Combining the (2.20) and (2.26), and choosing $\varepsilon_5 := \max\{\varepsilon_3,\varepsilon_4\}$, we obtain the following bounds on the probability $\mathbb{P}(\mathcal{T}[k]=T|Z_N>0)$:

$$\left(\frac{N}{N-k} - \#T_k\frac{C_\varepsilon N}{(N-k)^2}\right)(1-\varepsilon_5) \leqslant \frac{\mathbb{P}(\mathcal{T}[k]=T|Z_N>0)}{\#T_k\mathbb{P}(\mathcal{T}[k]=T)} \leqslant \frac{N}{(N-k)}(1+\varepsilon_5). \tag{2.27}$$

Now we bound the total variation distance between conditioned and infinite measures. Denote the conditioned measure as $\mathbb{P}_N(\cdot) := \mathbb{P}(\cdot \mid Z_N>0)$. From the upper bound in (2.27) we obtain that when k is small enough, the following inequality holds:

$$\begin{aligned}\mathbb{P}_N(\mathcal{T}[k]=T) - \mathbb{P}(\mathcal{T}^\infty[k]=T) &\leqslant \left(\frac{N}{N-k}(1+\varepsilon_5)-1\right)\#T_k\mathbb{P}(\mathcal{T}[k]=T) = \\ &= \left(\left(\frac{N}{N-k}-1\right)+\frac{N}{N-k}\varepsilon_5\right)\#T_k\mathbb{P}(\mathcal{T}[k]=T),\end{aligned}$$

and, on the other hand, from the lower bound in (2.27) we obtain

$$\begin{aligned}\mathbb{P}(\mathcal{T}^\infty[k]=T) - \mathbb{P}_N(\mathcal{T}[k]=T) &\leqslant \left(1-\frac{N}{N-k}(1-\varepsilon_5)+\#T_k\frac{C_\varepsilon N}{(N-k)^2}(1-\varepsilon_5)\right)\cdot \\ &\quad\cdot\#T_k\mathbb{P}(\mathcal{T}[k]=T) = \\ &= \left(\left(1-\frac{N}{N-k}\right)+\frac{N}{N-k}\varepsilon_5+\#T_k\frac{C_\varepsilon N}{(N-k)^2}(1-\varepsilon_5)\right)\cdot \\ &\quad\cdot\#T_k\mathbb{P}(\mathcal{T}[k]=T).\end{aligned}$$

Comparing both bounds we see that all summands are positive, except of $\left(1-\frac{N}{N-k}\right)$, thus we can inverse the sign and derive the bound for an absolute value:

$$\begin{aligned}\mathbb{P}(\mathcal{T}^\infty[k]=T) - \mathbb{P}_N(\mathcal{T}[k]=T) &\leqslant \left(\left(\frac{N}{N-k}-1\right)+\varepsilon_5\frac{N}{N-k}+\#T_k\frac{C_\varepsilon N}{(N-k)^2}(1-\varepsilon_5)\right)\cdot \\ &\quad\cdot\#T_k\mathbb{P}(\mathcal{T}[k]=T).\end{aligned}$$

Summing those bounds over all trees we get:

$$\sum_T \left|\mathbb{P}_N(\mathcal{T}[k]=T)-\mathbb{P}(\mathcal{T}^\infty[k]=T)\right| \leqslant \sum_T \left(\left(\frac{N}{N-k}-1\right)+\varepsilon_5\frac{N}{N-k}+\#T_k\frac{C_\varepsilon N}{(N-k)^2}(1-\varepsilon_5)\right)\cdot \\ \cdot \#T_k\mathbb{P}(\mathcal{T}[k]=T). \tag{2.28}$$

Letting k be equal to δN we further rewrite (2.29):

$$\sum_T \left|\mathbb{P}_N(\mathcal{T}[\delta N]=T)-\mathbb{P}(\mathcal{T}^\infty[\delta N]=T)\right| \leqslant \sum_T \left(\frac{\delta}{1-\delta}+\frac{\varepsilon_5}{1-\delta}+\frac{C_\varepsilon \#T_k}{N}\frac{1-\varepsilon_5}{(1-\delta)^2}\right)\cdot \\ \cdot \#T_k\mathbb{P}(\mathcal{T}[k]=T). \tag{2.29}$$

Since we have a measure on the set of infinite trees, then

$$\sum_T \mathbb{P}(\mathcal{T}^\infty[k]=T)=\sum_T \#T_k\mathbb{P}(\mathcal{T}[k]=T)=1,$$

and from Theorem 2.2.15 we have

$$\sum_T (\#T_k)^2\mathbb{P}(\mathcal{T}[k]=T)=1+k\sigma^2.$$

Therefore we can rewrite (2.29) for $k=\delta N$, when $\delta>0$ is small, and obtain

$$\sum_T \left|\mathbb{P}_N(\mathcal{T}[\delta N]=T)-\mathbb{P}(\mathcal{T}^\infty[\delta N]=T)\right| \leqslant \frac{\delta}{1-\delta}+\frac{\varepsilon_5}{1-\delta}+C'_\varepsilon\delta\frac{1-\varepsilon_5}{(1-\delta)^2}+\frac{C_\varepsilon}{N}\frac{1-\varepsilon_5}{(1-\delta)^2},$$

where $C'_\varepsilon=C_\varepsilon\sigma^2$. Hence, for any $\varepsilon_6>0$ we can find large N and small $\delta>0$, such that

$$\sum_T \left|\mathbb{P}_N(\mathcal{T}[\delta N]=T)-\mathbb{P}(\mathcal{T}^\infty[\delta N]=T)\right| \leqslant \varepsilon_6. \tag{2.30}$$

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. Then, by definition of the total variation distance and (2.30) we have

$$d_{TV}\left(\mathbb{P}_N\restriction_{\delta N},\mathbb{P}_\infty\restriction_{\delta N}\right)\leqslant\frac{1}{2}\varepsilon_6.$$

Hence by Theorem 2.2.5 there exists a coupling of 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. □

We now turn to the tree with an extra edge. Consider a rooted tree $\mathcal{T}$ with root s on n vertices and consider the SI spreading process $T=(T_j)_{j=1}^n$ on this tree with power-law

weights with $\alpha \in (1/2, 1)$. Let the degree of the root be equal to d_s. Denote subtrees hanging from the root as $\mathcal{T}_0, \ldots, \mathcal{T}_{d_s}$. Then, by Lemma 2.4.2,

$$\kappa(\mathcal{T}, s) = |\mathcal{T}| - \max\{|\mathcal{T}_i| : 0 \leqslant i \leqslant d_s\}. \tag{2.31}$$

Now add an extra edge e between s and a randomly chosen vertex in the tree. Then, we obtain a graph, denoted as $\mathcal{T}_{+e}$, which consists of a cycle C_k of length k with N rooted trees $\mathcal{T}_0, \mathcal{T}_1, \ldots, \mathcal{T}_N$, where $N > d_s$ a.s., attached to it by edges $e_0, e_1, \ldots, e_N$ (see Figure2.3). Then, by Lemma 2.4.2, we have

$$\kappa(\mathcal{T}_{+e}, s) = |\mathcal{T}_{+e}| - \max\{|\mathcal{T}_i| : 0 \leqslant i \leqslant N\}. \tag{2.32}$$

This fact is used in the following Section.

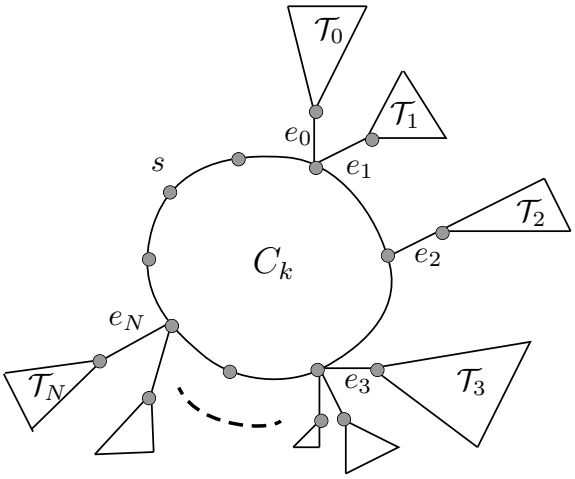


Figure 2.3: *Schematic structure of a graph $\mathcal{T}_{+e}$ of a tree $\mathcal{T}$ with an extra edge e: a cycle C_k with hanging trees $\mathcal{T}_0, \mathcal{T}_1, \ldots, \mathcal{T}_N$.*

2.5.2 Main Theorem

The following theorem shows that in the critical Galton-Watson tree w.h.p. there is a large jump on the average spreading curve near the beginning, but one extra edge lets the process to infect a non-zero fraction of the tree without large jumps.

Theorem 2.5.3. *Consider the CGW tree $\mathcal{T}^N$ conditioned on $Z_N > 0$ and the SI spreading process with power-law weights with $\alpha \in (1/2, 1)$. Denote as $\mathcal{T}^N_{+e}$ the tree $\mathcal{T}^N$ with an extra edge attached to a root and one of the vertices of $\mathcal{T}^N$ uniformly at random, and $|\mathcal{T}^N|$ denotes the number of vertices in $\mathcal{T}^N$. Then as $N \to \infty$,*

1. *the sequence of r.v.* $\kappa(\mathcal{T}^N, s)$ *is tight;*

2. *for any* $\varepsilon > 0$ *there exists* $\delta > 0$*, such that*

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

Proof. 1) Consider first the infinite CGW tree $\mathcal{T}^\infty$. Following the construction of $\mathcal{T}^\infty$ we denote the spine as γ and label the unconditioned CGW trees attached to the root as $t_1, t_2, \dots, t_\nu$, and the rest of the tree as t_0.

Let $n := \delta N$, where $\delta > 0$ is given. The number of the unconditioned trees ν is represented by shifted size-biased i.i.d. random variables

$$\nu = \hat{\xi} - 1.$$

Let $\mu := \mathbb{E}\nu$, then it is straightforward to show that $\mu = \sigma^2 < \infty$. By Markov inequality for any given $C_1 > 0$

$$\mathbb{P}(\nu > C_1) < \frac{\sigma^2}{C_1}.$$

Hence, for any $\varepsilon_1 > 0$ there exists $C_1 > 0$, such that

$$\mathbb{P}(\nu < C_1) > 1 - \varepsilon_1.$$

Then, using the law of total probability we bound the total size of these trees in the following way:

$$\begin{aligned}\mathbb{P}\Big(\sum_{i=1}^{\nu} |t_i| > C_1 K\Big) &= \mathbb{P}\Big(\sum_{i=1}^{\nu} |t_i| > C_1 K \;\Big|\; \nu < C_1\Big)\mathbb{P}(\nu < C_1) + \\ &+ \mathbb{P}\Big(\sum_{i=1}^{\nu} |t_i| > C_1 K \;\Big|\; \nu > C_1\Big)\mathbb{P}(\nu > C_1) < \\ &< \mathbb{P}\Big(\sum_{i=1}^{C_1} |t_i| > C_1 K\Big) + \varepsilon_1.\end{aligned} \tag{2.33}$$

It remains to show that the total size of C_1 trees is bounded. Using the union bound we can write:

$$\begin{aligned}\mathbb{P}\Big(\sum_{i=1}^{C_1} |t_i| > C_1 K\Big) &< \mathbb{P}(\text{at least one } |t_i| > K) \leqslant \\ &< C_1 \mathbb{P}(|t_i| > K),\end{aligned} \tag{2.34}$$

where by Theorem 2.2.12 and Lemma 2.2.3 we derive that for large K,

$$\begin{aligned}\mathbb{P}(|t_1| > K) &\leqslant C_2 \sum_{k=K+1} k^{-3/2} \leqslant C_2 \int_K^\infty x^{-3/2} dx = \\ &= \frac{C_3}{\sqrt{K}},\end{aligned} \tag{2.35}$$

where $C_3 > 0$ is constant that depends on ξ. Hence, for any $\varepsilon_2 > 0$ and given $C_1 > 0$ there exists large K, such that

$$\mathbb{P}\Big(\sum_{i=1}^{\nu} |t_i| < C_1 K\Big) > 1 - \varepsilon_2. \tag{2.36}$$

Since the total size of trees t_i, where $i \geqslant 1$, is bounded, then it follows that their height is bounded too. Then by Lemma 2.5.2, for large N and any $\varepsilon_3 > 0$ we can find $\delta > 0$ and a coupling of $\mathcal{T}^N$ and $\mathcal{T}^\infty$, such that the tree $\mathcal{T}^N[n]$ is same as $\mathcal{T}^\infty[n]$ with probability at least $(1 - \varepsilon_3)$, hence by (2.36) for any $\varepsilon_4 > 0$ and large N there exists K such that in the conditioned tree $\mathcal{T}^N$:

$$\mathbb{P}(\sum_{i=1}^{\nu} |t_i| < C_1 K) > 1 - \varepsilon_4. \tag{2.37}$$

In order to prove the statement of the theorem, we need to show that the size of t_0 is the largest among t_i, where $0 \leqslant i \leqslant \nu$. Thus, we show that the size of the tree $|\mathcal{T}^N|$ is of order N^2 w.h.p. By Bayes' formula:

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

where $\mathcal{T}$ denotes the CGW tree. Since the condition $Z_N > 0$ implies that there exists at least one vertex at each distance k from the root for $k \leqslant N$, then

$$\mathbb{P}(|\mathcal{T}^N| = k) = 0. \tag{2.38}$$

The event $(Z_N > 0 \mid |\mathcal{T}| = M)$ is equivalent to the event $(H(\mathcal{T}) > N \mid |\mathcal{T}| = M)$, hence by Theorem 2.2.12, Theorem 2.2.13 and Theorem 2.2.14, for large N and for any $M > N$, we have the following inequality holds:

$$\mathbb{P}(|\mathcal{T}^N| = M) < C_4 e^{-c_1 \frac{N^2}{M}} N M^{-3/2}, \tag{2.39}$$

where $c_1, C_4 > 0$ are constants that depend on ξ and N. Prove that the probability of having small $\mathcal{T}^N$ is small. Let $h_1 > 0$, then by (2.38) and (2.39) we have

$$\mathbb{P}(|\mathcal{T}^N| < h_1 N^2) < C_4 \sum_{m < h_1 N^2} e^{-c_1 \frac{N^2}{m}} N m^{-3/2},$$

and, by Lemma 2.2.3, the sum can be bounded with an integral:

$$\begin{aligned}
\mathbb{P}(|\mathcal{T}^N| < h_1 N^2) < C_4 \int_N^{h_1 N^2} e^{-c_1 \frac{N^2}{m}} N m^{-3/2} dm = \frac{C_4}{N^2} \int_{1/N}^{h_1} e^{-c_1/x} x^{-3/2} dx = \\
= \frac{C_4}{\sqrt{c_1}} \int_{c_1 N}^{c_1/h_1} e^{-y} y^{-1/2} dy < \frac{C_4}{\sqrt{c_1}} \int_{\infty}^{c_1/h_1} e^{-y} y^{-1/2} dy < \\
< \frac{C_4}{\sqrt{c_1}} \Gamma\left(1/2, \frac{c_1}{h_1}\right),
\end{aligned}$$

By Lemma 2.2.2, as $h_1 \to 0$,

$$\mathbb{P}(|\mathcal{T}^N| < h_1 N^2) < C_4 h_1^{-1/2} e^{-c_1/h_1}, \tag{2.40}$$

which gives us that for any $\varepsilon_5 > 0$ there exists $h_1 > 0$, such that

$$\mathbb{P}(|\mathcal{T}^N| > h_1 N^2) > 1 - \varepsilon_5. \tag{2.41}$$

Since by (2.36) the total size of trees t_i, where $i \geqslant 1$, is bounded with the absolute constant w.h.p. then for large enough N the size of t_0 has the same order of magnitude as $|\mathcal{T}^N|$ w.h.p. Hence we can apply formula (2.31), where $\mathcal{T}_i = t_i$ for $i \geqslant 0$ and derive that for any $\varepsilon_6 > 0$ there exists $L > 0$ and for any large enough N:

$$\mathbb{P}(\kappa(\mathcal{T}^N, s) < L) > 1 - \varepsilon_6.$$

2) In the following part we follow a similar line to prove the statement. First consider the infinite tree $\mathcal{T}^\infty$. Denote as γ_k, where $k \geqslant 1$, the initial part of γ between the root and the vertex at depth k inclusive. Following the construction of $\mathcal{T}^\infty$ label the unconditioned trees attached to γ_k in the breadth-first order as $t_1, t_2, \ldots, t_{\nu_1}, \ldots, t_{\nu_k}$. Denote the number of unconditioned trees $t_1, \ldots, t_{\nu_k}$ as $S_k = \sum_{i=1}^{k} \nu_i$, where $\nu_i = \hat{\xi}_i - 1$ is the shifted size-biased version of the offspring distribution ξ.

Let $n := \delta N$ and $n' := \delta^2 N$. Consider $S_{n'}$ unconditioned trees $t_1, \ldots, t_{\nu_{n'}}$ that hang off $\gamma_{n'}$ and let $t_0 := \mathcal{T}^\infty \setminus (\gamma_{n'} \cup t_1 \cup \cdots \cup t_{\nu_{n'}})$. We show that t_i, where $1 \leqslant i \leqslant \nu_{n'}$ do not go deeper than n'th generation in $\mathcal{T}^\infty$ w.h.p. or, in other words, they have height at most $(n - n')$. Indeed by Theorem 2.2.13 for any $\varepsilon_1' > 0$ and small $\delta > 0$ there exists large enough N such that

$$\mathbb{P}(H(t_1) > n - n') = \mathbb{P}(H(t_1) > (1-\delta)\delta N) < \frac{2}{\sigma^2(1-\delta)\delta N}(1 + \varepsilon_1').$$

We will now prove that the number of unconditioned trees $S_{n'}$ is bounded with linear function in n' w.h.p. Remember that $\mu = \mathbb{E}(\nu) < \infty$, then by the LLN for any $\varepsilon_2' > 0$ we can find such $K_1' > 0$ such that

$$\mathbb{P}(S_{n'} < K_1' \mu n') > 1 - \varepsilon_2'.$$

Hence, using the law of total probability we have

$$\mathbb{P}\big(\exists i \in \{1, \ldots, \nu_{n'}\} : \ H(t_i) > n - n'\big) < \mathbb{P}\big(\exists i \in \{1, \ldots, K_1' \mu n'\} : \ H(t_i) > n - n'\big) + \varepsilon_2'.$$

Using the union bound we can show that

$$\mathbb{P}\big(\exists i \in \{1, \ldots, K_1' \mu n'\} : \ H(t_i) > n - n'\big) < \frac{2K_1'\mu\delta}{\sigma^2(1-\delta)}(1 + \varepsilon_1') < C_1' \frac{\delta}{1-\delta}(1 + \varepsilon_1').$$

Thus, for any $\varepsilon_3' > 0$ there exists $\delta > 0$ small enough, such that

$$\mathbb{P}\big(H(t_i) < n - n' \ \forall i \in \{1, \dots, \nu_{n'}\}\big) > 1 - \varepsilon_3'.$$

We now prove that the total size of $S_{n'}$ unconditioned trees is of order n'^2 with high probability. Using the law of total probability we can write

$$\begin{aligned}\mathbb{P}(\sum_{i=1}^{S_{n'}} |t_i| > K_2' n'^2) = \mathbb{P}(\sum_{i=1}^{S_{n'}} |t_i| > K_2' n'^2 \mid S_{n'} < K_1'\mu n')\mathbb{P}(S_{n'} < K_1'\mu n') + \\ + \mathbb{P}(\sum_{i=1}^{S_{n'}} |t_i| > K_2' n'^2 \mid S_{n'} > K_1'\mu n')\mathbb{P}(S_{n'} > K_1'\mu n') \leqslant \\ < \mathbb{P}(\sum_{i=1}^{K_1'\mu n'} |t_i| > K_2' n'^2) + \varepsilon_2'.\end{aligned} \tag{2.42}$$

Thus, for an upper bound it remains to show that the size of $K_1'\mu n'$ trees has order at most $K_2' n'^2$ with high probability. By Theorem 2.2.12 the tail distribution of $|t_i|$ can be bounded below as follows:

$$\begin{aligned}\mathbb{P}(|t_1| \geqslant x) \geqslant C_2' \sum_{n=x}^{\infty} n^{-3/2} \geqslant \\ \geqslant C_2' \int_x^{\infty} z^{-3/2} dz = C_3' x^{-1/2}.\end{aligned}$$

Combining it with (2.35) we obtain that $|t_1|$ has power-law tail and then by Theorem 2.2.8 the size $|t_1|$ belongs to the domain of attraction of the stable law Y with $\alpha = 1/2$ that has the following density:

$$p_Y(y) = \frac{1}{\sqrt{2\pi}} y^{-3/2} e^{-\frac{1}{2y}}. \tag{2.43}$$

Hence for large enough $K_2' > 0$ and large N we have

$$\begin{aligned}\mathbb{P}\Big(\sum_{i=1}^{K_1'\mu n'} |t_i| > K_2' n'^2\Big) \sim \mathbb{P}\Big(Y > \frac{K_2'}{(K_1'\mu)^2}\Big) = \frac{1}{\sqrt{2\pi}} \int_{K_2'/(K_1'\mu)^2}^{\infty} y^{-3/2} e^{\frac{1}{2y}} dy = \\ = \frac{1}{\sqrt{2\pi}} \int_0^{(K_1'\mu)^2/2K_2'} z^{-1/2} e^{-z} dz = \frac{1}{\sqrt{2\pi}} \gamma\left(1/2, \frac{(K_1'\mu)^2}{2K_2'}\right).\end{aligned}$$

Then, by Lemma 2.2.1, we have that for large $K_2' > 0$:

$$\mathbb{P}\Big(\sum_{i=1}^{K_1'\mu n'} |t_i| > K_2' n'^2\Big) \sim \frac{K_1'\mu}{2\sqrt{\pi K_2'}}.$$

Therefore, for any $\varepsilon_4' > 0$ we can find large $K_2' > 0$ such that

$$\mathbb{P}\Big(\sum_{i=1}^{S_{n'}} |t_i| < K_2' n'^2\Big) > 1 - \varepsilon_4'. \tag{2.44}$$

The lower bound is given as follows. By the LLN for any $\varepsilon_5' > 0$ we can find small $k_2' > 0$ such that

$$\mathbb{P}(S_{n'} > k_1' \mu n') > 1 - \varepsilon_5',$$

and using the law of total probability as before we can write for small enough $k_2' > 0$ and large N:

$$\mathbb{P}(\sum_{i=1}^{S_{n'}} |t_i| < k_2' n'^2) < \varepsilon_5' + \mathbb{P}(\sum_{i=1}^{k_1' \mu n'} |t_i| < k_2' n'^2). \tag{2.45}$$

Thus, it remains to show that the size of $k_1' \mu n'$ trees has order at least $k_2' n'^2$ with high probability. Since all $|t_i|$ are i.i.d. and belong to the domain of attraction of the stable law Y from (2.43) then we have for large N:

$$\mathbb{P}\Big(\sum_{i=1}^{k_1' \mu n'} |t_i| < k_2' n'^2\Big) \sim \mathbb{P}\Big(Y < \frac{k_2'}{(k_1'\mu)^2}\Big) = \frac{1}{\sqrt{2\pi}} \int_0^{k_2'/(k_1'\mu)^2} y^{-3/2} e^{\frac{1}{2y}} dy =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{(k_1'\mu)^2/2k_2'}^{\infty} z^{-1/2} e^{-z} dz = \frac{1}{\sqrt{2\pi}} \Gamma\left(1/2, \frac{(k_1'\mu)^2}{2k_2'}\right).$$

Then, by Lemma 2.2.2, we have that for small $k_2' > 0$:

$$\mathbb{P}\Big(\sum_{i=1}^{k_1' \mu n'} |t_i| < k_2' n'^2\Big) \sim \frac{\sqrt{k_2'}}{\sqrt{\pi} k_1' \mu} \exp(-\frac{(k_1'\mu)^2}{2k_2'}).$$

Therefore, for any $\varepsilon_6' > 0$ we can find small enough $k_2' > 0$ such that

$$\mathbb{P}\Big(\sum_{i=1}^{S_{n'}} |t_i| > k_2' n'^2\Big) > 1 - \varepsilon_6'. \tag{2.46}$$

By Lemma 2.5.2, for large N and any $\varepsilon_7' > 0$ we can find $\delta > 0$ and a coupling of $\mathcal{T}^N$ and $\mathcal{T}^\infty$, such that the tree $\mathcal{T}^N[n]$ is same to $\mathcal{T}^\infty[n]$ with probability at least $(1 - \varepsilon_7')$. There exists an image of γ_n in $\mathcal{T}^N$, and all trees that hang off $\gamma_{n'}$ do not go deeper than generation Z_n w.h.p. Hence for any $\varepsilon_8' > 0$ and large N there exist $k_2', K_2' > 0$ such that in the conditioned tree $\mathcal{T}^N$:

$$\mathbb{P}\Big(\frac{1}{n'^2} \sum_{i=1}^{S_{n'}} |t_i| \in [k_2', K_2']\Big) > 1 - \varepsilon_8'. \tag{2.47}$$

Now we are ready to prove that the extra edge is attached to the subtree t_0 in $\mathcal{T}^N$ w.h.p. The probability of the edge e to be attached to a vertex in the tree t_0 can be bounded as follows:

$$\mathbb{P}(e \in t_0) > \mathbb{P}\Big(e \in t_0 \;\Big|\; \frac{|\mathcal{T}^N_{+e}\backslash t_0|}{|\mathcal{T}^N_{+e}|} < \delta'\Big)\mathbb{P}\Big(\frac{|\mathcal{T}^N_{+e}\backslash t_0|}{|\mathcal{T}^N_{+e}|} < \delta'\Big). \tag{2.48}$$

Since the attachment is uniform, then it automatically follows:

$$\mathbb{P}\Big(e \in t_0 \;\Big|\; \frac{|\mathcal{T}^N_{+e}\backslash t_0|}{|\mathcal{T}^N_{+e}|} < \delta'\Big) > 1 - \delta'.$$

Using (2.41) and (2.47) we have

$$\mathbb{P}\Big(\frac{|\mathcal{T}^N_{+e}\backslash t_0|}{|\mathcal{T}^N_{+e}|} < \frac{K_2'\delta^4}{h_1}\Big) > 1 - \varepsilon_6' - \varepsilon_8',$$

and since δ can be chosen to be small enough, we have that for any $\varepsilon_9' > 0$ there exist $\delta' > 0$ and large N such that

$$\mathbb{P}(e \in t_0) > 1 - \varepsilon_9'.$$

Now since the statement of the theorem about the fraction of $\kappa(\mathcal{T}^N_{+e}, s)$ and $|\mathcal{T}^N_{+e}|$, the larger volume of the tree can decrease this fraction, thus we need to prove that the probability to have a very large tree $\mathcal{T}^N$ in N^2 scale is also small. Consider a large $H_1 > 0$, then by (2.39) we have:

$$\begin{aligned}\mathbb{P}(|\mathcal{T}^N_{+e}| > H_1N^2) &< C_2 \sum_{m>H_1N^2} e^{-c_1\frac{N^2}{m}} Nm^{-3/2} \\ &< C_2 \int_{H_1N^2}^{\infty} e^{-c_1\frac{N^2}{m}} Nm^{-3/2}dm = C_2 \int_{H_1}^{\infty} e^{-\frac{c_1}{x}} x^{-3/2}dx = \\ &= \frac{C_2}{\sqrt{c_1}} \int_0^{c_1/H_1} e^{-y}y^{-1/2}dy = \frac{C_2}{\sqrt{c_1}}\gamma\left(1/2, \frac{c_1}{H_1}\right)\end{aligned}$$

and, by Lemma 2.2.1, for large H_1 we have

$$\mathbb{P}(|\mathcal{T}^N_{+e}| > H_1N^2) < C_2H_1^{-1/2}.$$

Hence, for any $\varepsilon_{10}' > 0$ we can find $H_1 > 0$, such that

$$\mathbb{P}(|\mathcal{T}^N_{+e}| < H_1N^2) > 1 - \varepsilon_{10}'. \tag{2.49}$$

We have proven that all the trees $t_1, \ldots, t_{\nu_{n'}}$ are short enough and have small total volume comparative to the total size of the tree $\mathcal{T}^N_{+e}$ w.h.p. Now use the formula (2.32) for $\kappa(\mathcal{T}^N_{+e}, s)$.

In this case, if the other endpoint of e is in t_0, then the largest subtree is in t_0 w.h.p. then if we cut off the whole t_0 we have the lower bound on $\kappa(\mathcal{T}^N_{+e}, s) > |\mathcal{T}^N_{+e}| - |t_0| = |\gamma_{n'}| + \sum_{i=1}^{S'_n} |t_i| > \sum_{i=1}^{S'_n} |t_i|$ given in (2.46) and the volume of the whole graph is bounded above in (2.49). Combining those bounds we obtain

$$\mathbb{P}\Big(\frac{\kappa(\mathcal{T}^N_{+e}, s)}{|\mathcal{T}^N_{+e}|} > \frac{k'_2\delta^4}{H_1}\Big) > 1 - \varepsilon'_8 - \varepsilon'_{10}.$$

Hence, choosing δ small enough we have that for any $\varepsilon'_{11} > 0$ we can find small enough $\delta'' > 0$ and large N such that

$$\mathbb{P}\Big(\frac{\kappa(\mathcal{T}^N_{+e}, s)}{|\mathcal{T}^N_{+e}|} > \delta''\Big) > 1 - \varepsilon'_{11}.$$

This finishes 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 [34]). This means that, for $\lambda = 0$, with a decent positive probability the SI spreading will encounter bottlenecks

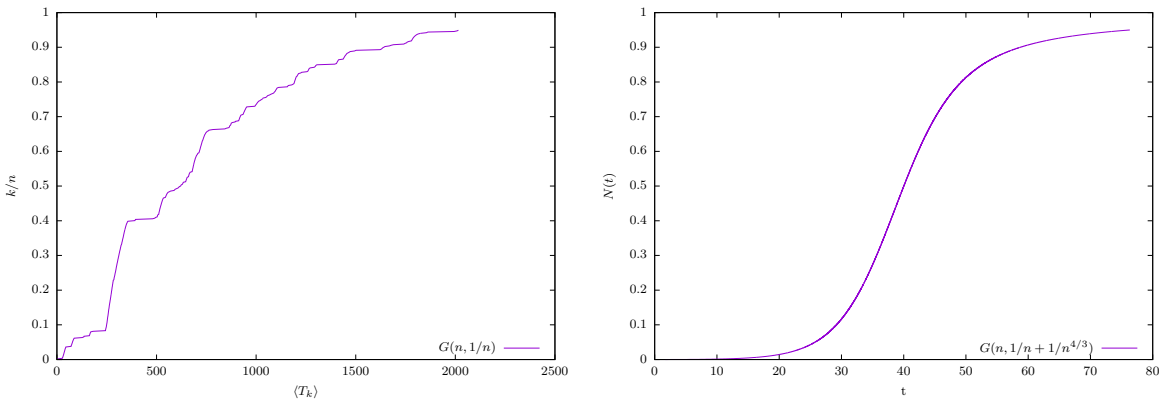
everywhere during the process, and, because of these bottlenecks at random locations, the averaged spreading curve will not converge and will produce jumps. However, at large $\lambda > 0$, a typical largest cluster will have one or more extra edges, hence, on a typical realization of the cluster, we expect to see the smoothing effect.

In order to apply our Theorem 2.4.5, we need to prove that for the largest connected cluster $\mathcal{C}_{(n)}$ of $G_p(n, p_n(\lambda))$ the $\kappa(\mathcal{C}_{(n)}, s)$ is typically positive for these clusters. Of course, in this case we have to choose the root from the 2-core - the maximum induced subgraph of $\mathcal{C}_{(n)}$ which has minimum degree two. 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 $n^{2/3}$.

This structural result is essentially provided by [1,2], where the scaling limit of large near-critical clusters as metric spaces is described, built on the 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 $k \geq 2$, the 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 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 spreading on temporal networks

3.1 Introduction

Recently much effort has been made to understand the main governing factors of spreading on networks [19,31,39]. Topological aspects like community structure [26] and temporal activity patterns like burstiness [9] severly influence the speed of spreading [32, 39, 50].

Different approaches have been applied to tackle this problem. One direction is to treat related models analytically and by numerical simulations [19, 31, 32, 35, 46, 69]. 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 [53,69] records of face to face encounters [61] or mobile call billing information [39, 50].

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 [55] that in mobile call networks subscribers 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 a 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 [56]. 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 exists 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, 39]) that homogeneous interaction patterns are modeled 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 described 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 (periodic) [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 the deterministic SI spreading process on the temporal network $\mathcal{G}$ as a dynamic process on the underlying network Γ. Let us recall 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 takes place 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 modeling 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. [64, 73]). The infinite square lattice with next nearest neighbour interactions (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 two-dimensional *finite* square lattice with NNN (Z^2) is defined as an induced sub-network of an infinite lattice on n^2 nodes $V(Z^2) = \{[i, j] : 1 \leqslant i, j \leqslant n\}$.

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 [45] as a model of social networks with 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 $O(\log n)$ [71]. 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 of edges 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 $n \to \infty$ [40]. The important properties of all four above models are summarized in the Table 3.1. Degree distributions of networks Z^2, ER and BA are known, when the degree distribution of network K can be well described by the log-normal distribution and it is stable with respect to the size of the network.

	Degree distribution	Small-world	Percolation threshold	Communities
Z^2	$Constant$	-	+	-
ER	$Poisson$	+	+	-
BA	$Power-law$	+	-	-
K	$Log-normal$	+	+	+

Table 3.1: *Main properties of the considered network models.*

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.

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,

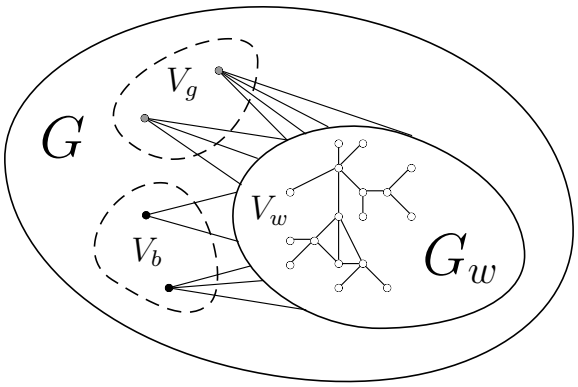


Figure 3.1: *Schematic structure of networks G_w and G. White nodes are denoted as V_w, grey nodes as V_g and black nodes as V_b.*

periodic boundary conditions in time are applied for the data - when the process reaches the end of the data time span it immediately proceeds again from the beginning of the time span. Since the dataset is large enough, the burstiness coefficient is not much affected by this conditioning. We assume that when a call takes place, 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 Figure 3.1.

The grey and black nodes of degree $k \geqslant 2$ in the network G 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 each network G by removing all grey and black nodes of degree 1. Then we consider the largest connected component $LCC(G_w)$ in the network G_w and $LCC(G)$ in the network G. We denote as $\langle k_w \rangle$ the average degree of white nodes.

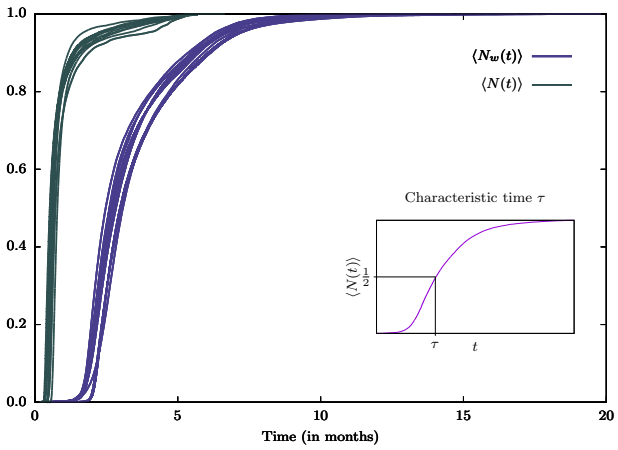


Figure 3.2: *Average spreading curves $\langle N_w(t)\rangle$ (dark blue) for 10 typical networks of cities G_w with average degree $\langle k_w\rangle = 4 \pm 0.2$. Respective average spreading curves $\langle N(t)\rangle$ for networks of same cities G are also given in dark grey. The inset illustrates 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 $LCC(G_w)$ or $LCC(G)$ 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 in Figure 3.2 along with the illustration of the notion of characteristic time.

3.4 Empirical results

We consider the sample of $N \simeq 300$ cities with the population between 10,000 and 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 (see Figure 3.4, (a)), however, the dependence of the size on the population of cities is vague (see Figure 3.4, (b)), which is explained by differences in the coverage of the company in different cities. Thus, it follows from the data that the number of bridges, or the total number of grey and black nodes $|V_g| + |V_b|$, is also

proportional to the size of the networks G, such that

$$|V_g| + |V_b| \simeq 5|V_w|, \tag{3.1}$$

which is further used in Section 3.5.

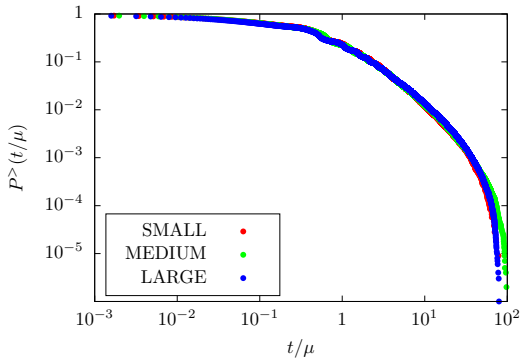


Figure 3.3: *The inter-event time distributions of three cities of different sizes of V_w. A small sized city has $|LCC(G_w)| \sim 10^3$, a medium sized city has $|LCC(G_w)| \sim 10^4$ and a large size city has $|LCC(G_w)| \sim 10^5$.*

The average inter-event time for each city i, where $1 \leqslant i \leqslant N$, is denoted 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 tail distributions $P^{>}(t/\mu) = \int_t^\infty p(s/\mu)ds$ of inter-event times in the cities are found to be similar for cities of different population size (see Figure 3.3), which is explained by the fact that people follow same patterns of calls irrespective of the size of their community.

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 and observe that the cities with similar average degrees $\langle k_w \rangle$ have similar degree distributions. In Figure 3.5, (a), a sample of tail distributions of node degrees $P^{>}(k) = \sum_k^\infty p(s)ds$ is presented. It is also seen that spreading curves have the same shape for the cities with similar average degrees. The sample of average spreading curves for the same cities with $\langle k_w \rangle = 3 \pm 0.2$ is given in Figure 3.5, (b), along with the illustration of the notion of characteristic time.

In agreement with our expectations we see a radical decrease of the characteristic times,

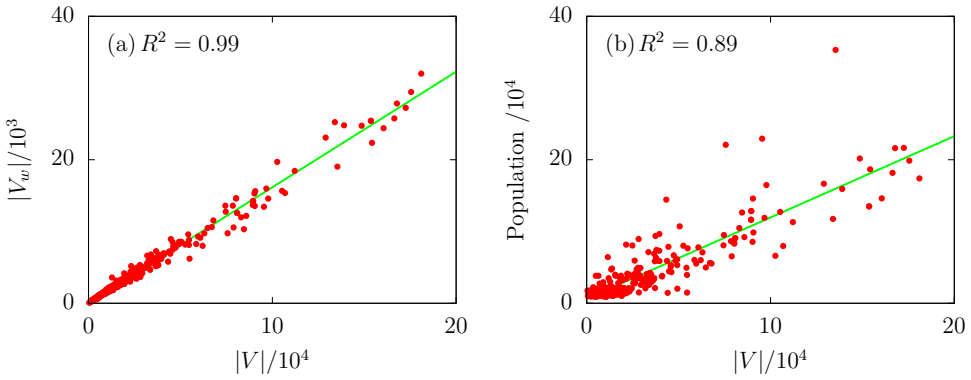
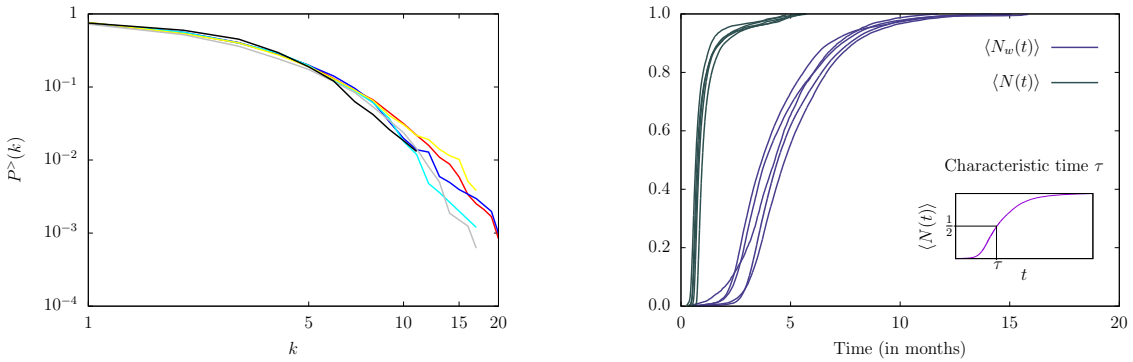


Figure 3.4: *(a) Dependence between the size of the network G and the number of white nodes $|V_w|$ for each city. (b) The relationship between the size of the network G and the population for each city. The linear trend is vague, which represents the difference in coverage proportion of the company in each city.*

when we switch from G_w to G (Figure 3.6). Remember that P_∞ is the relative size of the $LCC(G_w)$ to the total size of G_w. Then on the insets of Figure 3.6 we see that the dependence of $\langle k_w \rangle$ over the P_∞ shows the typical percolation behaviour. We observe when $P_\infty < 20\%$, the networks show scattered results in characteristic time (see Figure 3.6, (a)), thus we perform a truncation of those cities and present the results in Figure 3.6, (b).

We find a clear power-law dependence of the characteristic time τ_w on the average degree $\langle k_w \rangle$ in networks G_w (see Figure 3.6, blue plots). At 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 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 (see Figure 3.6, red plots). However, the dependence of the characteristic time τ on the average degree $\langle k_w \rangle$ in networks G is less apparent. 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 effect of the topology of the underlying network. In the next section we model this effect on the null network models.



(a) Degree distributions *(b) Spreading curves*

Figure 3.5: *(a) Tail distributions of node degrees and (b) average spreading curves $\langle N_w(t)\rangle$ (dark blue) for 6 typical networks of cities G_w with average degree $\langle k_w\rangle = 3 \pm 0.2$ are given. Respective average spreading curves $\langle N(t)\rangle$ for networks G of same cities are also given in dark grey in (b). The inset of (b) illustrates the definition of the characteristic time τ.*

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 note from the data that the number of white nodes is proportional to the total number of bridges, which is given by (3.1).

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: 1) the *power-law distribution* $pow(t_{min}, \alpha)$ with the following density function:

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

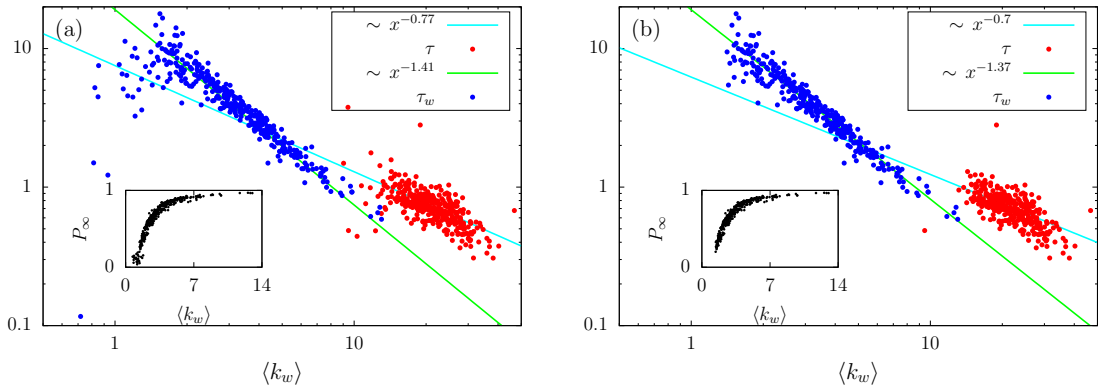


Figure 3.6: *The characteristic times τ_w (blue dots) and τ (red dots) relative to average degree of white nodes $\langle k_w \rangle$ in the networks G_w and G. The insets show the dependence of P_∞ on the average degree of white nodes $\langle k_w \rangle$. (a) The whole sample of cities is plotted, with large scattering around $\langle k_w \rangle \sim 1$. In this case the networks become very sparse and the size of LCC is comparatively small. (b) The same plot is obtained by deleting those cities with $P_\infty < 20\%$, which gives clearer results.*

and 2) the *shifted exponential distribution* $exp(t_{min}, \lambda)$ with the following density function:

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

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

3.5.2 Simulation

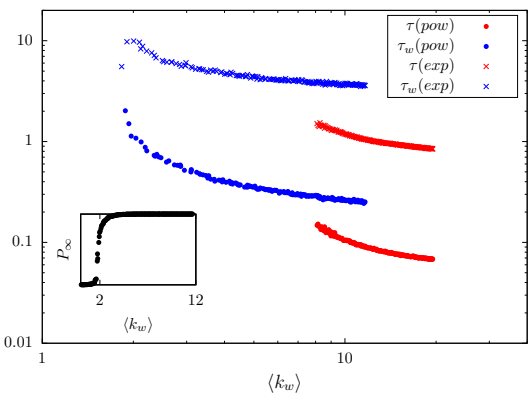
The simulation is performed on networks $\overline{G}_w^p$ and $\overline{G}^p$ constructed from a network $\overline{G}$, given by four network models defined earlier: 1) Finite square lattice with NNN (Z^2); 2) Erdős-Rényi model (ER); 3) Barabási-Albert model (BA); 4) the model of Kumpula et.al. (K). Each considered network model has $n = 5000$ nodes (network Z^2 has 4900 nodes) and the parameters of these models are tuned in such a way that the average degree d 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, and in each run the starting node is chosen at random from the set of white nodes in $LCC(\overline{G}_w^p)$ of degree greater or equal to two. The first set of runs has power-law inter-event times distribution on edges with $t_{min} = 0.008$ and $\alpha = 1.2$, and the second set has shifted exponential distribution

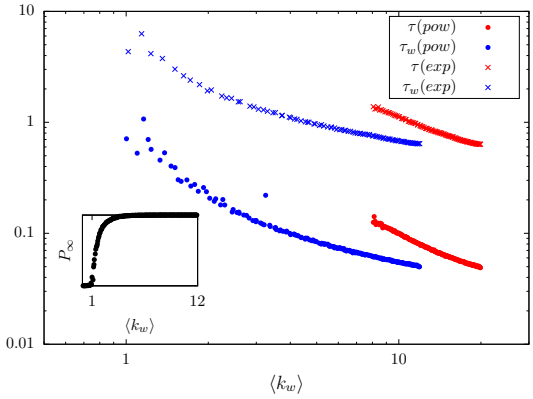
with the same t_{min} and the parameter λ is chosen in the way that two distributions have same mean μ.

The results are presented in Figure 3.7. All four network models show faster spreading with bursty interaction patterns modeled by the power-law inter-event times ($\tau_w(pow)$ and $\tau(pow)$) compared to the process with shifted exponential inter-event times ($\tau_w(exp)$ and $\tau(exp)$), which was also captured in [35]. The introduction of bridges decreases characteristic time in all models, which agrees with the empirical results in Section 3.4. It is observable that the large diameter and regularity in the structure of the underlying network influence much the speed of spreading.

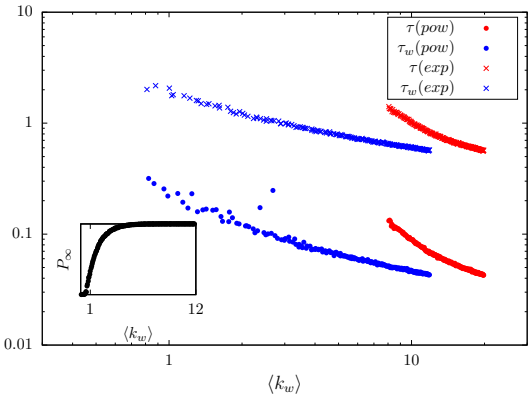
Following the observation of bursty activity patterns, in Figure 3.8 we show the comparison between the characteristic times τ_w and τ for four considered models with power-law inter-event times. We observe that BA model is the fastest in terms of characteristic time, which is explained by the presence of large degree nodes and since the percolation threshold is zero, these nodes remain in the diluted network as well. It is known that the presence of a community structure slows down spreading (see e.g. [72]). Therefore, in the model K 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, but since it has more of large degree nodes, it overtakes ER in the spreading speed. The model Z^2 is the slowest due to its regular structure and large diameter. The introduction of bridges smears the topology of initial networks and thus spreading results in almost the same behaviour for all four models.



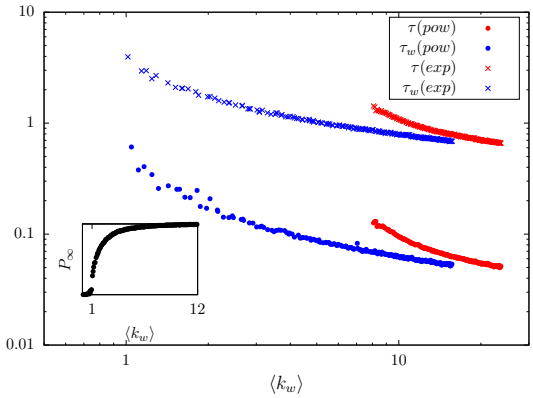
(a) Regular lattice with NNN



(b) Erdős-Rényi network



(c) Barabási-Albert network



(d) Network of Kumpula et.al.

Figure 3.7: *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 shifted exponential inter-event times (×) are considered. The following models are considered: (a) Square lattice with NNN interactions; (b) Erdős-Rényi network; (c) The Barabási-Albert network; (d) Network of Kumpula et.al. The data are plotted only for those values of $\langle k \rangle$ when the $LCC(\overline{G}_w^p)$ exists and the results show that the process with shifted exponential inter-event times is slower on each model. The introduction of bridges decreases the characteristic time and the decrease depends on the initial topology of the network.*

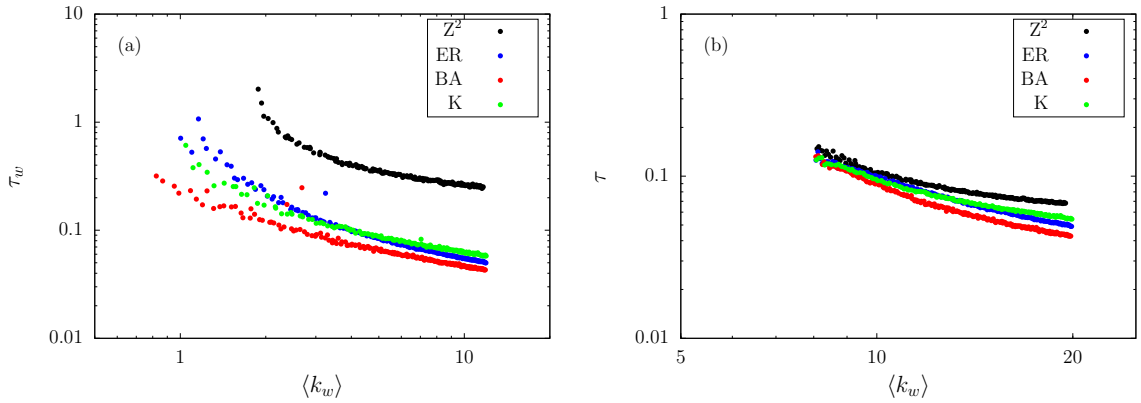


Figure 3.8: *Comparison of characteristic times τ_w and τ relative to average degree of white nodes $\langle k_w \rangle$ in the simulated networks $\overline{G}_w^p$ and $\overline{G}^p$ with four different initial networks $\overline{G}$. Each network has 5000 nodes and 25000 bridges are added and power-law distributed inter-event times are used. The data are shown only for those values of $\langle k \rangle$ when the $LCC(\overline{G}_w^p)$ exists and 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 takes place when the average degree of a network is large. The 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 a low average degree. In this case the LCC has a low number of cycles and the network locally has the structure of a tree. In Chapter 2 we have derived that the 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. Basing on this model, we have found evidence that the topology of the underlying network has effect in spreading. According to our expectations the presence of large degree nodes and absence of communities increases the speed of SI spreading. When the average degree of the network is less than one, it falls apart into a number of disconnected components, each of which having a tree-like structure. The introduction of bridges makes this network connected, but topology becomes homogeneous irrespective of 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 the 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 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 $n/2$ vertices 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 2.2.7 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, by dominating the original process with a delayed process that has d active edges in the front of epidemic and by constructing the appropriate process Q we can prove that the expected time to infect k vertices is bounded for the case of $k < \kappa_d(G, s)$, where $\kappa_d(G, s)$ is a uniquely defined number for any rooted graph G. However, this means we would approximate the original process by spreading on the d infinite rays that have common root. This construction is more complex and it is hard to find an example of tree-like graphs where such $\kappa_d(G, s)$ 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. Random cutting and records in deterministic and random trees. *Random Structures & Algorithms*, 29(2):139–179, 2006.

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

[35] 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.

[36] 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.

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

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

[39] 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.

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

[41] 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.

[42] 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.

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

[44] 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.

[45] 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.

[46] 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.

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

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

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

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

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

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

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

[54] 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.

[55] 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.

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

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

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

[59] B. Roehner and P. Winiwarter. Aggregation of independent paretian random variables. *Advances in Applied Probability*, 17(2):465–469, 1985.

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

[61] 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.

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

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

[64] 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.

[65] 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.

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

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

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

[69] 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.

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

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

[72] 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.

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