On some diffusion and spanning problems in configuration model

Kumar Gaurav

UPMC SORBONNE UNIVERSITÉS

THÈSE DE DOCTORAT DE
l'UNIVERSITÉ PIERRE ET MARIE CURIE

Spécialité : **Mathématiques Appliquées**
(École doctorale: Sciences mathématiques de Paris-Centre (ED 386))

Sur certains problèmes de diffusion et de connexité dans le modèle de configuration

Présentée par **Kumar GAURAV**

pour obtenir le grade de
Docteur de l'Université Pierre et Marie Curie

soutenue le 18 November 2016, devant le jury composé de :

M. Bartłomiej BŁASZCZYSZYN	Directeur de thèse
M. Laurent DECREUSEFOND	Rapporteur
M. Armand MAKOWSKI	Rapporteur
Mme. Irina KOURKOVA	Examinateur
M. Viet Chi TRAN	Examinateur

To my loving parents

Acknowledgements

First and foremost, I would like to thank my PhD advisor, Bartłomiej Błaszczyszyn, not only for his guidance in my research, but also for his infinite patience and motivation during the writing of my thesis. I was really fortunate to have him introduce me to the world of research.

I would also like to express my sincere gratitude to Laurent Decreusefond and Armand Makowski, who took time out of their busy schedules to thorougly review my thesis, and provid insightful comments for improving my manuscript as well as expanding on my research. I would also like to thank Irina Kourkova and Viet Chi Tran for accepting to be a part of my PhD jury.

I had the good fortune to be a part of Inria DYOGENE team during my PhD, and I would like to acknowledge the group's stimulating environment for independent thinking. I thank my labmates, in particular, Ana, Anne, Christelle, Emilie, Francois, Marc, Matthieu, Miodrag, Paul, Pierre, Remi - for all the discussions we had, related to research or otherwise. I also thank Helene Milome for her kind help in administrative tasks.

Lastly, I would like to express my love and gratitude for my parents, for always being there. I also thank my brothers for supporting me during the difficult phases of my PhD. And a very special thanks is reserved for my love, Sanu, for bringing a joy to the present, and a hope for the future.

Résumé

Un certain nombre de systèmes dans le monde réel, comprenant des agents interagissant, peut être utilement modélisé par des graphes, où les agents sont représentés par les sommets du graphe et les interactions par les arêtes. De tels systèmes peuvent être aussi divers et complexes que les réseaux sociaux (traditionnels ou virtuels), les réseaux d'interaction protéine-protéine, internet, réseaux de transport et les réseaux de prêts interbancaires. Une question importante qui se pose dans l'étude de ces réseaux est: dans quelle mesure, les statistiques locales d'un réseau déterminent sa topologie globale. Ce problème peut être approché par la construction d'un graphe aléatoire contraint d'avoir les mêmes statistiques locales que celles observées dans le graphe d'intérêt. Le modèle de configuration est un tel modèle de graphe aléatoire conçu de telle sorte qu'un sommet uniformément choisi présente une distribution de degré donnée. Il fournit le cadre sous-jacent à cette thèse.

En premier lieu nous considérons un problème de propagation de l'influence sur le modèle de configuration, où chaque sommet peut être influencé par l'un de ses voisins, mais à son tour, il ne peut influencer qu'un sous-ensemble aléatoire de ses voisins. Notre modèle étendu est décrit par le degré total du sommet typique et le nombre de voisins il est capable d'influencer. Nous donnons une condition stricte sur la distribution conjointe de ces deux degrés, qui permet à l'influence de parvenir, avec une forte probabilité, à un ensemble non négligeable de sommets, essentiellement unique, appelé la composante géante influencée, à condition que le sommet de la source soit choisi à partir d'un ensemble de bons pionniers. Nous évaluons explicitement la taille relative asymptotique de la composant géante influencée, ainsi que de l'ensemble des bons pionniers, à condition qu'ils soient non-négligeable. Notre preuve utilise l'exploration conjointe du modèle de configuration et de la propagation de l'influence jusqu'au moment où une grande partie est influencée, une technique introduite dans Janson et Luczak (2008). Notre modèle peut être vu comme une généralisation de la percolation classique par arêtes ou par sites sur le modèle de

configuration, avec la différence résultant de la conductivité orientée des arêtes dans notre modèle. Nous illustrons ces résultats en utilisant quelques exemples, en particulier, motivés par le marketing viral - un phénomène connu dans le contexte des réseaux sociaux.

Ensuite, nous considérons les sommets isolés et les arêtes longues de l'arbre couvrant minimal du modèle de configuration dont les arêtes sont indépendemment pondérée par des nombres non-négatifs interprétés comme des longueurs. En utilisant la méthode de Stein-Chen, nous calculons la distribution asymptotique du nombre de sommets qui sont séparés du reste du graphe par une certaine distance critique, par exemple alpha. Cette distribution donne la loi d'échelle de la longueur de la plus longue arête du graphe de plus proche voisin. Nous utilisons ensuite les résultats de Fountoulakis (2007) sur la percolation pour démontrer que, après la suppression de toutes les arêtes d'une longueur supérieure à alpha, le sous-graphe obtenu est connexe, sauf pour les sommets isolés. Cela nous amène à conclure que l'arête la plus long de l'arbre couvrant minimal et celle du graphe de plus proche voisin coïncident avec une forte probabilité.

Enfin, nous étudions une question plus générale, à savoir si une certaine comparaison basée sur des statistiques locales du graphe conduirait à la comparaison des propriétés topologiques globales, de sorte que des résultats pour certains graphes plus complexes pourraient être obtenus par leur comparaison à des graphes plus simples à étudier. À cette fin, nous introduisons un ordre convexe sur les graphes aléatoires et nous discutons des implications, notamment la façon dont cet ordre peut conduire à la comparaison des probabilités de percolation dans certaines situations.

Abstract

A number of real-world systems consisting of interacting agents can be usefully modelled by graphs, where the agents are represented by the vertices of the graph and the interactions by the edges. Such systems can be as diverse and complex as social networks (traditional or online), protein-protein interaction networks, internet, transport network and inter-bank loan networks.

One important question that arises in the study of these networks is: to what extent, the local statistics of a network determine its global topology. This problem can be approached by constructing a random graph constrained to have some of the same local statistics as those observed in the graph of interest. One such random graph model is *configuration model*, which is constructed in such a way that a uniformly chosen vertex has a given degree distribution. This is the random graph which provides the underlying framework for this thesis.

As our first problem, we consider propagation of influence on configuration model, where each vertex can be influenced by any of its neighbours but in its turn, it can only influence a random subset of its neighbours. Our (enhanced) model is described by the total degree of the typical vertex and the number of neighbours it is able to influence. We give a tight condition, involving the joint distribution of these two degrees, which allows with high probability the influence to reach an essentially unique non-negligible set of the vertices, called a *big influenced component*, provided that the source vertex is chosen from a set of *good pioneers*. We explicitly evaluate the asymptotic relative size of the influenced component as well as of the set of good pioneers, provided it is non-negligible. Our proof uses the joint exploration of the configuration model and the propagation of the influence up to the time when a big influenced component is completed, a technique introduced in Janson & Luczak (2008). Our model can be seen as a generalization of the classical Bond and Node percolation on configuration model, with the difference stemming from the oriented conductivity of edges in our model.

We illustrate these results using a few examples which are interesting from either theoretical or real-world perspective. The examples are, in particular, motivated by the viral marketing phenomenon in the context of social networks.

Next, we consider the isolated vertices and the longest edge of the minimum spanning tree of a weighted configuration model. Using Stein-Chen method, we compute the asymptotic distribution of the number of vertices which are separated from the rest of the graph by some critical distance, say α. This distribution gives the scaling of the length of the longest edge of the nearest neighbour graph with the size of the graph. We then use the results of Fountoulakis (2007) on percolation to prove that after removing all the edges of length greater than α, the subgraph obtained is connnected but for the isolated vertices. This leads us to conclude that the longest edge of the minimum spanning tree and that of nearest neighbour graph coincide w.h.p. .

Finally, we investigate a more general question, that is, whether some ordering based on local statistics of the graph would lead to an ordering of the global topological properties, so that the bounds for more complex graphs could be obtained from their simplified versions. To this end, we introduce a *convex order* on random graphs and discuss some implications, particularly how it can lead to the ordering of percolation probabilities in certain situations.

Overview of the thesis

The desire for understanding the mechanics of complex networks, describing a wide range of systems in nature and society, motivated many applied and theoretical investigations of the last two decades. Erdős-Rényi random graph, which had been introduced much earlier in the 60's, regained importance in this context since it exhibits some of the same global properties as observed in real-world networks, such as,

- phase-transition: At some critical value, a small change in the parameters determining the local interaction of agents can precipitate a huge change in the global network structure, for example, in the level of connectivity, and
- *small-world* property: average topological distance between any two vertices of the graph varies very slowly with the size of the graph (typically on logarithmic scale).

The analysis for this type of random graphs is generally quite tractable. This prompted the introduction of progressively more structure and complexity on top of these random graphs in order to better approximate real-world networks, while keeping the analysis tractable. The random graph model that forms the underlying framework in this paper is configuration model, which can be thought of as a generalization of the classical Erdős-Rényi random graph. Given a vertex set $[n] = 1, 2, \ldots, n$ and a sequence of non-negative integers, $\mathbf{d}^{(n)} = (d_i)_1^n$, with $\sum_{i=1}^n d_i$ even, the multi-graph version of configuration model, $G^*(n, (d_i)_1^n)$, is constructed simply by giving d_i half-edges to each vertex i, and then randomly matching all the half-edges. We will discuss the construction in more detail in Chapter 1.

One of the most widely studied problems on a random graph concerns the *percolation*, a reason being that a lot of practically relevant problems can be framed in terms of some percolation problem. Broadly speaking, study of percolation on a graph involves analysing the change in the global topology of the graph

with the removal of edges or nodes in a certain manner. The phenomenon of percolation in the context of configuration model is the major unifying theme recurring throughout the analysis in this thesis.

We give below a brief summary of the thesis.

Chapter 1. *Random Graphs: an Overview.* In this chapter, we introduce some of the essential notions in the theory of random graphs. We start by recalling some results on Galton-Watson branching process, which approximates to some extent, the process of exploration of a connected component in a random graph. The phase-transition in the extinction probability of this branching process anticipates the phase-transition in the existence of a giant connected component, a component which scales linearly with the size of the graph, in random graphs. We recall this fundamental result for both Erdős-Rényi graph and configuration model. Its analysis, particularly in the latter case, heavily influences our analysis in Chapter 2. We also recall the results on the total connectivity of both Erdős-Rényi graph and configuration model. In the former case, the result is another instance of phase-transition, and its analysis influences our general approach in Chapter 4, while the result in the latter case is generalized in the course of Chapter 4. Next, we recall some results on percolation in configuration model, which can be seen as a special case of the diffusion dynamic of Chapter 2 as we show in Chapter 3, but more importantly, it is directly relevant to our analysis in Chapter 4. We finish by recalling the definition of *Local Weak Convergence*, which underlies our elementary discussion in the last chapter on convex ordering in random graphs.

Chapter 2. *Viral Marketing in Configuration Model.* A problem closely related to that of percolation concerns the diffusion of information in the graph. One common assumption in such a context is that the information flows from one agent to another, independently of their neighbors (sometimes referred to as *independent cascade*). Under this assumption, there is no statistical difference between the set of agents who have a huge influence in the graph and the set of agents who are most susceptible to being influenced. It is a reasonable assumption when modelling the spread of an epidemic in a population, but not for modelling the viral marketing phenomenon in online social networks, where the aforementioned sets of agents would have markedly different local statistics. In this chapter, we generalize the diffusion mechanism on a random graph to more closely mimic that of viral marketing, which leads to the topological separation of *influencer* and *susceptible* agents.

More precisely, we consider propagation of influence on configuration model, where each vertex can be influenced by any of its neighbours but in its turn, it can only influence a random subset

of its neighbours. Our (enhanced) model is described by the total degree, of a uniformly chosen vertex and the number of neighbours it is able to influence. We give a tight condition, involving the joint distribution of these two degrees, which allows with high probability the influence to reach an essentially unique non-negligible set of *susceptible* vertices, provided that the source vertex is chosen from a set of *good pioneers* or *influencers*. We explicitly evaluate the asymptotic relative size of the influenced component of susceptible agents as well as of the set of influencers, provided it is non-negligible. This latter (we believe technical) assumption allows us to identify the set of influencers with a "big source component" in a dual process with the "reversed" dynamic. This dual process is not required in the case of independent cascade since one could identify the set of influencers with the big susceptible component itself. The results here roughly imply that under certain conditions, each influencer agent can influence all the susceptible agents, while each susceptible agent can be influenced by all the influencers.

We study both the forward and the dual propagation process using the joint exploration of the configuration model and the propagation of the influence up to the time when a big influenced component is completed, a technique introduced in Janson & Luczak (2008) [38]. Our model can be seen as a generalization of the classical Bond and Node percolation on configuration model, with the difference stemming from the oriented conductivity of edges in our model.

This chapter is based on [13], a joint work with Bartłomiej Błaszczyszyn. We thank René Schott for introducing us to the influence propagation dynamic analysed in this paper through a pre-print of [20] and thus motivating this study. We also thank Marc Lelarge for his useful suggestions regarding the analytical tools for exploration on configuration model and pointing us to [38], which has heavily influenced our approach.

Chapter 3. *Viral Marketing: Examples, Applications and Numerical Studies.* In this chapter, we illustrate the results of the previous chapter using a few examples. In particular, we consider two kinds of underlying graphs - one where the total degree of an agent is assumed to have Poisson distribution, and the second where it is assumed to have Power-law distribution. We recall that Poisson distribution approximates the degree distribution in Erdős-Rényi graph, while Power-law degree distribution is often observed in real-world networks. On these underlying graphs, the agents are assumed to have varying attitudes towards propagating the information. We analyze three cases, in particular — (1) Bernoulli transmissions, when a member influences each of its friend with

probability p; (2) Node percolation, when a member influences all its friends with probability p and none with probability $1-p$; (3) Coupon-collector transmissions, when a member randomly selects one of his friends K times with replacement.

We frame the above illustrations in the context of decision-making process of a firm looking to start an online marketing campaign.

This chapter is based on [33], with Bartłomiej Błaszczyszyn and Paul Keeler.

Chapter 4. *Isolated Vertices and the Longest Edge of the Minimum Spanning Tree of Weighted Configuration Model.* In this chapter, we study a problem closely related to the question of survival of a population in case of an epidemic spread. A population which is too closely knit together socially has a smaller chance of survival than a population which has at least a few *socially isolated* individuals. To model this, we introduce i.i.d. weights to the edges of totally connected configuration model (which can be thought of as representing social distance between members of a species). Using Stein-Chen method in a manner similar to that of Lindvall (1992) [44] in the context of Erdős-Rényi graph, we compute the asymptotic distribution of the number of vertices which are separated from the rest of the graph by some critical distance, say α. This distribution gives the scaling of the length of the longest edge of the nearest neighbour graph with the size of the graph. Now, the subgraph obtained after removing all the edges of length greater than α is equivalent to that obtained after bond percolation with probability π_n^α. This identification allows us to use the results of Fountoulakis (2007) [29] on percolation to prove that the resultant subgraph (after edge removal) is connnected, but for the isolated vertices. This leads us to conclude that the longest edge of the minimum spanning tree and that of nearest neighbour graph coincide w.h.p. .

This chapter is based on work in progress with Bartłomiej Błaszczyszyn. This study was motivated by similar results obtained by Penrose (1997) [51] in the context of Gilbert graph in Euclidean space. We thank Remco van der Hofstad for his useful suggestions and his notes [54] which helped us greatly in the proof of connectivity of configuration model after removing edges exceeding a certain length.

Chapter 5. *Future Work: Convex comparison of Random Graphs.* In this chapter, we motivate the introduction of stochastic ordering, in particular, convex ordering in the context of random graphs, for future work. In particular, we demonstrate how the convex order can lead to the ordering

of percolation probabilities in configuration model. This discussion is inspired by the successful application of convex order in the context of point processes in euclidean space by Błaszczyszyn and Yogeshwaran ([14], [12]).

Notation

Throughout the thesis, we will deal with the notion of *convergence in probability*: a sequence X_n of random variables is said to converge in probability to a limiting random variable X when, for every $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0. \tag{1}$$

We write this as $X_n \xrightarrow{\text{p}} X$.

An event happening *with high probability* implies that it happens with probability tending to 1 as $n \to \infty$. We will generally abbreviate 'with high probability' by 'whp'. o_p and O_p notation is also used in a standard way. For example, $X_n = o_p(n)$ means that $X_n/n \xrightarrow{\text{p}} 0$.

For a set A, we will denote the number of elements in A by $|A|$. For a graph G, we usually refer to the number of vertices in G as the size of the graph. In this case, $|G|$ will denote the number of vertices in G, that is, the size of G.

Contents

1

Random Graphs: an Overview

We start this chapter by introducing Galton-Watson branching process in Section 1.1. Branching process approximation of random graphs, even when not used for proofs explicitly, is quite useful in suggesting some global properties of random graphs, particularly those relating to phase-transition. Next in Section 1.2, we introduce (arguably) the simplest of random graphs, Erdős-Rényi Graph, which consists of n nodes with each pair of nodes independently connected with probability p. We will denote the resulting graph by $ER(n,p)$ throughout the thesis. This graph exhibits a very interesting phase transition in the size of its largest connected component as p increases. Throughout the thesis, whenever the number of vertices in a connected component scale linearly with the size of the graph, we will refer to it as the giant component. The results as well as the tools developed for proofs in this case can usually be extended to more general random graphs with some modifications. In Section 1.3, we introduce one such general random graph - configuration model, which provides the underlying framework for the rest of the thesis. We describe the phase-transition in its largest connected component in Section 1.3.1, the proof of which will guide our proofs in Chapter 2. We describe some simple ways of diffusion, including percolation,

on this model in Section 1.3.2. We will generalize the bond percolation introduced here in Chapters 2 and 3, and also use the results for bond percolation directly in the proofs of Chapter 4. $ER(n,p)$ exhibits another phase-transition, that in its total connectivity, when p scales logarithmically with the size of the graph. This result is described in Section 1.2.2, and the methods used in its proof influence our approach in Chapter 4. In Section 1.3.3, we give a total connectivity result for configuration model, the proof of which inspires the proof of its generalization in Chapter 4. Finally, in Section 1.4, we introduce the notion of Local Weak Convergence which we use to motivate the introduction of a convex order in random graphs in Chapter 5.

1.1 Galton-Watson Branching Processes

The Galton-Watson branching process was first used by Sir Francis Galton to describe a family tree and compute the probability of extinction of family names. This branching process has later been studied in extensive detail and generality.

Informally, the Galton-Watson process assumes that at every generation of population, each individual independently gives birth to a random number of children.

More formally, a Galton-Watson process is a Markov process $\{Z_n\}$, with $Z_0 = 1$, and it evolves subsequently such that given Z_n,

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_i^{(n)}, \tag{1.1}$$

where $\{\xi_i^{(n)} : n, i \in \mathbb{N}\}$ is a set of i.i.d. random variables, independent of Z_0, each having probability distribution $\boldsymbol{p} = \{p_k\}_{k\geq 0}$. Let ξ also be a random variable having distribution $\boldsymbol{p}$, and its generating function be given by

$$\phi_\xi(s) := \mathbb{E}\left[s^\xi\right] = \sum_{k\geq 0} p_k s^k, \tag{1.2}$$

for $s \in [0,1]$.

Now, let the probability of population extinction be given by

$$p_{ext} := \mathbb{P}(\exists n \geq 1, Z_n = 0). \tag{1.3}$$

Then, assuming $p_1 < 1$ (since the case where $p_1 = 1$ is trivial), we have the following theorem.

Theorem 1.1 *(Survival vs. Extinction). The extinction probability, p_{ext}, is given by the smallest solution in $[0,1]$ of*

$$s = \phi_\xi(s) \tag{1.4}$$

In particular, the following regimes can happen:

1. *Subcritical regime: If $\mathbb{E}[\xi] < 1$, then $p_{ext} = 1$.*
2. *Critical regime: If $\mathbb{E}[\xi] = 1$ and ξ is not deterministic, then $p_{ext} = 1$.*
3. *Supercritical regime: If $\mathbb{E}[\xi] > 1$, then $p_{ext} < 1$.*

This is one of the foremost examples of the phenomenon of phase-transition, a huge qualitative change in global behaviour precipitated by a tiny change in a parameter determining local behaviour, and is intimately related to many other phase-transitions we will see in this thesis. In particular, the above theorem, via branching process approximation, often suggests the *critical* condition for the appearance of a giant component in random graphs, as we will see later.

We give below another important result which mirrors the dichotomy between the sizes of the largest and the second largest components of a random graph.

Theorem 1.2 *Assume that $p_1 < 1$. Then, $\lim_{n\to\infty} Z_n \in \{0, \infty\}$ almost surely.*

To prove the above, it is first shown that the distribution of a supercritical branching process conditioned on extinction is equivalent to that of an (unconditioned) dual subcritical branching process. A similar duality is also observed in random graphs.

For the proofs of above theorems, and further details on the theory of branching processes, we refer to Harris (1963) [53].

1.2 Erdős-Rényi Random Graph

In this section, we recall some well-known results on the classical random graph model introduced and analysed by Erdős and Rényi in [27] and [28], which are in the same spirit as our results in Chapter 2.

The neighborhood of any vertex in Erdős-Rényi graph looks like that of the root of Galton-Watson branching process (with Poisson offspring distribution), described in the previous section and a phase

transition corresponding to the survival-extinction phase transition of Galton-Watson process occurs here. This is the phase transition in the size of the largest connected component of the graph, which we describe next.

1.2.1 Phase transition: Emergence of a giant component

The Erdős-Rényi random graph $ER(n,p)$ has vertex set $[n] = 1,2,\ldots,n$, and, each pair of vertices is connected by an edge with probability p, independently of other pairs. Remark that the degree distribution of a uniformly chosen vertex in $ER(n,p)$ is given by Binomial distribution with parameters $n-1$ and p. If we take $ER(n,p)$ to be sparse, i.e. $p = \lambda/n$, this degree distribution converges to Poisson distribution with parameter λ, when $n \to \infty$. In this case, we have the following phase-transition theorem on the emergence of a giant connected component in the graph.

Theorem 1.3 *([16, 28, 37]). Let C_1 and C_2 be the largest and the second largest components of the Erdős-Rényi random graph, $ER(n,\lambda/n)$. Denoting by $|C|$ the size, i.e. the number of vertices, of a component C, we have,*

1. *Subcritical regime: $\lambda < 1$. For some constant c depending on λ, the following holds:*

$$\lim_{n\to\infty} \mathbb{P}(|C_1| \leq c\log n) = 1. \tag{1.5}$$

2. *Supercritical regime: $\lambda > 1$. Denote by $p_{ext}(\lambda)$ the extinction probability of a Galton- Watson branching process with $Poi(\lambda)$ offspring distribution, i.e., the unique root in $(0,1)$ of the equation $x = \exp(-\lambda(1-x))$. Then for some constant $c > 0$ depending on λ, and all $\delta > 0$, one has the following:*

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{|C_1|}{n} - (1-p_{ext}(\lambda))\right| \leq \delta, \text{ and } |C_2| \leq c\log n\right) = 1. \tag{1.6}$$

The relation to *Poisson* branching process is self-evident in the result: the *critical* boundary separating the different regimes as well as the size of the largest component in supercritical regime are determined by the offspring distribution of the approximation branching process, and the size of the second largest component in supercritical regime is analogous to the size of the largest component in subcritical regime.

There is a non-trivial scaling result for the *critical regime* ($\lambda = 1$), which we do not discuss here, but refer to Bollobás (1984) [17] as well as [45, 46, 1, 39] for more recent results. We also refer to an excellent discussion of the above phase-transition result in Alon and Spencer (2000) [3], and Draief and Massoulie (2010) [26].

1.2.2 Connectivity

In the context of information diffusion in a graph, an important problem is to ensure that the information reaches the entire population, or in the case of an epidemic spread, prevent the epidemic from reaching the entire population (through vaccination or other strategies). This requires one to analyse how connectivity appears in a random graph with a change in system parameters. In the previous section, we saw that when the average asymptotic degree λ of an Erdős-Rényi random graph is of constant order $\lambda > 1$, the size of the largest connected component of the graph, even though of order n, is still strictly less than n, so that the graph is disconnected. In this section, we will see that the connectivity appears in Erdős-Rényi graph when λ scales logarithmically with n.

More precisely, one evaluates the probability of connectivity of $ER(n,p)$ when np scales as $\log n + c$ for some constant c. The strategy is to first evaluate the probability that there are no isolated vertices using Poisson approximation by the Stein–Chen method, and then show that, with high probability, disconnections are possible due to isolated vertices only. This is the same strategy that we adapt in Chapter 4 to prove that with high probability, the longest edge of the minimum spanning tree of configuration model coincides with that of its nearest neighbour graph, and both have an isolated vertex as one end-point. Therefore we walk through the main intermediate steps of the analysis in this section, but for complete proofs, we refer to Chapter 3 of Draief and Massoulie (2010) [26], and Lindvall (1992) [44].

The intuition behind the first part is as follows. Denoting the degree of vertex i by d_i, $I_i := \mathbb{1}(d_i = 0)$ is a Bernoulli random variable with parameter $(1-p)^{n-1}$ and the number of isolated vertices is given by $N_n = \sum_{i=1}^{n} I_i$. Ignoring the weak dependence among I_i's,

$$N_n \overset{d}{\approx} Binomial(n, (1-p)^{n-1}).$$

Now, when $np = \log n + c$,

$$\mathbb{E}N_n = n\left(1 - \frac{\log n + c}{n}\right)^{n-1} \to e^{-c}, \tag{1.7}$$

as $n \to \infty$. Therefore, by Poisson approximation of a Binomial random variable, we have that for large n,

$$N_n \overset{d}{\approx} Poisson(e^{-c}).$$

The Poisson approximation for the sum of dependent Bernoulli random variables is made rigorous using the Stein-Chen method as given below. Here, P_a denotes Poission r.v. with parameter a, $\mathscr{L}(\boldsymbol{V})$ denotes the law of random vector $\boldsymbol{V}$, and $d_{var}(.,.)$ denotes the *distance in total variation*.

Theorem 1.4 (Stein-Chen) *For a finite or countable set V, let $(I_j)_{j\in V}$ be a family of not necessarily independent Bernoulli variables with respective parameters $(\pi_j)_{j\in V}$. Let $X = \sum_{j\in V} I_j$ and $\lambda = \sum_{j\in V} \pi_j$. Assume that there exist random variables $(J_{ij})_{i,j\in V, j\neq i}$ defined on the same probability space as $(I_j)_{j\in V}$ and such that for all $i \in V$, the following equality of distributions holds:*

$$\mathscr{L}((J_{ij})_{i,j\in V,j\neq i}) = \mathscr{L}((I_j)_{j\in V,j\neq i}|I_i = 1). \tag{1.8}$$

Then

$$d_{var}(\mathscr{L}(X), P_\lambda) \leq 2\frac{1-e^{-\lambda}}{\lambda}\sum_{i\in V}\pi_i(\pi_i + \sum_{j\in V, j\neq i}\mathbb{E}(|I_j - J_{ij}|). \tag{1.9}$$

Lemma 1.5 *For $\lambda, \mu \geq 0$,*

$$d_{var}(P_\lambda, P_\mu) \leq 2|\lambda - \mu|.$$

Thus, the Poisson approximation of isolated vertices using the Stein-Chen method leads to the following theorem.

Theorem 1.6 *Assume that for some fixed $c \in \mathbb{R}$, $np = \log n + c$. Then the distribution of the number X of isolated nodes in $ER(n,p)$ converges in variation, as $n \to \infty$, to the Poisson distribution with parameter e^{-c}.*

It remains to be shown that given the assumption on p, the probability that the graph contains connected components of sizes between 2 and $n/2$ goes to zero. This will imply that the probability that the graph is connected is asymptotically equivalent to the probability that it has no isolated nodes, that is,

Theorem 1.7 *Let $c \in \mathbb{R}$ be given, and assume that p is such that $np = \log n + c$. One then has the limit*

$$\lim_{n\to\infty} \mathbb{P}(ER(n,p) \text{ connected}) = e^{-e^{-c}}. \tag{1.10}$$

A direct consequence of the above theorem is that if the average asymptotic degree λ asymptotically dominates $\log n$, i.e. if c is replaced by c_n such that $c_n \to \infty$ as n increases, then the Erdős-Rényi graph is connected with high probability. This represents a phase-transition in the connectivity of Erdős-Rényi graph.

1.3 Configuration Model

Recall from the previous section that the asymptotic degree distribution of a uniformly chosen vertex in $ER(n, \lambda/n)$ is given by Poisson distribution with parameter λ. In configuration model, we let this degree distribution to be arbitrary. This covers a wide range of random graphs of practical or theoretical interest, for example, random regular graphs (where every vertex has a fixed degree) and power-law random graphs. In fact, some of the earliest applications of configuration model were in the study of these two examples (Bender and Canfield (1978) [10], and Bollobás (2001) [15]).

Given a vertex set $[n] = 1, 2, \ldots, n$ and a sequence of non-negative integers, $\mathbf{d}^{(n)} = (d_i^{(n)})_1^n$, with $\sum_{i=1}^n d_i^{(n)}$ even, we are interested in a simple graph, $G(n, (d_i^{(n)})_1^n)$, chosen uniformly at random from all graphs where each vertex i has degree $d_i^{(n)}$. Unlike $ER(n,p)$, in $G(n, (d_i^{(n)})_1^n)$, the pairs of vertices are not connected independently of each other, therefore it is not easy to construct this graph directly. But it is possible to first construct the multi-graph version of this model, which is what we call configuration model and denote by $G^*(n, (d_i^{(n)})_1^n)$, and then prove that the relevant results proven on $G^*(n, (d_i^{(n)})_1^n)$ hold true for $G(n, (d_i^{(n)})_1^n)$.

$G^*(n, (d_i^{(n)})_1^n)$ is constructed simply by giving $d_i^{(n)}$ half-edges to each vertex i, and then randomly matching all the half-edges. Evidently, this procedure can produce self-loops and multi-edges, but $G^*(n, (d_i^{(n)})_1^n)$, conditioned to be simple, will be the simple graph $G(n, (d_i^{(n)})_1^n)$. Denote the total number of half-edges by

$$l_n = \sum_{i=1}^n d_i^{(n)}. \tag{1.11}$$

Then, the total number of configurations, i.e. all the different ways of pairing l_n half-edges is given by

$$(l_n - 1)!! \tag{1.12}$$

where we recall that $m!!$ denotes the double factorial of m, which, if m is even, is the product of all even integers less than or equal to m but greater than or equal to 2, and if m is odd, is the product of all odd integers less than or equal to m and greater than or equal to 1.

In the rest of the thesis, we will assume the following regularity conditions for $\mathbf{d}^{(n)}$, which were introduced by Molloy and Reed in [48], but the formulation below is based on the one used by Janson and Luczak in [38].

Condition 1.8 *For each n, $\mathbf{d}^{(n)} = (d_i^{(n)})_1^n$ is a sequence of non-negative integers such that $\sum_{i=1}^n d_i^{(n)}$ is even and, for some probability distribution $(p_r)_{r=0}^\infty$ over integers, independent of n, the following hold.*

(i) The degree density condition: $\frac{|\{i:d_i^{(n)}=k\}|}{n} \to p_k$ for every k as $n \to \infty$.

(ii) Finite expectation property: $\lambda := \sum_{k\geq 0} k p_k \in (0, \infty)$.

(iii) Second moment property: $\sum_{i=1}^n (d_i^{(n)})^2 = O(n)$.

If D_n denotes the degree of a uniformly chosen vertex in $G^*(n, (d_i^{(n)})_1^n)$, and D is a random variable with distribution $(p_k)_{k\geq 0}$, then Condition 1.8 (i) is equivalent to saying that

$$D_n \xrightarrow{\text{p}} D, \tag{1.13}$$

i.e., D is the asymptotic degree distribution of a uniformly chosen vertex in $G(n, (d_i^{(n)})_1^n)$.

Evidently, Condition 1.8 (ii) is equivalent to saying that

$$\lambda = \mathbb{E}D \in (0, \infty), \tag{1.14}$$

and Condition 1.8 (iii) is equivalent to

$$\mathbb{E}D_n^2 = O(1), \tag{1.15}$$

which, in particular, implies that the random variables D_n are uniformly integrable, and thus,

$$\mathbb{E}D_n \to \mathbb{E}D \tag{1.16}$$

as $n \to \infty$.

Now we would like to show that if we want to prove a result on $G(n,(d_i^{(n)})_1^n)$, then generally, it suffices to prove it on $G^*(n,(d_i^{(n)})_1^n)$. More precisely, we would like to prove the following result.

Theorem 1.9 *Let $\mathbf{d}^{(n)} = (d_i^{(n)})_1^n$ be a given fixed degree sequence satisfying Condition 1.8. Then, an event $\mathcal{E}_n$ occurs whp for $G(n,(d_i^{(n)})_1^n)$ when it occurs with high probability for $G^*(n,(d_i^{(n)})_1^n)$.*

This follows as a corollary of the following theorem.

Theorem 1.10 *([36]). Consider a random graph $G^*(n,(d_i^{(n)})_1^n)$ where the degree sequence $(d_i^{(n)})_1^n$ satisfies Condition 1.8. Then*

$$\liminf_{n\to\infty} \mathbb{P}(G^*(n,(d_i^{(n)})_1^n)) \text{ is simple }) > 0. \tag{1.17}$$

We refer to Janson (2014) [36] for the proofs.

To make the notation a little less cumbersome for the rest of the thesis, we will drop the superscript in $d_i^{(n)}$ and simply use d_i instead.

1.3.1 Phase transition: Emergence of a giant component

The phase-transition result regarding the emergence of a giant component in $ER(n,\lambda/n)$ (Section 1.2.1) has been generalized to configuration model by Molloy and Reed in [48] and [49], where the *critical* boundary of the phase-transition is given by

$$\mathbb{E}[D(D-2)] > 0. \tag{1.18}$$

Superficially, this condition appears quite different from the case of $ER(n,\lambda/n)$, where the *critical* boundary of the phase-transition is determined by the average aympototic degree. But the similarity becomes evident when one examines the approximating branching process. Let us start the exploration of a connected component of configuration model from a vertex i. Then, the number of neighbours, analogous to the first generation in the approximating branching process, has distribution $(p_k)_{k\geq 0}$. But this is not true from the second generation onwards. A first generation vertex with degree k is k times as likely to be chosen as one with degree 1, so the distribution of the number of children of a first generation vertex

becomes size-biased and is given by

$$q_{k-1} = \frac{kp_k}{\lambda}, \tag{1.19}$$

for all $k \geq 1$. The $k-1$ on the left-hand side comes from the fact that we used up one edge connecting to the vertex. The same logic applies, upto an approximation, to the subsequent generations. Now, the expectation of this size-biased distribution is given by

$$\nu := \sum_{k=0}^{\infty} kq_k = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}. \tag{1.20}$$

Therefore, the *critical* boundary of the approximating branching process, $\nu > 1$, becomes (1.18). Remark that in the case of $ER(n, \lambda/n)$, whose asymptotic degree distribution is the Poisson distribution, $p_k = e^{-\lambda}\lambda^k/k!$, the size-biased distribution is given by

$$q_{k-1} = e^{-\lambda}\frac{k\lambda^k}{\lambda k!} = e^{-\lambda}\frac{\lambda^{k-1}}{(k-1)!}, \tag{1.21}$$

which is same as the asymptotic degree distribution! This is the reason that we get a deceptively simple *critical* boundary condition in the case of $ER(n, \lambda/n)$.

Now we can state the phase-transition theorem for configuration model, analogous to Theorem 1.3 for $ER(n, \lambda/n)$.

Theorem 1.11 *([49], [38]). Let C_1 and C_2 be the largest and the second largest components of $G(n, (d_i)_1^n)$. Assume that $p_0 + p_2 < 1$. Then, we have,*

1. *Subcritical regime: $\mathbb{E}[D(D-2)] = \sum k(k-2)p_k < 0$. For all $\delta > 0$, the following holds:*

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{|C_1|}{n} \leq \delta\right) = 1. \tag{1.22}$$

2. *Supercritical regime: $\mathbb{E}[D(D-2)] = \sum k(k-2)p_k > 0$. For some constants $c_1, c_2 > 0$ depending on D, and for all $\delta > 0$, one has the following:*

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{|C_1|}{n} - c_1\right| \leq \delta, \text{ and } |C_2| \leq c_2 \log n\right) = 1. \tag{1.23}$$

In the above theorem, c_1 is related to the extinction probability of the approximating branching process in the same way as in the case of $ER(n, \lambda/n)$, i.e. if $p_{ext}(q)$ denotes the extinction probability of branching process having distribution $(q_k)_{k\geq 0}$ (as given by (1.19)), then

$$c_1 = 1 - p_{ext}(q). \tag{1.24}$$

The original proof of the above theorem was given by Molloy and Reed in [49], along with a stronger result for subcritical regime, but another arguably more elegant approach was proposed by Janson and Luczak in [38]. It is the latter approach that we adapt for the phase-transition results in Chapter 2.

Both the case $p_0+p_2 = 1$ and the *critical regime* have also been investigated in Janson and Luczak (2008) [38]. For results regarding the behavior in the *critical regime*, also see [40, 32].

1.3.2 Information diffusion

Various means of information diffusion in random graphs have been investigated in depth, particularly those inspired by percolation theory. We describe them here for an underlying graph, G_n.

In the case of *bond percolation*, for some $\pi \in (0, 1)$, each edge of G_n is present with probability π independently of every other edge. In other words, we randomly delete the edges (i.e. the bonds) of G_n. In the case of *site percolation*, every vertex (site) of G_n is isolated with probability $1 - \pi$, independently of every other vertex (or in other words, we delete this vertex). The random subgraph thus obtained is the spanning subgraph of G_n that does not contain the deleted edges (bond percolation) or the edges attached to the deleted vertices (site percolation). A question of particular interest is if this random subgraph exhibits a phase-transition. The diffusion dynamic that we investigate in Chapters 2 and 3 can be considered as a generalization of these percolation models.

Bond percolation is sometimes referred to as *independent cascade model* in the context of information diffusion. Another diffusion dynamic, which is particularly interesting from the perspective of opinion formation in social networks is that of *symmetric threshold model*, where the probability of an individual getting influenced depends on the number of its neighbors who are already influenced. We will not consider this model in this thesis, but refer to [22] and [23] for an excellent analysis.

In the case of Erdős-Rényi random graphs, percolation is equivalent to retaining a fraction π of the edges or sites, which results in another Erdős-Rényi random graph. Remarkably, we have a similar result

for configuration model. In the rest of this section, we discuss the case of bond percolation alone since we use its results directly for our connectivity proof in Chapter 4, but the case of site percolation is similar.

Consider the configuration model, $G^*(n,(d_i)_1^n)$, and a sequence of percolation probabilities, $(\pi_n)_{n\in\mathbb{N}}$. Let $\mathbf{D}^\pi = (D_i^\pi)_1^n$ be the random degree sequence and $G^*(n,(D_i^\pi)_1^n)$ be the graph induced by the random deletion of the edges of $G^*(n,(d_i)_1^n)$ with probability π_n.

Then, we have the following lemma from Fountoulakis (2007) [29].

Lemma 1.12 *Conditional on $\mathbf{D}^\pi = (d_i^\pi)_1^n$ and $\sum_{i=1}^n d_i^\pi$ being even, the subgraph induced by the random deletion of the edges of $G^*(n,(d_i)_1^n)$ with probability π_n, is also a configuration model, $G^*(n,(d_i^\pi)_1^n)$, on the same set of vertices and with degree sequence $(d_i^\pi)_1^n$.*

This lemma allows one to use the results already proved for configuration model, now for the subgraph obtained after percolation. In particular, we have a phase-transition on the percolated subgraph, analogous to that in Theorem 1.11

We assume that Condition 1.8 holds for $(d_i)_1^n$ and also that $\mathbb{E}[D(D-2)] = \sum k(k-2)p_k > 0$ and $p_0 + p_2 < 1$, so that we are in the supercritical regime of Theorem 1.11 and there exists a unique giant component in $G^*(n,(d_i)_1^n)$. Also, let

$$\pi^c := \frac{\mathbb{E}D}{\mathbb{E}D(D-1)}. \tag{1.25}$$

Then, corresponding to Theorem 1.11, we have

Theorem 1.13 *([18], [30], [35]) Let C_1^π denote the largest component and C_2^π the second largest component of $G^*(n,(D_i^\pi)_1^n)$. Then we have,*

1. *Subcritical regime: $\limsup_{n\to\infty} \pi_n < p^c$. For all $\delta > 0$, the following holds:*

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{|C_1^\pi|}{n} \le \delta\right) = 1. \tag{1.26}$$

2. *Supercritical regime: $\liminf_{n\to\infty} \pi_n > p^c$. For some constants $c_1, c_2 > 0$ depending on D and $(\pi_n)_{n\in\mathbb{N}}$, and for all $\delta > 0$, one has the following:*

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{|C_1^\pi|}{n} - c_1\right| \le \delta, \text{ and } |C_2^\pi| \le c_2 \log n\right) = 1. \tag{1.27}$$

The proofs in [18], [30] and [35] are given for constant percolation probability π, but extension to the case of $(\pi_n)_{n\in\mathbb{N}}$ is relatively straightforward.

1.3.3 Connectivity

From Section 1.3.1, we have that if Condition 1.8 (i) holds with $p_1 = p_2 = 0$, then the average number of children of the approximating size-biased branching process, ν, trivially satisfies $\nu \geq 2 > 1$, so that we are always in the supercritical regime of $G^*(n,(d_i)_1^n)$. Moreover, $p_1 = p_2 = 0$ implies that the extinction probability of the approximating branching process, $p_{ext}(q)$, is 1 and since $c_1 = 1 - p_{ext}(q) = 0$, then by Theorem 1.11, the largest component, C_1, satisfies

$$\frac{|C_1|}{n} \xrightarrow{\text{P}} 1. \tag{1.28}$$

In [54], van der Hofstad extends this to the statement that $G^*(n,(d_i)_1^n)$ is with high probability connected, i.e., $C_1 = [n]$ and $|C_1| = n$, and remarkably without even assuming Condition 1.8 for $\mathbf{d}^{(n)}$.

However, we do not have a result for a phase-transition in the connectivity of configuration model as in Section 1.2.2 for $ER(n,p)$. But from a practical point of view, an important advantage in the case of configuration model is that it is possible for the graph to be connected while the average degree is bounded, whereas for $ER(n,p)$ to be whp connected, the average degree must tend to infinity, as seen in Section 1.2.2. Many real-world networks are connected, therefore this property makes the configuration model often more suitable for modelling than other random graphs.

Let N_1 and N_2 denote the number of degree-1 and degree-2 vertices, respectively, in $G^*(n,(d_i)_1^n)$. We first state the disconnectivity results for $G^*(n,(d_i)_1^n)$ when either $N_1 >> n^{1/2}$, or when $p_2 = \mathbb{P}(D=2) > 0$. The main result in this section is Theorem 1.16, which states that for all possible degree sequences with $N_1 = N_2 = 0$, $G^*(n,(d_i)_1^n)$ is whp connected. As remarked earlier, we do not need Condition 1.8 in this case.

Proposition 1.14 *(Disconnectivity of $G^*(n,(d_i)_1^n)$ when $N_1 >> n^{1/2}$). Let Condition 1.8 (i)-(ii) hold, and assume that $N_1 >> n^{1/2}$. Then,*

$$\lim_{n\to\infty} \mathbb{P}(G^*(n,(d_i)_1^n) \textit{ connected}) = 0. \tag{1.29}$$

Proposition 1.15 *(Disconnectivity of $G^*(n,(d_i)_1^n)$ when $p_2 > 0$). Let Condition 1.8 (i)-(ii) hold, and assume that $p_2 > 0$. Then,*

$$\limsup_{n\to\infty} \mathbb{P}(G^*(n,(d_i)_1^n) \text{ connected}) < 1. \tag{1.30}$$

Theorem 1.16 *(Connectivity of $G^*(n,(d_i)_1^n)$). Assume that $d_i \geq 3$ for every $i \in [n]$. Then $G^*(n,(d_i)_1^n)$ is connected whp, i.e.,*

$$\mathbb{P}(G^*(n,(d_i)_1^n) \text{ disconnected}) = O(1/n). \tag{1.31}$$

In Chapter 4, we generalize the above result and prove whp connectivity when $N_1 < \rho_1 n^{1/3}$ and $N_2 < \rho_2 n^{2/3}$ for some constants ρ_1 and ρ_2. However, we will assume Condition 1.8 in that case since we will need the bound on the second largest component of configuration model in supercritical regime as given by Theorem 1.11.

1.3.4 Diameter of weighted configuration model

We will now give the result for the diameter of weights of configuration model, i.e. configuration model-whose edges are given i.i.d. exponentially distributed weights with parameter 1. This result given here is proved by Amini and Lelarge in [5].

Recall that for a graph G, the diameter of the graph is given by

$$\text{diam}(G) := \max\{\text{dist}(a,b), a, b \in V, \text{dist}(a,b) < \infty\} \tag{1.32}$$

where $dist$ is the usual graph distance and V is the vertex set of graph G.

Also recall from Section 1.3.1 that if $\nu > 1$, then a giant component exists whp in configuration model. We assume this for the following theorem. Let $d_{min} := \min\{k | p_k > 0\}$ such that for $k < d_{min}$; $|\{i, d_i = k\}| = 0$, for all n sufficiently large. We state the result only for the case $d_{min} \geq 3$.

Theorem 1.17 *([5]) Consider a random graph $G(n,(d_i)_1^n)$ with i.i.d. exponential 1 weights on its edges, where the degree sequences $(d_i)_1^n$ satisfy the above conditions on d_{min} and ν along with Condition 1.8. Then we have*

$$\frac{\text{diam}(G(n,(d_i)_1^n))}{\log n} \xrightarrow{p} \frac{1}{\nu - 1} + \frac{2}{d_{min}} \tag{1.33}$$

The first term in the above theorem, $1/(\nu-1)$, is closely linked to the typical distance, i.e. distance between two uniformly vertices, of $G(n,(d_i)_1^n)$ (see [11]). We will see in Chapter 4 that the second term is closely linked to the length of the longest edge of nearest neighbour graph (or minimum spanning tree) of $G(n,(d_i)_1^n)$.

1.4 Local weak convergence

We briefly introduce here the idea of local weak convergence in the context of random rooted graphs, which provides the framework for our discussion on convex comparison of random graphs in Chapter 5. We refer to Aldous and Steele (2002) [2] for more details and applications of this idea.

We first recall that a rooted graph is a graph with a distinguished vertex called root, and a rooted graph isomorphism from G to G' is a graph isomorphism which maps the root of G to the root of G'. We now recall what is meant by local weak convergence for unweighted graphs.

Definition 1.18 (Local Weak Convergence) *Let $\mathscr{G}_*$ be the set of rooted graphs. For $G \in \mathscr{G}_*$, we denote by $N_k(G)$ the restriction of G to those vertices which are separated from the root by at most k edges.*

We say that G_n **converges** *to G_∞ in $\mathscr{G}_*$ if for each $k \in \mathbb{N}$, there exists an $n_0 = n_0(k, G_\infty)$ such that for all $n \geq n_0$ there exists a rooted graph isomorphism from $N_k(G_\infty)$ to $N_k(G_n)$.*

This definition determines a topology that makes $\mathscr{G}_*$ into a complete separable metric space. Local weak convergence is the usual weak convergence with respect to this space.

In the context of random graphs that we have introduced in this chapter, it is natural to ask if $[ER(n,\lambda/n); n \in \mathbb{N}]$ and $[G^*(n,(d_i)_1^n); n \in \mathbb{N}]$ converge to their approximating branching processes in the local weak sense. The answer is indeed yes. We do not formally state the results here since they are only relevant for the informal discussion in Chapter 5, but refer to Dembo and Montanari (2009) [24] for the details in the case of configuration model (of which Erdős-Rényi model is a special case).

2

Viral Marketing in Configuration Model

This chapter is organized as follows. In Section 2.1, we motivate the introduction of a new mechanism of information propagation and discuss our results, methodology and related work. In the next section, we formally describe our model and formulate the results. In Sections 2.3 and 2.4 we analyze, respectively, the original and reversed dynamic of influence propagation. The relations between the two dynamics are explored in Section 2.5.

2.1 Introduction

A motivation for our work can come from the phenomenon of viral marketing in social networks: A person after getting acquainted with an advertisement (or a news article or a Gangnam style video, for that matter) through one of his "friends", may decide to share it with some (not necessarily all) of his friends, who will, in turn, pass it along to some of their friends, and so on. The campaign is successful if starting from a relatively small number of initially targeted persons, the influence (or information) can spread as

an epidemic "infecting" a non-negligible fraction of the population. This mode of diffusion can be seen as a generalization of *independent cascade model* (or, bond percolation), introduced in Section 1.3.2), which is a good model for phenomena similar to the spread of an epidemic in a population, but not so much for the diffusion of information (marketing campaign, news article etc) in an online social network.

Another motivation from a more theoretical perspective is that this generalization demonstrates an asymmetry, not exhibited by bond percolation, between the set of agents who have a huge influence in the graph and the set of agents who are most susceptible to being influenced. This asymmetry arises because the processes representing *forward* propagation of influence and *backward* tracking of influence are not statistically identical.

We analyse this new mechanism of information propagation on the underlying graph given by configuration model (introduced in Section 1.3), because due to its structure, it becomes possible to simultaneously construct the underlying graph and explore the process of information propagation.

2.1.1 Enhanced Configuration Model

In order to generalize the diffusion process in configuration model, we enhance the (classical) configuration model introduced in Section 1.3 by considering *two types of half-edges*. *Transmitter half edges* of a given vertex represent links through which this vertex will influence (pass the information once it has it) to its neighbours. Its *receiver half-edges* represent links through which this vertex will not propagate the information to its neighbours. The neighbours receive the information both through their transmitter and receiver half-edges matched to a transmitter half edge of the information sender. The two types of half-edges are not distinguished during the uniform pair-wise matching of all half-edges, but only to trace the propagation of information. Assuming the usual consistency conditions for the numbers of transmitter and receiver half-edges, the *enhanced configuration model* is asymptotically (when the number of vertices n goes to infinity) described by the vector of two, not necessarily independent, integer valued random variables, representing the *transmitter* and *receiver degree* of the typical vertex. Equivalently, we can consider the *total vertex degree*, representing the total number of friends of a person and its transmitter degree, representing the number of friends he/she is able to influence.

Remark that this model is a generalization of the following more or less classical cases.

(i) (Pure) Configuration Model: A vertex can influence all its neighbours, that is, all the half-edges are transmitter half-edges (see [48, 49, 38]).

(ii) Bond Percolation: A vertex can influence each of its neighbours independently with probability p. Note that there is a difference between such a bond-percolation model (which has oriented edges) and the "usual" bond percolation studied on the configuration model (cf [29, 35]) in which edges of the underlying graph are appropriately thinned to get (non-directed) *bond percolated* components. It is easy however to see that one can couple both models so that the information propagation starting from *one* arbitrarily fixed vertex reaches exactly all the vertices of its bond percolated component.

(iii) Node Percolation: A vertex can influence all of its neighbours with probability p and none with probability $1-p$. Again, there is a slight difference from the previous studies in [29, 35]. This time one can couple both models so that the information propagation started from *any* vertex reaches all the vertices of its *node percolated* component and in addition, some extra leaves which would have been *closed* in the node percolated graph.

(iv) Coupon-Collector propagation: A vertex chooses, some number of times, one of its neighbors (with replacement, i.e., forgetting his previous choices) to whom it propagates the message.

2.1.2 Results

We consider the advertisement campaign started from some initial target (source vertex) and following the aforementioned dynamic on a realization of the enhanced configuration model of the total number of vertices n. The results are formulated with high probability (whp), i.e. with probability approaching one as $n \to \infty$.

First, in Theorem 2.2 we give a condition (2.4) involving the total degree and the transmitter degree distributions of the enhanced configuration model, which if satisfied, would allow whp the advertisement campaign to reach a non-negligible ($O(n)$) fraction of the vertices, called a *big (influenced) component*, provided that the initial target is chosen from a set of *good pioneers*. We explicitly evaluate the asymptotic size of this component relative to n. Further in this case, in Theorem 2.3 we show that asymptotically the *big component is essentially the same* regardless of the good pioneer chosen. The essential uniqueness of the big component means that the subsets of influenced vertices reached from two different good pioneers differ by at most $o(n)$ vertices whp. The condition (2.4) is also tight in the sense that if not satisfied then the set of good pioneers is asymptotically negligible (Theorem 2.4). Finally, under condition (2.4) we calculate the relative size of the set of good pioneers, provided it is non-null. This latter (we believe

technical) assumption allows us to identify the set of good pioneers with a "big source component" in a dual process with the "reversed" dynamic and thus calculate its size (Theorem 2.5 and Corollary 2.8). Our duality relation says also that any given node from the big influenced component can be influenced *only* (up to a negligible fraction of nodes) by the set of good pioneers; we call this *strong inter-connectivity* of the two sets.

Remark here that in the case of (classical) configuration model, the big influenced component and the set of good pioneers are one and the same, that is, the giant component of the graph. Moreover, the aforementioned strong inter-connectivity is different from the strong connectivity usually considered in directed graphs.

2.1.3 Methodology

A standard technique for the analysis of diffusion of information on the configuration model consists in simultaneous exploration of the model and the propagation of the influence. We adopt this technique and, more precisely, the approach proposed in [38] for the study of the giant component of the (classical) configuration model. The approach originally used in [48, 49] to study the giant component of the (classical) configuration model involved approximating the initial phase of graph exploration process by a branching process, but this approach proves unwieldy for our model. In the approach proposed in [38], instead of the branching process approximation, one uses a "fluid limit" analysis of the process up to the time when the exploration of the big component is completed. We tailor this method to our specific dynamic of influence propagation and calculate the relative size of the big influenced component, as well as prove its essential uniqueness.

A fundamental difference with respect to the study of the giant component of the classical model stems from the directional character of our propagation dynamic. Precisely, the edges matching a transmitter and a receiver half-edge can relay the influence from the transmitter half-edge to the receiver one, but not the other way around. This means that the good pioneers do not need to belong to the big (influenced) component, and vice versa. In this context, we introduce a *reverse dynamic*, in which a message (think of an "acknowledgement") can be sent in the reversed direction on every edge (from an arbitrary half-edge to the receiver one), which traces all the possible sources of influence of a given vertex. This reversed dynamic can be studied using the same approach as the original one. In particular, one can establish the essential uniqueness of the big component of the reversed process as well as calculate its relative

size. Interestingly, this relative size coincides with the probability of the non-extinction of the branching process approximating the initial phase of the original exploration process, whence the hypothesis that the big component of the reverse process coincides with the set of good pioneers. We prove this conjecture under some additional (technical) assumption.

2.1.4 Related Work

The propagation of influence through a network has been previously studied in various contexts. The giant component of the (classical) configuration model, for example, has been extensively studied (see [48, 49, 38]). Strongly connected components in the *directed configuration model* (different from our model, see [54] for an overview of this model) was studied in [21], and previously in the directed Erdős-Rényi graph in [41].

The configuration model has also formed the base network for other dynamics of influence propagation, the most basic of which is the (bond and node) percolation (see [29, 35]). Another one relevant to the phenomenon of viral networking in social networks is discussed in [4, 43], where a vertex in the network gets influenced only if a certain proportion of its neighbours have already been influenced. This interesting propagation dynamic is further studied by introducing *cliques* in configuration model to observe the impact of clustering on the size of the population influenced (see [22, 23]). This dynamic is a kind of *pull model* where influence propagation depends on whether a vertex decides to receive the influence from its neighbours, while percolation is an example of *push model*, where the influence propagation depends on whether a vertex decides to transmit the influence. As mentioned in the introduction, we study a more generalized form of push model. A propagation dynamic where every influenced node, at all times, keeps choosing one of its neighbours uniformly at random and transmits the message to it is studied on a d-regular graph in [31]. This dynamic is close in its spirit to the one we considered in this chapter, however [31] focusses on the temporal evolution of the process unlike in this chapter. The process considered there stops when *all* nodes receive the message, and this stopping time is studied in the chapter. The same dynamic but restricted to some (possibly random) maximal number of transmissions allowed for each vertex is considered in [20] on a complete graph. This can be thought as a special case of our dynamic (although we study it on a different underlying graph) if we assume that the transmitter and receiver degrees correspond to the number of collected and non-collected coupons, respectively, in the classical coupon collector problem with the number of coupons being the vertex degree and the number of trials

being the number of allowed transmissions.

We report some numerical results regarding bond and node Percolation as well as the coupon-collector dynamics in Chapter 3. In a more applied context, an heuristic analysis of dynamic similar to ours is done on a real-world network (*facebook*) in [19].

Our method of introducing a reverse process to derive results for the original one is reminiscent of the usage of a backward "susceptibility" branching process in [9] to study epidemic curves. Directional character of our propagation dynamic and a systematic usage of the "fluid limit" analysis instead of the branching process approximation differentiate the two approaches.

2.2 Notation and Results

Given a degree sequence $(d_i^{(n)})_1^n$ for n vertices labelled 1 to n, we represent the degree, d_i, of each vertex i as the sum of two degrees: transmitter degree, $d_i^{(t)}$ and receiver degree, $d_i^{(r)}$.

We will asssume the following set of consistency conditions for our enhanced configuration model, which are analogous to those assumed for configuration model in Section 1.3.

Condition 2.1 *For each n, $\mathbf{d}^{(n)} = (d_i)_1^n$,is a sequence of non-negative integers such that $\sum_{i=1}^n d_i := 2m$ is even and for each i, $d_i = d_i^{(r)} + d_i^{(t)}$. For $k \in \mathbb{N}$, let $u_{k,l} = |\{i : d_i^{(r)} = k, d_i^{(t)} = l\}|$, and $D_n^{(r)}$ and $D_n^{(t)}$ be the receiver and transmitter degrees respectively of a uniformly chosen vertex in our model, i.e., $\mathbb{P}(D_n^{(r)} = k, D_n^{(t)} = l) = u_{k,l}/n$. Let $D^{(r)}$ and $D^{(t)}$ be two random variables taking value in non-negative integers with joint probability distribution $(p_{v,w})_{(v,w)\in\mathbb{N}^2}$, and $D := D^{(r)} + D^{(t)}$. Then the following hold.*

(i) $\frac{u_{k,l}}{n} \to p_{k,l}$ *for all* $(k,l) \in \mathbb{N}^2$.

(ii) $\mathbb{E}[D] = \mathbb{E}[D^{(r)} + D^{(t)}] = \sum_{k,l}(k+l)p_{k,l} \in (0,\infty)$. *Let* $\lambda_r = \mathbb{E}[D^{(r)}]$, $\lambda_t = \mathbb{E}[D^{(t)}]$ *and* $\lambda = \lambda_r + \lambda_t$.

(iii) $\sum_{i=1}^n (d_i)^2 = O(n)$.

(iv) $\mathbb{P}(D = 1) > 0$.

Let $g(x,y) := \mathbb{E}[x^{D^{(r)}} y^{D^{(t)}}]$ be the joint probability generating function of $(p_{v,w})_{(v,w)\in\mathbb{N}^2}$. Further let

$$h(x) := x \left.\frac{\partial g(x,y)}{\partial y}\right|_{y=x} = \mathbb{E}[D^{(t)} x^D], \tag{2.1}$$

and

$$H(x) := \lambda x^2 - \lambda_r x - h(x). \tag{2.2}$$

If two neighbouring vertices x and y are connected via the pairing of a transmitter half-edge of x with any half-edge of y, then x has the ability to directly influence y. More generally, for any two vertices x and y in the graph and $k \geq 1$, if there exists a set of vertices $x_0 = x, x_1, \ldots, x_{k-1}, x_k = y$ such that $\forall i : 1 \leq i \leq k$, x_{i-1} has the ability to directly influence x_i, we say that x has the ability to influence y and denote it by $x \to y$; in other words, y can be influenced starting from the initial source x. Let $C(x)$ be the set of vertices of $G(n, (d_i)_1^n)$ which are influenced starting from an initial source of influence, x, until the process stops, i.e.,

$$C(x) = \left\{ y \in v(G(n, (d_i)_1^n)) : x \to y \right\}, \tag{2.3}$$

where $v(G(n, (d_i)_1^n))$ denotes the set of all the vertices of $G(n, (d_i)_1^n)$. We use $|.|$ to denote the number of elements in a set here, although at other times we also use the symbol to denote the absolute value, which would be clear from the context.

2.2.1 Forward influence propagation

We have the following theorems for the forward influence propagation process.

Theorem 2.2 *Suppose that Condition 2.1 holds and consider the random graph $G(n, (d_i)_1^n)$, letting $n \to \infty$. If*

$$\mathbb{E}[D^{(t)}D] > \mathbb{E}[D^{(t)} + D] \tag{2.4}$$

then there is a unique $\xi \in (0, 1)$ such that $H(\xi) = 0$ and there exists at least one x_n in $G(n, (d_i)_1^n)$ such that

$$\frac{|C(x_n)|}{n} \xrightarrow{p} 1 - g(\xi, \xi) > 0. \tag{2.5}$$

We denote $C(x_n)$ constructed in the proof of Theorem 2.2 by C^*. For every $\epsilon > 0$, let

$$\mathbb{C}^s(\epsilon) := \left\{ x \in v(G(n, (d_i)_1^n)) : |C(x)| / n < \epsilon \right\}$$

and

$$\mathbb{C}^L(\epsilon) := \left\{ x \in v(G(n, (d_i)_1^n)) : |C(x) \triangle C^*| / n < \epsilon \right\},$$

where $\triangle$ denotes the symmetric difference. Remark that C^* and thus $\mathbb{C}^L(\epsilon)$ is defined only under condition (2.4).

Theorem 2.3 *Under assumptions of Theorem 2.2, we have that*

$$\forall \epsilon, \quad \frac{|\mathbb{C}^s(\epsilon)| + \left|\mathbb{C}^L(\epsilon)\right|}{n} \xrightarrow{p} 1. \tag{2.6}$$

Theorem 2.4 *Suppose that Condition 2.1 holds and* $\mathbb{E}[D^{(t)}D] \leq \mathbb{E}[D^{(t)} + D]$. *Then*

$$\forall \epsilon, \quad \frac{|\mathbb{C}^s(\epsilon)|}{n} \xrightarrow{p} 1. \tag{2.7}$$

Remark 1 The above results say that asymptotically ($n \to \infty$), under assumptions of Theorem 2.2, whp, there is essentially *one and only one big graph component* (of size $O(n)$) that can possibly be influenced starting propagation from a given vertex in the graph. Moreover, the condition (2.4) in Theorem 2.2 is *necessary* in the following sense: when it is not satisfied, then the chance of influencing a big component by randomly choosing the initial node goes to zero. In other words, whp, at most $o(n)$ nodes can influence a big component.

What the above results do not tell, however, is the relative size of the set of vertices which, under the condition (2.4), are indeed able to reach this big component (we call them *good pioneers*). This is the question we turn to next.

2.2.2 Pioneers — Branching process heuristic

Our analysis technique to obtain the above results involves the simultaneous exploration of the configuration model and the propagation of influence. Another commonly used method to explore the components of configuration model is to make the branching process approximation in the initial stages of the exploration process. Although we won't explicitly follow this path in this chapter, an heuristic analysis of the branching process approximation of our propagation model provides some important insights about the size of the set of good pioneers (using Theorem 1.1).

Coming to the approximation, if we start the exploration with a uniformly chosen vertex i, then the number of its neighbours that it does not influence and those that it does, denoted by the random vector

$(D_i^{(r)}, D_i^{(t)})$, will have a joint distribution $(p_{v,w})$. But since the probability of getting influenced is proportional to the degree, the number of neighbours of a first-generation vertex excluding its parent (the vertex which influenced it) won't follow this joint distribution. Their joint distribution as well the joint distribution in the subsequent generations, denoted by $(\widetilde{D}^{(r)}, \widetilde{D}^{(t)})$, is given by

$$\widetilde{p}_{v,w} = \frac{(v+1)\,p_{v+1,w} + (w+1)\,p_{v,w+1}}{\lambda}. \tag{2.8}$$

Note that Condition 2.1(iv) implies that $\mathbb{P}(\widetilde{D}^{(t)} = 0) > 0$, and therefore, from Theorem 1.1, this branching process gets extinct a.s. unless,

$$\begin{aligned}
&\mathbb{E}\left[\widetilde{D}^{(t)}\right] > 1;\\
\text{equivalently,}\quad &\sum_{v,w} w\widetilde{p}_{v,w} > 1,\\
&\sum_{v,w} \frac{w(v+1)\,p_{v+1,w} + w(w+1)\,p_{v,w+1}}{\lambda} > 1,\\
&\mathbb{E}\left[D^{(r)}D^{(t)}\right] + \mathbb{E}\left[D^{(t)}\left(D^{(t)}-1\right)\right] > \mathbb{E}[D],\\
&\mathbb{E}\left[DD^{(t)}\right] > \mathbb{E}\left[D + D^{(t)}\right].
\end{aligned}$$

This condition for non-extinction of branching process remarkably agrees with the condition in Theorem 2.2 which determines the possibility of influencing a non-negligible proportion of population.

Further from Theorem 1.1, if this condition is satisfied, the extinction probability of the branching process which diverges from the first-generation vertex, $\widetilde{p}_{ext}$, is given by the smallest $x \in (0,1)$ which satisfies

$$\begin{aligned}
&\mathbb{E}\left[x^{\widetilde{D}^{(t)}}\right] = x;\\
\text{equivalently,}\quad &\sum_{v,w} \frac{x^w(v+1)\,p_{v+1,w} + (w+1)\,x^w p_{v,w+1}}{\lambda} = x,\\
&\mathbb{E}\left[D^{(r)}x^{D^{(t)}}\right] + \mathbb{E}\left[D^{(t)}x^{D^{(t)}-1}\right] = x\mathbb{E}[D],\\
&\mathbb{E}[D]\,x^2 - \mathbb{E}\left[D^{(t)}x^{D^{(t)}}\right] - x\mathbb{E}\left[D^{(r)}x^{D^{(t)}}\right] = 0.
\end{aligned} \tag{2.9}$$

Note that 0 is excluded as a solution since $\mathbb{P}(\widetilde{D}^{(t)} = 0) > 0$.

Finally, the extinction probability of the branching process starting from the root, p_{ext}, is given by

$$p_{ext} = \mathbb{E}\Big[(\widetilde{p}_{ext})^{D^{(t)}}\Big]. \tag{2.10}$$

Since the root is uniformly chosen, we would expect the proportion of the vertices which can influence a non-negligible proportion to be roughly $1 - p_{ext} = 1 - \mathbb{E}\Big[(\widetilde{p}_{ext})^{D^{(t)}}\Big]$. Indeed, we confirm this result using a more rigorous analysis involving the introduction and study of a reverse influence propagation which essentially traces all the possible sources of influence of a given vertex. This method of introducing a reverse process (in a way, dual to the original process) to derive results for the original process has not been seen in a related context in the existing literature to the best of our knowledge, although the analysis of this dual process uses the familiar tools used for the original process.

2.2.3 Dual Back-Propagation Process

Let $\overline{g}(x) := \mathbb{E}[x^{D^{(t)}}]$, $\overline{h}(x) := \mathbb{E}[D^{(t)}x^{D^{(t)}}] + x\mathbb{E}[D^{(r)}x^{D^{(t)}}]$ and

$$\overline{H}(x) := \mathbb{E}[D]x^2 - \overline{h}(x) = \lambda x^2 - \overline{h}(x). \tag{2.11}$$

Let $\overline{C}(y)$ be the set of vertices of $G(n,(d_i)_1^n)$ starting from which y can be influenced, i.e., $\overline{C}(y) := \big\{x \in v(G(n,(d_i)_1^n)) : x \to y\big\}$. We have the following theorems for the dual backward propagation process.

Theorem 2.5 *Under assumptions of Theorem 2.2, there is a unique $\overline{\xi} \in (0,1)$ such that $\overline{H}(\overline{\xi}) = 0$ and there exists at least one y_n in $G^*(n,(d_i)_1^n)$ such that*

$$\frac{\big|\overline{C}(y_n)\big|}{n} \xrightarrow{p} 1 - \overline{g}(\overline{\xi}) > 0. \tag{2.12}$$

Remark that $\overline{H}(x) = 0$ is the same as equation (2.9) and therefore $\overline{\xi} \equiv \widetilde{p}_{ext}$ and $1 - \overline{g}(\overline{\xi}) \equiv p_{ext}$ from the branching process approximation.

We denote $\overline{C}(y_n)$ constructed in the proof of Theorem 2.5 by $\overline{C}^*$. For every $\epsilon > 0$, let

$$\overline{\mathbb{C}}^s(\epsilon) := \left\{ y \in v(G(n,(d_i)_1^n)) : \left|\overline{C}(y)\right| / n < \epsilon \right\},$$

and

$$\overline{\mathbb{C}}^L(\epsilon) := \left\{ y \in v(G(n,(d_i)_1^n)) : \left|\overline{C}(y) \vartriangle \overline{C}^*\right| / n < \epsilon \right\}.$$

Theorem 2.6 *Suppose that Condition 2.1 holds. If the condition (2.4) is satisfied then*

$$\forall \epsilon, \quad \frac{\left|\overline{\mathbb{C}}^s(\epsilon)\right| + \left|\overline{\mathbb{C}}^L(\epsilon)\right|}{n} \xrightarrow{p} 1, \tag{2.13}$$

and

$$\forall \epsilon, \quad \frac{\left|\overline{\mathbb{C}}^s(\epsilon)\right|}{n} \xrightarrow{p} 1 \tag{2.14}$$

otherwise.

Informally, the above theorem says that asymptotically ($n \to \infty$) and under the assumption (2.4), there is essentially one and only one big *source* component in the graph, to which a given vertex can possibly trace back while tracing all the possible sources of its influence. Moreover, if the assumption (2.4) is not satisfied, the chance of detecting a big source component when starting the backtracking from a randomly selected node goes to zero.

Finally, we have the following theorem which establishes the duality relation between the two processes.

Theorem 2.7 *Under assumptions of Theorem 2.2, for any $\epsilon > 0$ and $n \to \infty$,*

$$n^{-1}|\overline{\mathbb{C}}^L(\epsilon)| \left| n^{-1}|\overline{C}^*| - n^{-1}|\mathbb{C}^L(\epsilon)| \right| \leq \alpha\epsilon + R_n(\epsilon), \tag{2.15}$$

where $\alpha > 0$ and $R_n(\epsilon) \xrightarrow{p} 0$. The same statement holds with $\overline{\mathbb{C}}^L(\epsilon)$ exchanged with $\mathbb{C}^L(\epsilon)$ and $\overline{C}^*$ replaced by C^*.

The theorem leads to the following fundamental result of this chapter, where it all comes together and we are able to essentially identify, under one additional assumption apart from those in Theorem 2.2, the set of pioneers with the one big source component that we discovered above. In particular, this gives us

the relative size (w.r.t. n) of the set of pioneers, provided it is non-null, since we know the relative size of the source component.

Corollary 2.8 *Under assumptions of Theorem 2.2, for any $\epsilon > 0$ and $n \to \infty$, if there exists $a > 0$ such that $n^{-1}|\mathbb{C}^L(\epsilon)| > a$ whp, then*

$$n^{-1}|\mathbb{C}^L(\epsilon) \triangle \overline{C}^*| \leq \alpha'\epsilon + R'_n(\epsilon), \tag{2.16}$$

where $\alpha' > 0$ and $R'_n(\epsilon) \xrightarrow{p} 0$.

Remark 2 In particular, if $\mathbb{E}[D^{(t)}(D^{(t)}-2)] > 0$, then the configuration model with the degree sequence $(d_i^{(t)})_1^n$ will have a giant component $C^{(t)}$ whp. In other words, our enhanced configuration model will have a strongly connected giant component whp. In this case, whp $n^{-1}|\mathbb{C}^L(\epsilon)| \geq n^{-1}|C^{(t)}| > a$ for some $a > 0$, and thus the condition in the above corollary is satisfied.

2.2.4 Concluding Remarks

Let us conclude the presentation of the main results by the following remarks.

- Note that the condition (2.4), sufficient for the existence and essential uniqueness of the big influenced component, implies $\mathbb{E}[D(D-2)] > 0$. This latter condition is necessary and sufficient (provided Condition 2.1 holds; cf [38]) for the existence of a unique connected component of the underlying configuration model, usually called *big component*. Obviously, our big influenced component can exist only within this big component. Our condition (2.4) is also necessary in the sense explained in Remark 1.

- In contrast to [41, 21], we do *not* study the existence of a big *strongly connected* component (in which every node is reachable from every other node) in the directed graph, with which our enhanced configuration model can be identified, after replacing each edge through which the influence can travel in a given direction by a directed edge pointing in that direction (or, by two directed edges if the influence can travel in both directions), while deleting all edges which do not allow the influence to travel in either direction. Instead, we reveal the existence of two big *strongly interconnected components*: the set of pioneers and the set of influenced nodes, such that every pioneer can reach all influenced nodes and only these nodes, while every influenced node can be reached from and only from any pioneer. These two sets can have different size. In fact, numerical studies reported in

Chapter 3 show that in the case of node percolation there are more influenced nodes than pioneers, while the inverse is true for the coupon-collector dynamic. In the case of bond-percolation dynamic one can formally show that both sets have the same relative size, equal to the big component in the usual, non-directed bond percolation model (cf the coupling argument mentioned in Section 2.1.1 or direct calculations which follow next in this chapter). However, even in this case we do not know whether our two strongly interconnected sets coincide or even have a (relatively) big intersection, which would form a strongly connected component. We believe that this emergence of what we call two strongly interconnected components has not been observed earlier in any graph-theoretical model and therefore, adds a new dimension to our understanding of the phenomenon of viral information propagation in graph-theoretical models.

- There is a strong indication that in Corollary 2.8, we do not need the lower bound on $n^{-1}|\mathbb{C}^L(\epsilon)|$ for (2.16) to hold. One possible approach to prove this would be to make rigorous the branching process approximation heuristically illustrated in the previous section to provide insight (see [18], where the branching process approximation is used to find the largest component of Erdős-Rényi graph). This approach would most likely give only the lower bound on $n^{-1}|\mathbb{C}^L(\epsilon)|$ in Corollary 2.8, not the desired approximation of $n^{-1}|\mathbb{C}^L(\epsilon)|$ which we here obtain by the identification of $\mathbb{C}^L(\epsilon)$ with $\overline{C}^*$ in Corollary 2.8. Moreover, the introduction of the dual process which leads to the identification of $\mathbb{C}^L(\epsilon)$ with $\overline{C}^*$ is useful since this would provide us with important additional information regarding the structure of $\mathbb{C}^L(\epsilon)$, which we have not explored in this chapter.

  Another possible approach is to use a more general version of Glivenko-Cantelli lemma to study $\mathbb{C}^L(\epsilon_n)$, $\mathbb{C}^s(\epsilon_n)$, etc., where $\epsilon_n \to 0$ as $n \to \infty$.

2.3 Analysis of the Original Forward-Propagation Process

The following analysis is similar to the one presented in [38] and wherever the proofs of analogous lemmas, theorems etc. don't have any new point of note, we refer the reader to [38] without giving the proofs.

Throughout the construction and propagation process, we keep track of what we call *active transmitter* half-edges. To begin with, all the vertices and the attached half-edges are *sleeping* but once influenced, a vertex and its half-edges become *active*. Both sleeping and active half-edges at any time constitute what

we call *living* half-edges and when two half-edges are matched to reveal an edge along which the flow of influence has occurred, the half-edges are pronounced *dead*. Half-edges are further classified according to their ability or inability to transmit information as *transmitters* and *receivers* respectively. We initially give all the half-edges i.i.d. random maximal lifetimes with distribution given by $\tau \sim \exp(1)$, then go through the following algorithm.

C1 If there is no active half-edge (as in the beginning), select a sleeping vertex and declare it active, along with all its half-edges. For definiteness, we choose the vertex uniformly at random among all sleeping vertices. If there is no sleeping vertex left, the process stops.

C2 Pick an active transmitter half-edge and kill it.

C3 Wait until the next living half-edge dies (spontaneously, due to the expiration of its exponential lifetime). This is joined to the one killed in previous step to form an edge of the graph along which information has been transmitted. If the vertex it belongs to is sleeping, we change its status to active, along with all of its half-edges. Repeat from the first step.

Every time C1 is performed, we choose a vertex and trace the flow of influence from here onwards. Just before C1 is performed again, when the number of active transmitter half-edges goes to 0, we've explored the extent of the graph component that the chosen vertex can influence, that had not been previously influenced.

Let $S_T(t)$, $S_R(t)$, $A_T(t)$ and $A_R(t)$ represent the number of sleeping transmitter, sleeping receiver, active transmitter and active receiver half-edges, respectively, at time t. Therefore, $R(t) := A_R(t) + S_R(t)$ and $L(t) := A_T(t) + A_R(t) + S_T(t) + S_R(t) = A_T(t) + S_T(t) + R(t)$ denotes the number of receiver and living half-edges, respectively, at time t.

For definiteness, we will take them all to be right-continuous, which along with C1 entails that $L(0) = 2m - 1$. Subsequently, whenever a living half-edge dies spontaneously, C3 is performed, immediately followed by C2. As such, $L(t)$ is decreased by 2 every time a living half-edge dies spontaneously, up until the last living one die and the process terminates. Also remark that all the receiver half-edges, both sleeping and active, continue to die spontaneously.

The following consequences of *Glivenko-Cantelli* theorem are analogous to those given in [38] and we state them without proof.

Lemma 2.9 *As* $n \to \infty$,

$$\sup_{t\geq 0} \left| n^{-1} L(t) - \lambda e^{-2t} \right| \xrightarrow{p} 0. \tag{2.17}$$

Lemma 2.10 *As* $n \to \infty$,

$$\sup_{t\geq 0} \left| n^{-1} R(t) - \lambda_r e^{-t} \right| \xrightarrow{p} 0. \tag{2.18}$$

Let $V_{k,l}(t)$ be the number of sleeping vertices at time t which started with receiver and transmitter degrees k and l respectively . Clearly,

$$S_T(t) = \sum_{k,l} l V_{k,l}(t). \tag{2.19}$$

Among the three steps, only C1 is responsible for premature death (before the expiration of exponential life-time) of sleeping vertices. We first ignore its effect by letting $\widetilde{V}_{k,l}(t)$ be the number of vertices with receiver and transmitter degrees k and l respectively, such that all their half-edges would die spontaneously (without the aid of C1) after time t. Correspondingly, let $\widetilde{S}_T(t) = \sum_{k,l} l \widetilde{V}_{k,l}(t)$.

Then,

Lemma 2.11 *As* $n \to \infty$,

$$\sup_{t\geq 0} \left| n^{-1} \widetilde{V}_{k,l}(t) - p_{k,l} e^{-(k+l)t}) \right| \xrightarrow{p} 0. \tag{2.20}$$

for all $(k,l) \in \mathbb{N}^2$, *and*

$$\sup_{t\geq 0} \left| n^{-1} \sum_{k,l} \widetilde{V}_{k,l}(t) - g(e^{-t}, e^{-t}) \right| \xrightarrow{p} 0. \tag{2.21}$$

$$\sup_{t\geq 0} \left| n^{-1} \widetilde{S}_T(t) - h(e^{-t}) \right| \xrightarrow{p} 0. \tag{2.22}$$

Proof. Again, (2.20) follows from *Glivenko-Cantelli* theorem. To prove (2.22), note that by Condition 2.1(iii), $D_n = D_n^{(r)} + D_n^{(t)}$ are uniformly integrable, i.e., for every $\epsilon > 0$ there exists $K < \infty$ such that for all n,

$$\mathbb{E}(D_n; D_n > K) = \sum_{(k,l;k+l>K)} (k+l) \frac{u_{k,l}}{n} < \epsilon. \tag{2.23}$$

This, by Fatou's inequality, further implies that

$$\sum_{(k,l;k+l>K)} (k+l)p_{k,l} < \epsilon. \tag{2.24}$$

Thus, by (2.20), we have whp,

$$\begin{aligned}\sup_{t\geq 0}\left|n^{-1}\widetilde{S}_T(t)-h(e^{-t})\right| &= \sup_{t\geq 0}\left|\sum_{k,l} l(n^{-1}\widetilde{V}_{k,l}(t)-p_{k,l}e^{-(k+l)t})\right| \\ &\leq \sum_{(k,l;k+l\leq K)} l\sup_{t\geq 0}\left|(n^{-1}\widetilde{V}_{k,l}(t)-p_{k,l}e^{-(k+l)t})\right| + \\ &\quad \sum_{(k,l;k+l>K)} l(\frac{u_{k,l}}{n}+p_{k,l}) \\ &\leq \epsilon+\epsilon+\epsilon,\end{aligned}$$

which proves (2.22). A similar argument also proves (2.21).

Lemma 2.12 *If $d_{max} := \max_i d_i$ is the maximum degree of $G^*(n,(d_i)_1^n)$, then*

$$0 \leq \widetilde{S}_T(t)-S_T(t) < \sup_{0\leq s\leq t}(\widetilde{S}_T(s)+R(s)-L(s))+d_{max}. \tag{2.25}$$

Proof. Clearly, $V_{k,l}(t) \leq \widetilde{V}_{k,l}(t)$, and thus $S_T(t) \leq \widetilde{S}_T(t)$. Therefore, we have that $\widetilde{S}_T(t)-S_T(t) \geq 0$ and the difference increases only when C1 is performed. Suppose that happens at time t and a sleeping vertex of degree $j>0$ gets activated, then C2 applies immediately and we have $A_T(t) \leq j-1 < d_{max}$, and consequently,

$$\begin{aligned}\widetilde{S}_T(t)-S_T(t) &= \widetilde{S}_T(t)-(L(t)-R(t)-A_T(t)) \\ &< \widetilde{S}_T(t)+R(t)-L(t)+d_{max}.\end{aligned}$$

Since $\widetilde{S}_T(t)-S_T(t)$ does not change in the intervals during which C1 is not performed, $\widetilde{S}_T(t)-S_T(t) \leq \widetilde{S}_T(s)-S_T(s)$, where s is the last time before t that C1 was performed. The lemma follows.

Let

$$\widetilde{A}_T(t) := L(t) - R(t) - \widetilde{S}_T(t) = A_T(t) - (\widetilde{S}_T(t) - S_T(t)). \tag{2.26}$$

Then, Lemma 2.12 can be rewritten as

$$\widetilde{A}_T(t) \le A_T(t) < \widetilde{A}_T(t) - \inf_{s \le t} \widetilde{A}_T(s) + d_{max}. \tag{2.27}$$

Also, by Lemmas 2.9, 2.10 and 2.11 and (2.2),

$$\sup_{t \ge 0} \left| n^{-1} \widetilde{A}_T(t) - H(e^{-t}) \right| \xrightarrow{p} 0. \tag{2.28}$$

Lemma 2.13 *Suppose that Condition 2.1 holds and let $H(x)$ be given by (2.2).*

(i) If $\mathbb{E}[D^{(t)}D] > \mathbb{E}[D^{(t)}+D]$, then there is a unique $\xi \in (0,1)$, such that $H(\xi) = 0$; moreover, $H(x) < 0$ for $x \in (0,\xi)$ and $H(x) > 0$ for $x \in (\xi, 1)$.

(ii) If $\mathbb{E}[D^{(t)}D] \le \mathbb{E}[D^{(t)} + D]$, then $H(x) < 0$ for $x \in (0,1)$.

Proof. Remark that $H(0) = H(1) = 0$ and $H'(1) = 2\mathbb{E}[D] - \mathbb{E}[D^{(r)}] - \mathbb{E}[D^{(t)}D] = \mathbb{E}[D + D^{(t)}] - \mathbb{E}[D^{(t)}D]$. Furthermore we define $\phi(x) := H(x)/x = \lambda x - \lambda_r - \sum_{k,l} l p_{k,l} x^{k+l-1}$, which is a concave function on $(0,1]$, in fact, strictly concave unless $p_{k,l} = 0$ whenever $k + l \ge 3$ and $l \ge 1$, in which case $H'(1) = p_{0,1} + p_{1,1} + \sum_{k \ge 1} k p_{k,0} \ge p_{0,1} + p_{1,0} = \mathbb{P}(D = 1) > 0$, by Condition 2.1(iv).

In case (ii), we thus have ϕ concave and $\phi'(1) = H'(1) - H(1) \ge 0$, with either the concavity or the above inequality strict, and thus $\phi'(x) > 0$ for all $x \in (0,1)$, whence $\phi(x) < \phi(1) = 0$ for $x \in (0,1)$.

In case (i), $H'(1) < 0$, and thus $H(x) > 0$ for x close to 1. Further,

$$\begin{aligned} H'(0) &= -\lambda_r - \sum_{\{(k,l):k+l=1\}} l p_{k,l} \\ &= -\lambda_r - p_{0,1} \\ &\le -p_{1,0} - p_{0,1} < 0 \end{aligned}$$

by Condition 2.1(iv), which implies that $H(x) < 0$ for x close to 0. Hence there is at least one $\xi \in (0,1)$ with $H(\xi) = 0$. Now, since $H(x)/x$ is strictly concave and also $\phi(1) = H(1) = 0$, there is at most one such ξ. This proves the result.

Proof of Theorem 2.2. Let ξ be the zero of H given by Lemma 2.13(i) and let $\tau := -\ln \xi$. Then, by Lemma 2.13, $H(e^{-t}) > 0$ for $0 < t < \tau$, and thus $\inf_{t \le \tau} H(e^{-t}) = 0$. Consequently, (2.28) implies

$$n^{-1} \inf_{t \le \tau} \widetilde{A}_T(t) = n^{-1} \inf_{t \le \tau} \widetilde{A}_T(t) - \inf_{t \le \tau} H(e^{-t}) \xrightarrow{p} 0. \tag{2.29}$$

Further, by Condition 2.1(iii), $d_{max} = O(n^{1/2})$, and thus $n^{-1} d_{max} \to 0$. Consequently, by (2.27) and (2.29)

$$\sup_{t \le \tau} n^{-1} \left| A_T(t) - \widetilde{A}_T(t) \right| = \sup_{t \le \tau} n^{-1} \left| \widetilde{S}_T(t) - S_T(t) \right| \xrightarrow{p} 0, \tag{2.30}$$

and thus, by (2.28),

$$\sup_{t \ge 0} \left| n^{-1} A_T(t) - H(e^{-t}) \right| \xrightarrow{p} 0. \tag{2.31}$$

Let $0 < \epsilon < \tau/2$. Since $H(e^{-t}) > 0$ on the compact interval $[\epsilon, \tau - \epsilon]$, (2.31) implies that whp $A_T(t)$ remains positive on $[\epsilon, \tau - \epsilon]$, and thus C1 is not performed during this interval.

On the other hand, again by Lemma 2.13(i), $H(e^{-\tau-\epsilon}) < 0$ and (2.28) implies $n^{-1} \widetilde{A}_T(\tau+\epsilon) \xrightarrow{p} H(e^{-\tau-\epsilon})$, while $A_T(\tau + \epsilon) \ge 0$. Thus, with $\delta := |H(e^{-\tau-\epsilon})|/2 > 0$, whp

$$\widetilde{S}_T(\tau + \epsilon) - S_T(\tau + \epsilon) = A_T(\tau + \epsilon) - \widetilde{A}_T(\tau + \epsilon) \ge -\widetilde{A}_T(\tau + \epsilon) > n\delta, \tag{2.32}$$

while (2.30) implies that $\widetilde{S}_T(\tau) - S_T(\tau) < n\delta$ whp. Consequently, whp $\widetilde{S}_T(\tau + \epsilon) - S_T(\tau + \epsilon) > \widetilde{S}_T(\tau) - S_T(\tau)$, so C1 is performed between τ and $\tau + \epsilon$.

Let T_1 be the last time that C1 is performed before $\tau/2$, let x_n be the sleeping vertex declared active at this point of time and let T_2 be the next time C1 is performed. We have shown that for any $\epsilon > 0$, whp $0 \le T_1 \le \epsilon$ and $\tau - \epsilon \le T_2 \le \tau + \epsilon$; in other words, $T_1 \xrightarrow{p} 0$ and $T_2 \xrightarrow{p} \tau$.

We next use the following lemma.

Lemma 2.14 *Let T_1^* and T_2^* be two (random) times when C1 are performed, with $T_1^* \le T_2^*$, and assume that $T_1^* \xrightarrow{p} t_1$ and $T_2^* \xrightarrow{p} t_2$ where $0 \le t_1 \le t_2 \le \tau$. If C is the union of all the vertices informed between T_1^* and T_2^*, then*

$$|C|/n \xrightarrow{p} g(e^{-t_1}, e^{-t_1}) - g(e^{-t_2}, e^{-t_2}). \tag{2.33}$$

Proof. For all $t \geq 0$, we have

$$\sum_{i,j}(\widetilde{V}_{i,j}(t) - V_{i,j}(t)) \leq \sum_{i,j} j(\widetilde{V}_{i,j}(t) - V_{i,j}(t)) = \widetilde{S}_T(t) - S_T(t).$$

Thus,

$$\begin{aligned}|C| &= \sum (V_{k,l}(T_1^*-) - V_{k,l}(T_2^*-)) = \sum (\widetilde{V}_{k,l}(T_1^*-) - \widetilde{V}_{k,l}(T_2^*-)) + o_p(n) \\ &= ng(e^{-T_1^*}, e^{-T_1^*}) - ng(e^{-T_2^*}, e^{-T_2^*}) + o_p(n).\end{aligned}$$

Let C' be the set of vertices informed up till T_1 and C'' be the set of vertices informed between T_1 and T_2. Then, by Lemma 2.14, we have that

$$\frac{|C'|}{n} \xrightarrow{p} 0 \tag{2.34}$$

and

$$\frac{|C''|}{n} \xrightarrow{p} g(1,1) - g(e^{-\tau}, e^{-\tau}) = 1 - g(e^{-\tau}, e^{-\tau}). \tag{2.35}$$

Evidently, $C'' \subset C(x_n)$. Note that $C(x_n) = \left\{y \in v(G^*(n, (d_i)_1^n)) : x_n \to y\right\}$. It is clear that if $x_n \to y$, then $y \notin (C' \cup C'')^c$. Therefore, we have that $C(x_n) \subset C' \cup C''$, which implies that

$$\left|C''\right| \leq |C(x_n)| \leq \left|C'\right| + \left|C''\right|, \tag{2.36}$$

and thus, from (2.34) and (2.35),

$$\frac{|C(x_n)|}{n} \xrightarrow{p} 1 - g(e^{-\tau}, e^{-\tau}), \tag{2.37}$$

which completes the proof of Theorem 2.2.

Proof of Theorem 2.3. We continue from where we left in the proof of previous theorem, with the following Lemmas. Assumptions of Theorem 2.2 continue to hold for what follows in this section.

Lemma 2.15 *$\forall \epsilon > 0$, let*

$$\mathbb{A}(\epsilon) := \left\{ y \in v(G^*(n, (d_i)_1^n)) : \frac{|C(y)|}{n} \geq \epsilon \text{ and } \left| \frac{|C(y)|}{n} - (1 - g(\xi, \xi)) \right| \geq \epsilon \right\}.$$

Then,

$$\forall \epsilon, \quad \frac{|\mathbb{A}(\epsilon)|}{n} \xrightarrow{p} 0. \tag{2.38}$$

Proof. Suppose the converse is true. Then, there exists $\delta > 0$, $\delta' > 0$ and a sequence $(n_k)_{k>0}$ such that

$$\forall k, \quad \mathbb{P}\left(\frac{|\mathbb{A}(\epsilon)|}{n_k} > \delta\right) > \delta'. \tag{2.39}$$

Since the vertex initially informed to start the transmission process, say a, is uniformly chosen, we have

$$\forall n_k, \quad \mathbb{P}(a \in \mathbb{A}(\epsilon)) > \delta\delta' \tag{2.40}$$

and thus,

$$\forall k, \quad \mathbb{P}\left(\frac{|C'|}{n_k} \geq \epsilon \text{ or } \left|\frac{|C''|}{n_k} - (1 - g(\xi, \xi))\right| \geq \epsilon\right) > \delta\delta', \tag{2.41}$$

which contradicts (2.35).

Lemma 2.16 *For every $\epsilon > 0$, let*

$$\mathbb{B}(\epsilon) := \left\{y \in C' \cup C'' : |C(y)|/n \geq \epsilon \text{ and } |C(y) \triangle C^*|/n \geq \epsilon\right\}. \tag{2.42}$$

Then,

$$\forall \epsilon, \quad \frac{|\mathbb{B}(\epsilon)|}{n} \xrightarrow{p} 0. \tag{2.43}$$

Proof. Recall that for any three sets A, B and C, we have that $A \triangle B \subset (A \triangle C) \cup (B \triangle C)$. Therefore, for any $y \in C' \cup C''$, we have that

$$C(y) \triangle C^* \subset \left[C(y) \triangle (C' \cup C'')\right] \cup \left[C^* \triangle (C' \cup C'')\right]. \tag{2.44}$$

But recall that $C^* \subset C' \cup C''$ and by a similar argument, for every $y \in C' \cup C''$, $C(y) \subset C' \cup C''$. Thus,

$$C(y) \triangle C^* \subset \left[(C' \cup C'') \setminus C(y)\right] \cup \left[(C' \cup C'') \setminus C^*\right]. \tag{2.45}$$

Hence, if $|C(y) \triangle C^*|/n \geq \epsilon$, then either $|(C' \cup C'') \setminus C(y)|/n \geq \epsilon/2$ or $|(C' \cup C'') \setminus C^*|/n \geq \epsilon/2$.

Consequently,

$$\mathbb{B}(\epsilon) \subset \left\{ y \in v(G^*(n,(d_i)_1^n)) : \epsilon \le |C(y)|/n \le \left|(C' \cup C'')\right|/n - \epsilon/2 \right\} \\ \cup \left\{ y \in v(G^*(n,(d_i)_1^n)) : \left|(C' \cup C'') \setminus C^*\right|/n \ge \epsilon/2 \right\}.$$

Letting

$$e1 := \left| \left\{ y \in v(G^*(n,(d_i)_1^n)) : \epsilon \le |C(y)|/n \le \left|(C' \cup C'')\right|/n - \epsilon/2 \right\} \right| /n \\ E2 := \left\{ \left|(C' \cup C'') \setminus C^*\right|/n \ge \epsilon/2 \right\},$$

we have

$$\mathbb{B}(\epsilon)/n \le e1 + \mathbf{1}_{E2}. \tag{2.46}$$

Now, $e1 \xrightarrow{p} 0$ by (2.35) and Lemma 2.15, while $\mathbf{1}_{E2} \xrightarrow{p} 0$ because $\mathbb{P}(E2) \to 0$ by (2.34), (2.35) and (2.36). This concludes the proof.

Lemma 2.17 *Let T_3 be the first time after T_2 that C1 is performed and let z_n be the sleeping vertex activated at this moment. If C''' is the set of vertices informed between T_2 and T_3, then*

$$\frac{|C'''|}{n} \xrightarrow{p} 0. \tag{2.47}$$

Proof. Since $\widetilde{S}_T(t) - S_T(t)$ increases by at most $d_{max} = o_p(n)$ each time C1 is performed, we obtain that

$$\sup_{t \le T_3} (\widetilde{S}_T(t) - S_T(t)) \le \sup_{t \le T_2} (\widetilde{S}_T(t) - S_T(t)) + d_{max} = o_p(n). \tag{2.48}$$

Comparing this to (2.32) we see that for every $\epsilon > 0$, whp $\tau + \epsilon > T_3$. Since also $T_3 > T_2 \xrightarrow{p} \tau$, it follows that $T_3 \xrightarrow{p} \tau$. This in combination with Lemma 2.14 yields that

$$\frac{|C'''|}{n} \xrightarrow{p} 0.$$

Lemma 2.18 *For every $\epsilon > 0$, let*

$$\mathbb{C}(\epsilon) := \left\{ z \in \left(C' \cup C''\right)^c : |C(z)|/n \ge \epsilon \text{ and } |C(z) \,\triangle\, C^*|/n \ge \epsilon \right\}. \tag{2.49}$$

Then, we have that

$$\forall \epsilon, \quad \frac{|\mathbb{C}(\epsilon)|}{n} \xrightarrow{p} 0. \tag{2.50}$$

Proof. We start by remarking that by Lemma 2.15, it is sufficient to prove that

$$\frac{|\mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)|}{n} \xrightarrow{p} 0. \tag{2.51}$$

Now assume that there exist $\delta, \delta' > 0$ and a sequence $(n_k)_{k>0}$ such that

$$\forall k, \quad \mathbb{P}\left(\frac{|\mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)|}{n_k} > \delta\right) > \delta'. \tag{2.52}$$

Let

$$\mathscr{E}_1 := \{\text{Configuration Model completely revealed}\},$$

$$\mathscr{E}_2 := \{\text{Influence propagation revealed upto } C''\}$$

and $\mathscr{E}_3 := \mathscr{E}_1 \cap \mathscr{E}_2$.Then, denoting be z_{n_k} the vertex awakened by C1 at time T_2, we have that

$$\begin{aligned}
&\mathbb{P}\left(z_{n_k} \in \mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon) \middle| \mathscr{E}_3\right) \\
&\geq \frac{|\mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)|}{n_k - |C' \cup C''|} \mathbf{1}\left(\frac{|\mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)|}{n_k} > \delta\right) \\
&\geq \frac{|\mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)|}{n_k} \mathbf{1}\left(\frac{|\mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)|}{n_k} > \delta\right) \\
&\geq \delta \mathbf{1}\left(\frac{|\mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)|}{n_k} > \delta\right).
\end{aligned}$$

Taking expectations, we have

$$\mathbb{P}\left(z_{n_k} \in \mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)\right) \geq \delta\delta'. \tag{2.53}$$

But this leads to contradiction. Indeed, we have that

$$C(z_n) \triangle C^* \subset \left[C(z_n) \triangle (C' \cup C'' \cup C''')\right] \cup \left[C^* \triangle (C' \cup C'' \cup C''')\right]. \tag{2.54}$$

Again recall that $C^* \subset C' \cup C'' \cup C'''$ and by a similar argument, $C(z_n) \subset C' \cup C'' \cup C'''$ so that

$$C(z_n) \vartriangle C^* \subset \left[(C' \cup C'' \cup C''') \setminus C(z_n)\right] \cup \left[(C' \cup C'' \cup C''') \setminus C^*\right]. \tag{2.55}$$

Hence, if $|C(z_n) \vartriangle C^*| / n \geq \epsilon$, then either

$$\left|(C' \cup C'' \cup C''') \setminus C(z_n)\right| / n \geq \epsilon/2$$
$$\text{equivalently, } |C(z_n)| / n \leq \left|(C' \cup C'' \cup C''')\right| / n - \epsilon/2,$$

or,

$$\left|(C' \cup C'' \cup C''') \setminus C^*\right| / n \geq \epsilon/2.$$

Let

$$E3 := \left\{|C(z_n)| / n \leq \left|(C' \cup C'' \cup C''')\right| / n - \epsilon/2\right\}$$

and

$$E4 := \left\{\left|(C' \cup C'' \cup C''') \setminus C^*\right| / n \geq \epsilon/2\right\}.$$

Now assume that $z_n \in \mathbb{C}(\epsilon) \cap \mathbb{A}^c(\epsilon)$. This implies that either $E4$ holds or

$$\left\{1 - g(\xi, \xi) - \epsilon \leq \frac{|C(z_n)|}{n} \leq 1 - g(\xi, \xi) + \epsilon\right\} \cap E3$$

holds. But thanks to (2.34), (2.35) and Lemma 2.17, neither of these two events hold with asymptotically positive probability.

This completes the proof.

Finally, Lemma 2.16 and Lemma 2.18 allow us to state that

$$\forall \epsilon, \quad \frac{|\mathbb{C}^s(\epsilon)| + \left|\mathbb{C}^L(\epsilon)\right|}{n} \xrightarrow{p} 1, \tag{2.56}$$

which concludes the proof of Theorem 2.3.

Proof of Theorem 2.4. Following [38, proof of Theorem 2.3(ii)], let $T_0 = 0$ and T_1 be the next time C1 is performed. By (2.26) and the fact that $\widetilde{S}_T(t) - S_T(t)$ does not change in the intervals during which C1

is not performed,

$$\sup_{t\le T_2}\left|A_T(t)-\tilde{A}_T(t)\right| \le 2d_{max}\,. \tag{2.57}$$

Thus by (2.28) and and Lemma 2.13(ii) we have $1/n\tilde{A}_T(\epsilon) \xrightarrow{p} H(e^{-\epsilon}) < 0$, while $A(\epsilon) \ge 0$, and it follows from (2.57) that $T_2 \xrightarrow{p} 0$. Now if (2.7) does not hold then there exists a sub-sequence n_k for which the probability of choosing the first node (when performing C1 for the first time) outside $\mathbb{C}^s(\epsilon)$ is bounded away from zero. This contradicts $T_2 \xrightarrow{p} 0$ in view of Lemma 2.14.

2.4 Analysis of the Dual Back-Propagation Process

Now we introduce the algorithm to trace the possible sources of influence of a randomly chosen vertex. We borrow the terminology from the previous section, only in this case we put a *bar* over the label to indicate that we're talking about the dual process. The analysis also proceeds along the same lines as that of the original process, and we do not give the proof when it differs from the analogous proof in the previous section only by notation.

As before, we initially give all the half-edges i.i.d. random maximal lifetimes with distribution $\overline{\tau} \sim \exp(1)$ and then go through the following algorithm.

$\overline{\text{C}}1$ If there is no active half-edge (as in the beginning), select a sleeping vertex and declare it active, along with all its half-edges. For definiteness, we choose the vertex uniformly at random among all sleeping vertices. If there is no sleeping vertex left, the process stops.

$\overline{\text{C}}2$ Pick an active half-edge and kill it.

$\overline{\text{C}}3$ Wait until the next transmitter half-edge dies (spontaneously). This is joined to the one killed in previous step to form an edge of the graph. If the vertex it belongs to is sleeping, we change its status to active, along with all of its half-edges. Repeat from the first step.

Again, as before, $\overline{L}(0) = 2m-1$ and we have the following consequences of *Glivenko-Cantelli* theorem.

Lemma 2.19 *As* $n \to \infty$,

$$\sup_{t\ge 0}\left|n^{-1}\overline{L}(t)-\lambda e^{-2t}\right| \xrightarrow{p} 0. \tag{2.58}$$

Let $\overline{V}_{k,l}(t)$ be the number of sleeping vertices at time t which had receiver and transmitter degrees k and l respectively at time 0. It is easy to see that

$$\overline{S}(t) = \sum_{k,l} (ke^{-t} + l)\overline{V}_{k,l}(t). \tag{2.59}$$

Let $\widetilde{\overline{V}}_{k,l}(t)$ be the corresponding number if the impact of $\overline{\mathrm{C}}1$ on sleeping vertices is ignored. Correspondingly, let $\widetilde{\overline{S}}(t) = \sum_{k,l}(ke^{-t} + l)\widetilde{\overline{V}}_{k,l}(t)$.

Then,

Lemma 2.20 *As* $n \to \infty$,

$$\sup_{t\geq 0} \left| n^{-1}\widetilde{\overline{V}}_{k,l}(t) - p_{k,l}e^{-lt}) \right| \xrightarrow{p} 0. \tag{2.60}$$

for all $(k,l) \in \mathbb{N}^2$, *and*

$$\sup_{t\geq 0} \left| n^{-1} \sum_{k,l} \widetilde{\overline{V}}_{k,l}(t) - \overline{g}(e^{-t}) \right| \xrightarrow{p} 0. \tag{2.61}$$

$$\sup_{t\geq 0} \left| n^{-1}\widetilde{\overline{S}}(t) - \overline{h}(e^{-t}) \right| \xrightarrow{p} 0. \tag{2.62}$$

Proof. Again, (2.60) follows from *Glivenko-Cantelli* theorem.

To prove (2.62), note that by (3) of Condition(2.1), $D_n = D_n^{(r)} + D_n^{(t)}$ are uniformly integrable, i.e., for every $\epsilon > 0$ there exists $K < \infty$ such that for all n,

$$\mathbb{E}(D_n; D_n > K) = \sum_{(k,l;k+l>K)} (k+l)\frac{u_{k,l}}{n} < \epsilon. \tag{2.63}$$

This, by Fatou's inequality, further implies that

$$\sum_{(k,l;k+l>K)} (k+l)p_{k,l} < \epsilon. \tag{2.64}$$

Thus, by (2.60), we have whp,

$$\begin{aligned}\sup_{t\geq 0}\left|n^{-1}\widetilde{\overline{S}}(t)-\overline{h}(e^{-t})\right| &= \sup_{t\geq 0}\left|\sum_{k,l}(ke^{-t}+l)(n^{-1}\widetilde{\overline{V}}_{k,l}(t)-p_{k,l}e^{-lt})\right| \\ &\leq \sum_{(k,l;k+l\leq K)}(k+l)\sup_{t\geq 0}\left|(n^{-1}\widetilde{\overline{V}}_{k,l}(t)-p_{k,l}e^{-lt})\right| + \\ &\quad \sum_{(k,l;k+l>K)}(k+l)(\frac{u_{k,l}}{n}+p_{k,l}) \\ &\leq \epsilon+\epsilon+\epsilon,\end{aligned}$$

which proves (2.62). A similar argument also proves (2.61).

Lemma 2.21 *If $d_{max} := \max_i d_i$ is the maximum degree of $G^*(n,(d_i)_1^n)$, then*

$$0 \leq \widetilde{\overline{S}}(t)-\overline{S}(t) < \sup_{0\leq s\leq t}(\widetilde{\overline{S}}(s)-L(s))+d_{max}. \tag{2.65}$$

Proof. Clearly, $\overline{V}_{k,l}(t)\leq\widetilde{\overline{V}}_{k,l}(t)$, and thus $\overline{S}(t)\leq\widetilde{\overline{S}}(t)$. Therefore, we have that $\widetilde{\overline{S}}(t)-\overline{S}(t)\geq 0$ and the difference increases only when $\overline{\mathrm{C}}1$ is performed. Suppose that happens at time t and a sleeping vertex of degree $j>0$ gets activated, then $\overline{\mathrm{C}}2$ applies immediately and we have $\overline{A}(t)\leq j-1<d_{max}$, and consequently,

$$\begin{aligned}\widetilde{\overline{S}}(t)-\overline{S}(t) &= \widetilde{\overline{S}}(t)-(\overline{L}(t)-\overline{A}(t)) \\ &< \widetilde{\overline{S}}(t)-\overline{L}(t)+d_{max}.\end{aligned}$$

Since $\widetilde{\overline{S}}(t)-\overline{S}(t)$ does not change in the intervals during which $\overline{\mathrm{C}}1$ is not performed, $\widetilde{\overline{S}}(t)-\overline{S}(t)\leq\widetilde{\overline{S}}(s)-\overline{S}(s)$, where s is the last time before t that $\overline{\mathrm{C}}1$ was performed. The lemma follows.

Let

$$\widetilde{\overline{A}}(t) := \overline{L}(t)-\widetilde{\overline{S}}(t) = \overline{A}(t)-(\widetilde{\overline{S}}(t)-\overline{S}(t)). \tag{2.66}$$

Then, Lemma 2.21 can be rewritten as

$$\widetilde{\overline{A}}(t) \le \overline{A}(t) < \widetilde{\overline{A}}(t) - \inf_{s \le t} \widetilde{\overline{A}}(s) + d_{max}. \tag{2.67}$$

Also, by Lemmas 2.19 and 2.20 and (2.11),

$$\sup_{t \ge 0} \left| n^{-1} \widetilde{\overline{A}}(t) - \overline{H}(e^{-t}) \right| \xrightarrow{p} 0. \tag{2.68}$$

Lemma 2.22 *Suppose that Condition 2.1 holds and let $\overline{H}(x)$ be given by (2.11).*

(i) If $\mathbb{E}[D^{(t)}D] > \mathbb{E}[D^{(t)}+D]$, then there is a unique $\overline{\xi} \in (0,1)$, such that $\overline{H}(\overline{\xi}) = 0$; moreover, $\overline{H}(x) < 0$ for $x \in (0, \overline{\xi})$ and $\overline{H}(x) > 0$ for $x \in (\overline{\xi}, 1)$.

(ii) If $\mathbb{E}[D^{(t)}D] \le \mathbb{E}[D^{(t)} + D]$, then $\overline{H}(x) < 0$ for $x \in (0,1)$.

Proof. Remark that $\overline{H}(0) = \overline{H}(1) = 0$ and $\overline{H}'(1) = 2\mathbb{E}[D] - \mathbb{E}[(D^{(t)})^2] - \mathbb{E}[(D^{(r)})] - \mathbb{E}[D^{(r)}D^{(t)}] = \mathbb{E}[D + D^{(t)}] - \mathbb{E}[D^{(t)}D]$. Furthermore we define $\overline{\phi}(x) := \overline{H}(x)/x = \lambda x - \sum_{k,l} l p_{k,l} x^{l-1} - \sum_{k,l} k p_{k,l} x^l$, which is a concave function on $(0,1]$, in fact, strictly concave unless $p_{k,l} = 0$ whenever $l > 2$, or $l = 2$ and $k \ge 1$, in which case $\overline{H}'(1) = \sum_{k \ge 0} p_{k,1} + \sum_{k \ge 0} k p_{k,0} \ge p_{1,0} + p_{0,1} > 0$ by Condition 2.1(iv).

In case (ii), we thus have $\overline{\phi}$ concave and $\overline{\phi}'(1) = \overline{H}'(1) - \overline{H}(1) \ge 0$, with either the concavity or the above inequality strict, and thus $\overline{\phi}'(x) > 0$ for all $x \in (0,1)$, whence $\overline{\phi}(x) < \overline{\phi}(1) = 0$ for $x \in (0,1)$.

In case (i), $\overline{H}'(1) < 0$, and thus $\overline{H}(x) > 0$ for x close to 1. Further, in case (i),

$$\overline{H}'(0) = -\sum_k p_{k,1} - \sum_k k p_{k,0} \le -p_{1,0} - p_{0,1} < 0 \tag{2.69}$$

by Condition 2.1(iv), which implies that $\overline{H}(x) < 0$ for x close to 0. Hence there is at least one $\overline{\xi} \in (0,1)$ with $\overline{H}(\overline{\xi}) = 0$. Now, since $\overline{H}(x)/x$ is strictly concave and also $\overline{H}(1) = 0$, there is at most one such $\overline{\xi}$. This proves the result.

Proof of Theorem 2.5. Let $\overline{\xi}$ be the zero of $\overline{H}$ given by Lemma 2.22(i) and let $\overline{\tau} := -\ln \overline{\xi}$. Then, by Lemma 2.22, $\overline{H}(e^{-t}) > 0$ for $0 < t < \overline{\tau}$, and thus $\inf_{t \le \overline{\tau}} \overline{H}(e^{-t}) = 0$. Consequently, (2.68) implies

$$n^{-1} \inf_{t \le \overline{\tau}} \widetilde{\overline{A}}(t) = n^{-1} \inf_{t \le \overline{\tau}} \widetilde{\overline{A}}(t) - \inf_{t \le \overline{\tau}} \overline{H}(e^{-t}) \xrightarrow{p} 0. \tag{2.70}$$

Further, by Condition 2.1(iii), $d_{max} = O(n^{1/2})$, and thus $n^{-1}d_{max} \to 0$. Consequently, by (2.67) and (2.70)

$$\sup_{t\le\overline{\tau}} n^{-1}\left|\overline{A}(t)-\widetilde{\overline{A}}(t)\right| = \sup_{t\le\overline{\tau}} n^{-1}\left|\widetilde{\overline{S}}(t)-\overline{S}(t)\right| \xrightarrow{p} 0 \tag{2.71}$$

and thus, by (2.68),

$$\sup_{t\ge 0}\left|n^{-1}\overline{A}(t)-\overline{H}(e^{-t})\right| \xrightarrow{p} 0. \tag{2.72}$$

Let $0 < \epsilon < \overline{\tau}/2$. Since $\overline{H}(e^{-t}) > 0$ on the compact interval $[\epsilon, \overline{\tau}-\epsilon]$, (2.72) implies that whp $\overline{A}(t)$ remains positive on $[\epsilon, \overline{\tau}-\epsilon]$, and thus $\overline{\mathrm{C}}1$ is not performed during this interval.

On the other hand, again by Lemma 2.22(i), $\overline{H}(e^{-\overline{\tau}-\epsilon}) < 0$ and (2.68) implies $n^{-1}\widetilde{\overline{A}}(\overline{\tau}+\epsilon) \xrightarrow{p} \overline{H}(e^{-\overline{\tau}-\epsilon})$, while $\overline{A}(t)(\overline{\tau}+\epsilon) \ge 0$. Thus, with $\delta := \left|\overline{H}(e^{-\overline{\tau}-\epsilon})\right|/2 > 0$, whp

$$\widetilde{\overline{S}}(\overline{\tau}+\epsilon)-\overline{S}(\overline{\tau}+\epsilon) = \overline{A}(t)(\overline{\tau}+\epsilon)-\widetilde{\overline{A}}(\overline{\tau}+\epsilon) \ge -\widetilde{\overline{A}}(\overline{\tau}+\epsilon) > n\delta, \tag{2.73}$$

while (2.71) implies that $\widetilde{\overline{S}}(\overline{\tau})-\overline{S}(\overline{\tau}) < n\delta$ whp. Consequently, whp $\widetilde{\overline{S}}(\overline{\tau}+\epsilon)-\overline{S}(\overline{\tau}+\epsilon) > \widetilde{\overline{S}}(\overline{\tau})-\overline{S}(\overline{\tau})$, so $\overline{\mathrm{C}}1$ is performed between $\overline{\tau}$ and $\overline{\tau}+\epsilon$.

Let $\overline{T}_1$ be the last time that $\overline{\mathrm{C}}1$ is performed before $\overline{\tau}/2$, let y_n be the sleeping vertex declared active at this point of time and let $\overline{T}_2$ be the next time $\overline{\mathrm{C}}1$ is performed. We have shown that for any $\epsilon > 0$, whp $0 \le \overline{T}_1 \le \epsilon$ and $\overline{\tau}-\epsilon \le \overline{T}_2 \le \overline{\tau}+\epsilon$; in other words, $\overline{T}_1 \xrightarrow{p} 0$ and $\overline{T}_2 \xrightarrow{p} \overline{\tau}$.

We next use the following lamma.

Lemma 2.23 *Let $\overline{T}_1^*$ and $\overline{T}_2^*$ be two (random) times when $\overline{C}1$ are performed, with $\overline{T}_1^* \le \overline{T}_2^*$, and assume that $\overline{T}_1^* \xrightarrow{p} t_1$ and $\overline{T}_2^* \xrightarrow{p} t_2$ where $0 \le t_1 \le t_2 \le \overline{\tau}$. If $\overline{C}$ is the union of all the informer vertices reached between $\overline{T}_1^*$ and $\overline{T}_2^*$, then*

$$\left|\overline{C}\right|/n \xrightarrow{p} \overline{g}(e^{-t_1})-\overline{g}(e^{-t_2}). \tag{2.74}$$

Proof. For all $t \ge 0$, we have

$$\sum_{i,j}(\widetilde{\overline{V}}_{i,j}(t)-\overline{V}_{i,j}(t)) \le \sum_{i,j} j(\widetilde{\overline{V}}_{i,j}(t)-\overline{V}_{i,j}(t)) = \widetilde{\overline{S}}(t)-\overline{S}(t).$$

Thus,

$$\begin{aligned}\left|\overline{C}\right| &= \sum(\overline{V}_{k,l}(\overline{T}_1^*-) - \overline{V}_{k,l}(\overline{T}_2^*-)) = \sum(\widetilde{\overline{V}}_{k,l}(\overline{T}_1^*-) - \widetilde{\overline{V}}_{k,l}(\overline{T}_2^*-)) + o_p(n)\\ &= n\overline{g}(e^{-\overline{T}_1^*}) - n\overline{g}(e^{-\overline{T}_2^*}) + o_p(n).\end{aligned}$$

Let $\overline{C}'$ be the set of possible influence sources traced up till $\overline{T}_1$ and $\overline{C}''$ be the set of those traced between $\overline{T}_1$ and $\overline{T}_2$. Then, by Lemma 2.23, we have that

$$\frac{\left|\overline{C}'\right|}{n} \xrightarrow{p} 0 \tag{2.75}$$

and

$$\frac{\left|\overline{C}''\right|}{n} \xrightarrow{p} \overline{g}(1) - \overline{g}(e^{-\overline{\tau}}) = 1 - \overline{g}(e^{-\overline{\tau}}). \tag{2.76}$$

Evidently, $\overline{C}'' \subset \overline{C}(y_n)$ and $\overline{C}(y_n) \subset \overline{C}' \cup \overline{C}''$, therefore

$$\left|\overline{C}''\right| \leq \left|\overline{C}(y_n)\right| \leq \left|\overline{C}'\right| + \left|\overline{C}''\right| \tag{2.77}$$

and thus, from (2.75) and (2.76),

$$\frac{\left|\overline{C}(y_n)\right|}{n} \xrightarrow{p} 1 - \overline{g}(e^{-\overline{\tau}}), \tag{2.78}$$

which completes the proof.

Proof of Theorem 2.6. As in the previous section, we have the following set of Lemmas, which we state without proof since the only change is notational. As before, assumptions of Theorem 2.2 continue to hold.

Lemma 2.24 $\forall \epsilon > 0$, *let*

$$\overline{\mathbb{A}}(\epsilon) := \left\{ x \in \nu(G^*(n,(d_i)_1^n)) : \frac{\left|\overline{C}(x)\right|}{n} \geq \epsilon \text{ and } \left| \frac{\left|\overline{C}(x)\right|}{n} - (1 - \overline{g}(\overline{\xi})) \right| \geq \epsilon \right\}.$$

Then,

$$\forall \epsilon, \quad \frac{\left|\overline{\mathbb{A}}(\epsilon)\right|}{n} \xrightarrow{p} 0. \tag{2.79}$$

Lemma 2.25 *For every $\epsilon > 0$, let*

$$\overline{\mathbb{B}}(\epsilon) := \left\{ x \in \overline{C}' \cup \overline{C}'' : \left|\overline{C}(x)\right| / n \geq \epsilon \text{ and } \left|\overline{C}(x) \vartriangle \overline{C}^*\right| / n \geq \epsilon \right\}. \tag{2.80}$$

Then,

$$\forall \epsilon, \quad \frac{\left|\overline{\mathbb{B}}(\epsilon)\right|}{n} \xrightarrow{p} 0. \tag{2.81}$$

Lemma 2.26 *Let $\overline{T}_3$ be the first time after $\overline{T}_2$ that $\overline{C}1$ is performed and let w_n be the sleeping vertex activated at this moment. If $\overline{C}'''$ is the set of informer vertices reached between $\overline{T}_2$ and $\overline{T}_3$, then*

$$\frac{\left|\overline{C}'''\right|}{n} \xrightarrow{p} 0. \tag{2.82}$$

Lemma 2.27 *For every $\epsilon > 0$, let*

$$\overline{\mathbb{C}}(\epsilon) := \left\{ w \in \left(\overline{C}' \cup \overline{C}''\right)^c : \left|\overline{C}(w)\right| / n \geq \epsilon \text{ and } \left|\overline{C}(w) \vartriangle \overline{C}^*\right| / n \geq \epsilon \right\}. \tag{2.83}$$

Then, we have that

$$\forall \epsilon, \quad \frac{\left|\overline{\mathbb{C}}(\epsilon)\right|}{n} \xrightarrow{p} 0. \tag{2.84}$$

Finally, Lemma 2.25 and Lemma 2.27 allow us to conclude that

$$\forall \epsilon, \quad \frac{\left|\overline{\mathbb{C}}^s(\epsilon)\right| + \left|\overline{\mathbb{C}}^L(\epsilon)\right|}{n} \xrightarrow{p} 1. \tag{2.85}$$

The proof of (2.14) goes along the same lines as the proof of Theorem 2.4.

2.5 Duality Relation

The forward and backward processes are linked through the tautology: $y \in C(x) \iff x \in \overline{C}(y)$. To prove the Theorem 2.7, we consider the double sum: $\sum_{x,y \in \nu(G(n,(d_i)_1^n))} \mathbf{1}(y \in C(x))$.

From here onwards, we abridge $\nu(G(n,(d_i)_1^n))$ to $\nu(G)$. Assumptions of Theorem 2.2 continue to hold throughout this section. We start with the following Proposition.

Proposition 2.28 *We have,*

$$\mathbf{A}_n := \left| n^{-2} \sum_{x,y \in \nu(G)} \mathbf{1}(y \in C(x)) - n^{-2} \sum_{x,y \in \nu(G)} \mathbf{1}(x \in \overline{C}^*)\mathbf{1}(y \in C(x) \cap C^*) \right| \xrightarrow{p} 0,$$

when $n \to \infty$.

Proof. The Proposition follows from the following two Lemmas.

Lemma 2.29 *For any* $\epsilon > 0$ *and* $n \to \infty$,

$$\left| n^{-2} \sum_{x,y \in \nu(G)} \mathbf{1}(y \in C(x)) - n^{-2} \sum_{x,y \in \nu(G)} \mathbf{1}(y \in C(x) \cap C^*) \right| \leq 2\epsilon + R_n^1(\epsilon),$$

where $R_n^1(\epsilon) \xrightarrow{p} 0$.

Proof. For $\epsilon > 0$, we have

$$\begin{aligned}
& \left| n^{-2} \sum_{x,y} \mathbf{1}(y \in C(x)) - n^{-2} \sum_{x,y} \mathbf{1}(y \in C(x) \cap C^*) \right| \\
\leq \; & n^{-2} \sum_{x} \min(|C(x)|, |C(x) \vartriangle C^*|) \\
= \; & n^{-2} \sum_{x \in \mathbb{C}^s(\epsilon)} \min(|C(x)|, |C(x) \vartriangle C^*|) \\
& + n^{-2} \sum_{x \in \mathbb{C}^L(\epsilon)} \min(|C(x)|, |C(x) \vartriangle C^*|) \\
& + n^{-2} \sum_{x \notin \mathbb{C}^s(\epsilon) \cup \mathbb{C}^L(\epsilon)} \min(|C(x)|, |C(x) \vartriangle C^*|) \\
\leq \; & n^{-1} \sum_{x \in \mathbb{C}^s(\epsilon)} \epsilon + n^{-1} \sum_{x \in \mathbb{C}^L(\epsilon)} \epsilon + n^{-1} \sum_{x \notin \mathbb{C}^s(\epsilon) \cup \mathbb{C}^L(\epsilon)} 1 \\
\leq \; & \epsilon + \epsilon + \left(1 - \frac{|\mathbb{C}^s(\epsilon)| + |\mathbb{C}^L(\epsilon)|}{n} \right).
\end{aligned}$$

Taking $R_n^1(\epsilon) := 1 - \frac{|\mathbb{C}^s(\epsilon)| + |\mathbb{C}^L(\epsilon)|}{n}$ and using Theorem 2.3, we conclude the proof.

Lemma 2.30 *For any $\epsilon > 0$ and $n \to \infty$,*

$$\begin{aligned}
& \left| n^{-2} \sum_{x,y \in v(G)} \mathbf{1}(y \in C(x) \cap C^*) - n^{-2} \sum_{x,y \in v(G)} \mathbf{1}(x \in \overline{C}^*) \mathbf{1}(y \in C(x) \cap C^*) \right| \\
& \leq 2\epsilon + R_n^2(\epsilon),
\end{aligned}$$

where $R_n^2(\epsilon) \xrightarrow{p} 0$.

Proof. Since $y \in C(x) \iff x \in \overline{C}(y)$, we have

$$\sum_{x,y \in v(G)} \mathbf{1}(y \in C(x) \cap C^*) = \sum_{x,y \in v(G)} \mathbf{1}(x \in \overline{C}(y)) \mathbf{1}(y \in C^*) \tag{2.86}$$

and

$$\sum_{x,y} \mathbf{1}(x \in \overline{C}^*) \mathbf{1}(y \in C(x) \cap C^*) = \sum_{x,y} \mathbf{1}(x \in \overline{C}(y) \cap \overline{C}^*) \mathbf{1}(y \in C^*). \tag{2.87}$$

Consequently,

$$\begin{aligned}
&\left| n^{-2} \sum_{x,y} \mathbf{1}(y \in C(x) \cap C^*) - n^{-2} \sum_{x,y} \mathbf{1}(x \in \overline{C}^*) \mathbf{1}(y \in C(x) \cap C^*) \right| \\
&\leq n^{-2} \sum_{y} \mathbf{1}(y \in C^*) \min\left(|\overline{C}(y)|, |\overline{C}(y) \vartriangle \overline{C}^*| \right) \\
&\leq n^{-2} \sum_{y} \min\left(|\overline{C}(y)|, |\overline{C}(y) \vartriangle \overline{C}^*| \right).
\end{aligned}$$

The result follows by the arguments similar to those in the proof of Lemma 2.29, with $R_n^2(\epsilon) := 1 - \frac{|\overline{\mathbb{C}}^s(\epsilon)| + |\overline{\mathbb{C}}^L(\epsilon)|}{n}$.

Next, we have the following two Propositions, which lead to Theorem 2.7.

Proposition 2.31 *For any $\epsilon > 0$ and $n \to \infty$,*

$$\left| n^{-1} |\mathbb{C}^L(\epsilon)| - n^{-1} |\mathbb{C}^L(\epsilon) \cap \overline{C}^*| \right| \leq \alpha^1 \epsilon + R_n^3(\epsilon), \tag{2.88}$$

where $\alpha^1 > 0$ is a constant and $R_n^3(\epsilon) \xrightarrow{p} 0$. Analogously,

$$\left| n^{-1} |\overline{\mathbb{C}}^L(\epsilon)| - n^{-1} |\overline{\mathbb{C}}^L(\epsilon) \cap C^*| \right| \leq \alpha^2 \epsilon + R_n^4(\epsilon) \tag{2.89}$$

where $\alpha^2 > 0$ is a constant and $R_n^4(\epsilon) \xrightarrow{p} 0$.

Proof. Remark that

$$\begin{aligned}
\sum_{x,y \in v(G)} \mathbf{1}(y \in C(x)) = &\sum_{x \in v(G), y \in \overline{\mathbb{C}}^L(\epsilon)} \mathbf{1}(x \in \overline{C}(y)) + \sum_{x \in v(G), y \in \overline{\mathbb{C}}^s(\epsilon)} \mathbf{1}(x \in \overline{C}(y)) \\
&+ \sum_{x \in v(G), y \notin \overline{\mathbb{C}}^s(\epsilon) \cup \overline{\mathbb{C}}^L(\epsilon)} \mathbf{1}(x \in \overline{C}(y)).
\end{aligned}$$

Therefore, using the arguments similar to those in the proof of Lemma 2.29, we have

$$\left| n^{-2} \sum_{x,y \in v(G)} \mathbf{1}(y \in C(x)) - n^{-2} |\overline{C}^*| . |\overline{\mathbb{C}}^L(\epsilon)| \right| \leq 2\epsilon + R_n^2(\epsilon). \tag{2.90}$$

In the same way,

$$\left| n^{-2} \sum_{x,y \in v(G)} \mathbf{1}(y \in C^*) \mathbf{1}(x \in \overline{C}(y) \cap \overline{C}^*) - n^{-2} |\overline{C}^*| . |\overline{\mathbb{C}}^L(\epsilon) \cap C^*| \right| \leq 2\epsilon + R_n^2(\epsilon).$$

From the above two equations and using Proposition 2.28, we have

$$\left| n^{-2} |\overline{C}^*| . |\overline{\mathbb{C}}^L(\epsilon)| - n^{-2} |\overline{C}^*| . |\overline{\mathbb{C}}^L(\epsilon) \cap C^*| \right| \leq 4\epsilon + 2R_n^2(\epsilon) + A_n.$$

Now using Theorem 2.2 and taking $\alpha^2 := \frac{5}{1-\overline{g}(\overline{\xi},\overline{\xi})}$ and $R_n^4(\epsilon) = \frac{3R_n^2(\epsilon)+2A_n}{1-\overline{g}(\overline{\xi},\overline{\xi})}$, we have the second part of the proposition. The proof of the first part is similar, with $\alpha^1 := \frac{5}{1-g(\xi,\xi)}$ and $R_n^3(\epsilon) = \frac{3R_n^1(\epsilon)+2A_n}{1-g(\xi,\xi)}$.

Proposition 2.32 *For any $\epsilon > 0$,*

$$\left| n^{-2} \sum_{x,y \in v(G)} \mathbf{1}(y \in C(x)) - n^{-2} |C^* \cap \overline{\mathbb{C}}^L(\epsilon)| . |\mathbb{C}^L(\epsilon)| \right| \leq 3\epsilon + R_n^1(\epsilon) + R_n^2(\epsilon)$$

Proof. We can upper bound the double sum thus,

$$\begin{aligned}\sum_{x,y\in v(G)} \mathbf{1}(y \in C(x)) \leq & \sum_{x\in\mathbb{C}^L(\epsilon), y\in\overline{\mathbb{C}}^L(\epsilon)} \mathbf{1}(y \in C(x)) \\ & + \sum_{x\in\mathbb{C}^s(\epsilon), y\in v(G)} \mathbf{1}(y \in C(x)) \\ & + \sum_{x\in v(G), y\in\overline{\mathbb{C}}^s(\epsilon)} \mathbf{1}(y \in C(x)) \\ & + \sum_{x\notin\mathbb{C}^s(\epsilon)\cup\mathbb{C}^L(\epsilon), y\in v(G)} \mathbf{1}(y \in C(x)) \\ & + \sum_{x\in v(G), y\notin\mathbb{C}^s(\epsilon)\cup\mathbb{C}^L(\epsilon)} \mathbf{1}(y \in C(x)).\end{aligned}$$

The result follows, once again, by using the arguments similar to those in the proof of Lemma 2.29.

Proof of Theorem 2.7. Now, from Proposition 2.32 and (2.89) and (2.90) from Proposition 2.31, we can conclude the proof of Theorem 2.7, with $\alpha := 5 + \alpha^1$ and $R_n(\epsilon) := R_n^1(\epsilon) + 2R_n^2(\epsilon) + R_n^4(\epsilon)$. The second statement of Theorem 2.7 holds by the symmetry of the model.

Corollary 2.8 follows from both statements of Theorem 2.7.

3

Viral Marketing: Examples, Applications and Numerical Studies

In a viral marketing campaign, a firm initially targets a small set of pioneers and hopes that they would influence a sizeable fraction of the population by diffusion of influence through the network. In general, any marketing campaign might fail to go viral in the first try. As such, it would be useful to have some guide to evaluate the effectiveness of the campaign and judge whether it is worthy of further resources, and in case the campaign has potential, how to hit upon a good pioneer who can make the campaign go viral.

In this chapter, we use the results of the previous chapter and some key illustrative examples to provide an insight from a firm's perspective regarding how to evaluate the effectiveness of a marketing campaign and do cost-benefit analysis by collecting relevant statistical data from the pioneers it selects.

3.1 Introduction

The penetration of internet and the emergence of huge online social networks in the last decade has radically altered the way that people consume media and print, leading to an ongoing decline in importance of conventional channels and consequently, marketing through them. This radical shift has brought in its wake a host of opportunities as well as challenges for the advertisers. On the one hand, firms finally have the possibility to reach in a cost-effective way not only the past responsive customers, but indeed all the potentially responsive ones. But the firms, in general, have found it a hit-or-miss game to gain attention through the new medium, with the chance of a hit depending on the loyaity of its fan-base. What makes viral marketing tempting for the firms is that when it is a hit, it is a spectacular one. But continuing to spend resources on a campaign while hoping that it goes viral is a precarious strategy. The *fat-tail* uncertainty of viral marketing makes it inherently different from conventional marketing and calls for a fundamentally different approach to decision-making: which individuals, and how many, to initially target in the online network? What amount of resources to spend on these initially targeted *pioneers*? And most importantly, when to stop, admit the inefficacy of the current campaign and develop a new one?

Configuration model serves as a useful first approximation of an online social network, particularly when one does not have an access to a detailed information about the network structure. The diffusion dynamic studied in Chapter 2 can be stated in this context as follows: any individual in the network influences a random subset of its neighbours, the distribution of which depends on the effectiveness of the marketing campaign.

We first illustrate large-network-limit results proved in Chapter 2 for configuration model having two types of degree-distribution: Poisson and Power Law. Three examples illustrating the dynamic of influence propagation on these two networks are considered: (1) Bernoulli transmissions; (2) Node percolation; (3) Coupon-collector transmissions.

Based on the above analysis, we offer a practical decision-making guide for marketing on online networks which could be useful to firms with no prior access to detailed network structure. Specifically, we consider the naïve strategy of picking some number of pioneers at random from the population, spending some fixed amount of resources on each of them and waiting to see if the campaign goes viral, picking another batch if it does not. For this strategy, we suggest what statistical data the firm should collect from its pioneers, and based on these, how to estimate the effectiveness of the campaign and make a cost-benefit analysis.

We will continue to use the notation introduced in the previous chapter.

3.1.1 Related Work

The phenomenon of viral propagation was first studied in the context of the spread of epidemics on complex networks, whence the term *viral* marketing originates ([6], [50]). The impact of social network structure on the propagation of social and economic behavior has also been recognized ([8], [52]) and there is growing evidence of its importance ([7]).

In the context of viral marketing, the principal approach which has been developed tries to exploit the network structure to maximize the probability of marketing campaign going viral for each dollar spent. This approach relies on the availability of large databases containing detailed information regarding the network structure and the past instances of influence propagation to come up with the best predictor of the most influential individuals who should be targeted for future campaigns ([42], [25]). In our approach, we do not rely on locating the pioneers by data-mining the network. Instead, we suggest a way to measure the current campaign's effectiveness based on its ongoing diffusion in the network. This idea can possibly complement the data mining approach when the network information is freely accessible.

3.2 Examples

Let us consider the results of Chapter 2 in the context of a few illustrative network examples. In what follows, denote by $G_D(x) = \mathbb{E}[x^D]$ and $G_{D^{(t)}}(x) = \mathbb{E}[x^{D^{(t)}}]$ the probability generating function (pgf) of D and $D^{(t)}$, respectively. In the notation of Chapter 2, this means that

$$G_D(x) = g(x, x) \tag{3.1}$$

and

$$G_{D^{(t)}}(x) = \overline{g}(x). \tag{3.2}$$

3.2.1 Bernoulli transmissions

Let us assume some arbitrary distribution of the degree D satisfying Condition 1.3 (to guarantee the existence of the big component of the social graph). Suppose that each user decides independently for

each of its friends with probability $p \in [0,1]$ whether to transmit the influence to him or not. We call this model *CM with Bernoulli transmissions* and p the *transmission probability*. Note that given the total degree D, the transmitter degree $D^{(t)}$ is Binomial(D,p) random variable.

Proposition 3.1 *In the CM with a general degree distribution D satisfying Condition 1.3 and Bernoulli transmissions, the campaign can go viral if and only if the transmission probability p satisfies*

$$p > \frac{\mathbb{E}[D]}{\mathbb{E}[D^2]-\mathbb{E}[D]}\,. \tag{3.3}$$

In this latter case the fraction of the influenced population and the fraction of good pioneers are asymptotically equal to each other for large n. More precisely, when $n \to \infty$,

$$|C^*|/n, |\bar{C}^*|/n \xrightarrow{p} 1-G_D(\xi)\,, \tag{3.4}$$

where ξ is the unique zero of the function

$$\mathbb{E}[D]((x-1)/p+1)-G_D'(x)$$

in $(0,1)$.

Proof. Bernoulli transmissions along with (2.2) and (2.11) imply $H(x) = \mathbb{E}[D]x^2-(1-p)\mathbb{E}[D]x - pxG_D'(x)$ and $\overline{H}(x) = \mathbb{E}[D]x^2-G_D'(1-p(1-x))$. Moreover $\overline{G}_{D^{(t)}}(x) = G_D(1-p(1-x))$. Dividing $H(x)$ by px and substituting $y := 1-p(1-x)$ in $\overline{H}(x)$ and $\overline{G}_{D^{(t)}}(x)$ completes the proof.

Consider two specific network degree examples.

Example 3.2.1 (Poisson degree) *When D has Poisson distribution of parameter λ (in which case the CM is asymptotically equivalent to the Erdős-Rényi model in the local weak sense as defined in refSi.lwc) the condition (3.3) reduces to*

$$\lambda p > 1$$

and by (3.4), the fraction of the influenced population and good pioneers, whp, is equal to $(1-\xi)/p$,

where ξ is the unique zero of the function

$$(x-1)/p+1-\exp(\lambda(x-1)$$

in $(0,1)$.

More commonly observed degree-distributions in social networks have power-law tails.

Example 3.2.2 (Power-Law ("zipf") degree) *Assume D having distribution*

$$\mathbb{P}\{D=k\}=k^{-\beta}/\zeta(\beta) \quad k=1,2,\dots,$$

with $\beta>2$, where $\zeta(\beta)$ is the zeta function. Recall that the pgf of D is equal to $G_D(x)=Li_\beta(x)/\zeta(\beta)$, where $Li_\beta(x)=\sum_{k=1}^{\infty}k^{-\beta}x^k$ is the so-called poly-logarithmic function. In this case, condition (1.18) for the existence of the big component is equivalent to

$$\zeta(\beta-2)-2\zeta(\beta-1)>0,$$

which is approximately $\beta<3.48$. Condition (3.3) reduces to

$$p>\zeta(\beta-1)/(\zeta(\beta-2)-\zeta(\beta-1))$$

and by (3.4), the fraction of the influenced population and good pioneers, whp, is equal to $1-Li_\beta(\xi)$, where ξ is the unique zero of the function

$$x\zeta(\beta-1)((x-1)/p+1)-Li_{\beta-1}(x)$$

in $(0,1)$.

Recall from Proposition 3.1, that Bernoulli transmissions lead to the model where the fraction of the influenced population and the fraction of good pioneers are asymptotically equal to each other. In what follows we present two scenarios where the set of good pioneers and the influenced population have different size.

3.2.2 Enthusiastic and apathetic users or node percolation

Consider CM with a general degree distribution D satisfying (2.4), whose nodes either transmit the influence to all their friends (these are "enthusiastic" nodes) or do not transmit to any of their friends ("apathetic" ones). Let p denote the fraction of nodes in the network which are enthusiastic. Note that this model corresponds to the *node-percolation* [1] on the CM. Thus, in this model, given D, $D^{(t)} = D$ with probability p and $D^{(t)} = 0$ with probability $1-p$.

Proposition 3.2 *Consider node-percolation on the CM with a general degree distribution D satisfying Condition 1.3. The campaign can go viral if and only if the fraction p of enthusiastic users satisfies condition (3.3); the same as for the Bernoulli model. Moreover, in this case, the fraction of reached population, say α, is also the same whp as in the network with Bernoulli transmissions, i.e., as given in Proposition 3.1. However, the fraction of good pioneers, say $\overline{\alpha}$, whp is given by $\overline{\alpha} = 1-p(1-\alpha)$.*

Proof. Node percolation along with (2.2) and (2.11) imply $H(x) = \overline{H}(x) = \mathbb{E}[D]x^2 - (1-p)\mathbb{E}[D]x - pxG'_D(x)$. Moreover $\overline{G}_{D^{(t)}}(x) = pG_D(x)$. Substituting in Theorem 2.2 and Theorem 2.5 completes the proof.

Note that the *campaign on the network with enthusiastic and apathetic users can reach the same population as in the Bernoulli transmissions, however there are less good pioneers*.

3.2.3 Absentminded users or coupon-collector transmissions

Consider again CM with a general degree distribution D satisfying (2.4). Suppose that each user is willing (or allowed) to transmit K messages of influence. In this regard, it randomly selects K times one of his friends *with replacement* (as if he were forgetting his previous choices). An equivalent dynamic of the influence propagation can be formulated as follows: every influenced user, at all times, keeps choosing one of its friends uniformly at random and transmits the influence to him; it stops forwarding the influence after K transmissions.

In this model the transmission degree $D^{(t)}$ correspond to the number of collected coupons in the classical coupon collector problem with the number of coupons being the vertex degree D and the number of trials

[1] different than edge-percolation

K. The conditional distribution of $D^{(t)}$ given D can be expressed as follows:

$$\mathbb{P}\{\, D^{(t)} = k \,|\, D \,\} = \frac{D!}{(D-k)!D^{-K}} \left\{ {K \atop k} \right\},$$

where $\{ {K \atop k} \} = 1/k! \sum_{i=0}^{K} (-1)^i \binom{k}{i} (k-i)^K$ is the Stirling number of the second kind.

Calculating the pgf for this distribution is tedious and we do not present analytical results regarding this model but only simulations and estimation. As we shall see in Section 3.2.4, in this model *the influenced population is smaller than the population of good pioneers.*

3.2.4 Numerical examples

We will present now a few numerical examples of networks and diffusion models presented above.

3.2.4.1 Simulations

In all our examples we simulate the enhanced configuration model on $N = 1000$ nodes assuming some particular node degree D distribution and influence propagation mechanism modeled by the conditional distribution of the transmitter degree $D^{(t)}$. More precisely, we sample the individual node degrees and transmitter degrees $(D_i, D_i^{(t)})$ $i = 1 \ldots N$ independently from the joint distribution of $(D, D^{(t)})$ and use these values to construct an instance of our enhanced CM by uniform pairwise matching of the half-edges. We calculate the relative size of the influenced population and the set of good pioneers through the exploration of the influenced components for all nodes. [1] In fact, relative sizes of the populations reached from different pioneers concentrate very clearly, as shown on Figure 3.1, which illustrates the claims in Chapter 2.

[1] The simulations are run in *python* using the *networkx* package.

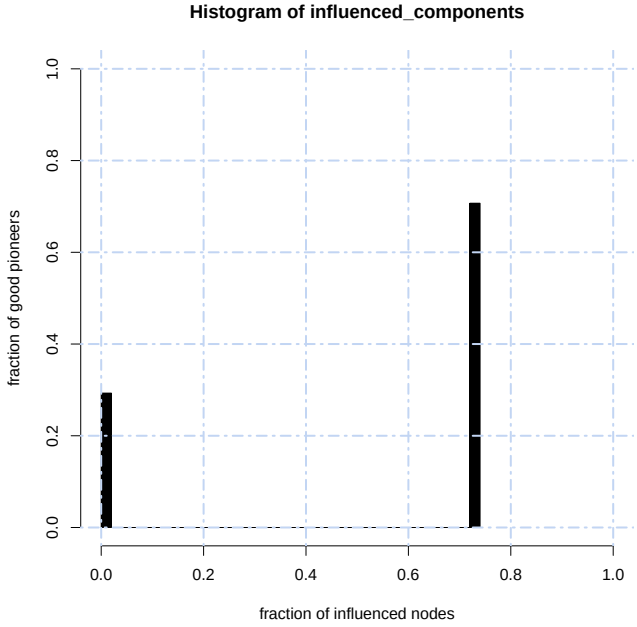

Figure 3.1: Concentration of the relative sizes of populations reached from different pioneers. CM with Poisson degree of mean $\lambda = 2$ and Bernoulli transmissions with $p = 0.8$.

3.2.4.2 Estimation

We adopt also the following "semi-analytic" approach: Using the sample $(D_i, D_i^{(t)})$, $i = 1, \ldots, N$ used to construct the CM, we consider estimators

$$\hat{G}_D(x) := \frac{1}{N} \sum_{i=1}^{N} x^{D_i} \tag{3.5}$$

$$\hat{G}_D^{(t)}(x) := \frac{1}{N} \sum_{i=1}^{N} x^{D_i^{(t)}} \tag{3.6}$$

$$\hat{H}(x) := \frac{1}{N} \sum_{i=1}^{N} \Big(D_i x^2 - (D_i - D_i^{(t)})x - D_i^{(t)} x^{D_i} \Big) \tag{3.7}$$

$$\hat{\overline{H}}(x) := \sum_{i=1}^{N} \Big(D_i x^2 - D_i^{(t)} x^{D_i^{(t)}} - (D_i - D_i^{(t)}) x^{D_i^{(t)}+1} \Big) \tag{3.8}$$

of the functions $G_D(x)$, $G_{D^{(t)}}(x)$, $H(x)$ and $\overline{H}(x)$, respectively. We calculate estimators $\hat{\alpha}$ and $\hat{\overline{\alpha}}$ of the fraction of the influenced population α and of good pioneers $\overline{\alpha}$ using Theorems 2.2 and 2.5 and the estimated functions $\hat{G}_D(x)$, $\hat{G}_{D^{(t)}}(x)$, $\hat{H}(x)$ and $\hat{\overline{H}}(x)$. (That is, we find numerically zeros $\hat{\xi}$ and $\hat{\overline{\xi}}$ of $\hat{H}(x)$ and $\hat{\overline{H}}(x)$, respectively, and use Theorems 2.2 and 2.5, with $\hat{G}_D(x)$ and $\hat{G}_{D^{(t)}}(x)$ replacing $G_D(x)$ and $G_{D^{(t)}}(x)$.)

Note that in the semi-analytic approach we do not need to know/construct the realization of the underlying model. This observation is a basis of a *campaign evaluation method* that we propose in Section 3.3. In fact, in reality one usually does not have the complete insight into the network structure and needs to rely on statistics collected from the initially contacted pioneers.

3.2.4.3 Analytic evaluation

Finally, for all models, except the "coupon-collector" one of Section 3.2.3, we calculate numerically the values of α and $\overline{\alpha}$ using the explicit forms of all the involved functions. (For the coupon-collector model we obtained the "true" values of α and $\overline{\alpha}$ from a sample of $(D_i, D_i^{(t)})$ of a larger size N.)

When comparing these analytic solutions to the simulation and semi-analytic estimates we see that in some cases $N = 1000$ is not big enough to match the theoretical values. One can easily consider larger samples, however we decided to stay with $N = 1000$ to show how the quality of the estimation varies over

different model assumptions. Also, $N = 1000$ seems to be near the lower range of the number of initial pioneers one needs to contact to produce a reasonable prognosis for the development of the campaign.

3.2.4.4 Case study

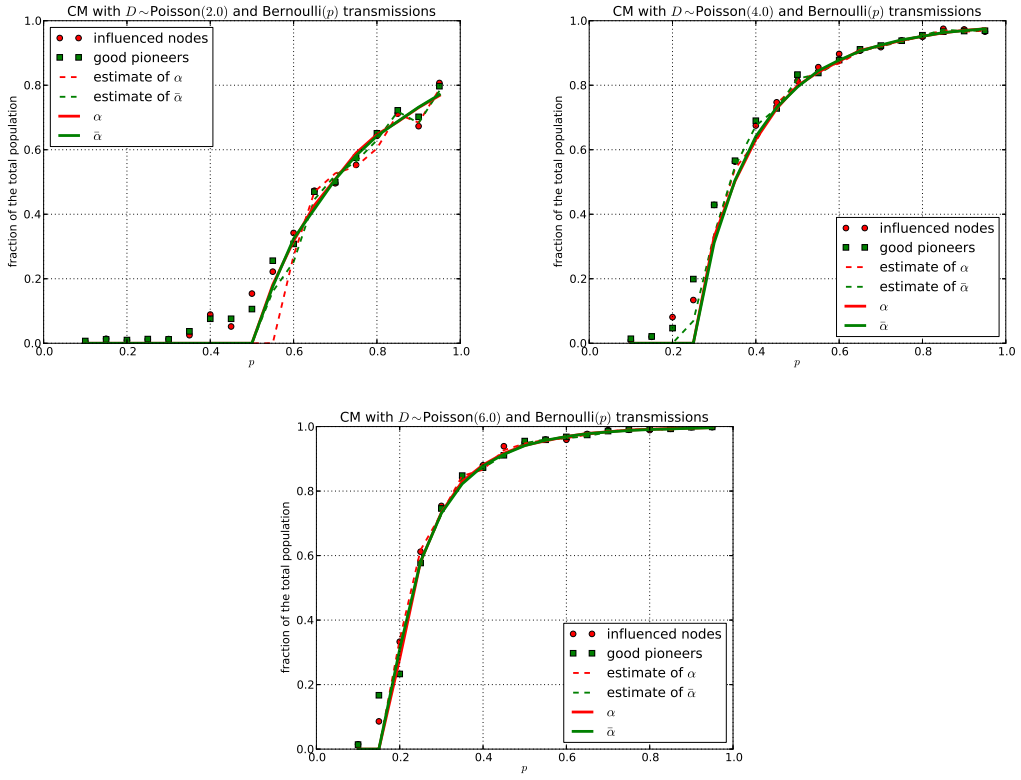

Figure 3.2: CM with Poisson degree of mean $\lambda = 2, 4, 6$ and Bernoulli transmissions with probability p. The set of good pioneers and the influenced population are of the same size. Their fraction is strictly positive for $p > 1/\lambda$.

Figures 3.2 and 3.3 present Bernoulli influence propagation on the CM with Poisson and Power-Law degree distribution of mean $\mathbb{E}[D] = 2, 4, 6$. Bernoulli transmissions imply the set of good pioneers and influenced population of the same size. The Power-Law degree with $\beta < 3$ leads to positive fraction of good pioneers and influenced component for all $p > 0$, while for the Poisson degree distribution one observes the phase transition at $p = 1/\lambda$. That is, the fractions of good pioneers and the influenced component are strictly positive if and only if $p > 1/\lambda$.

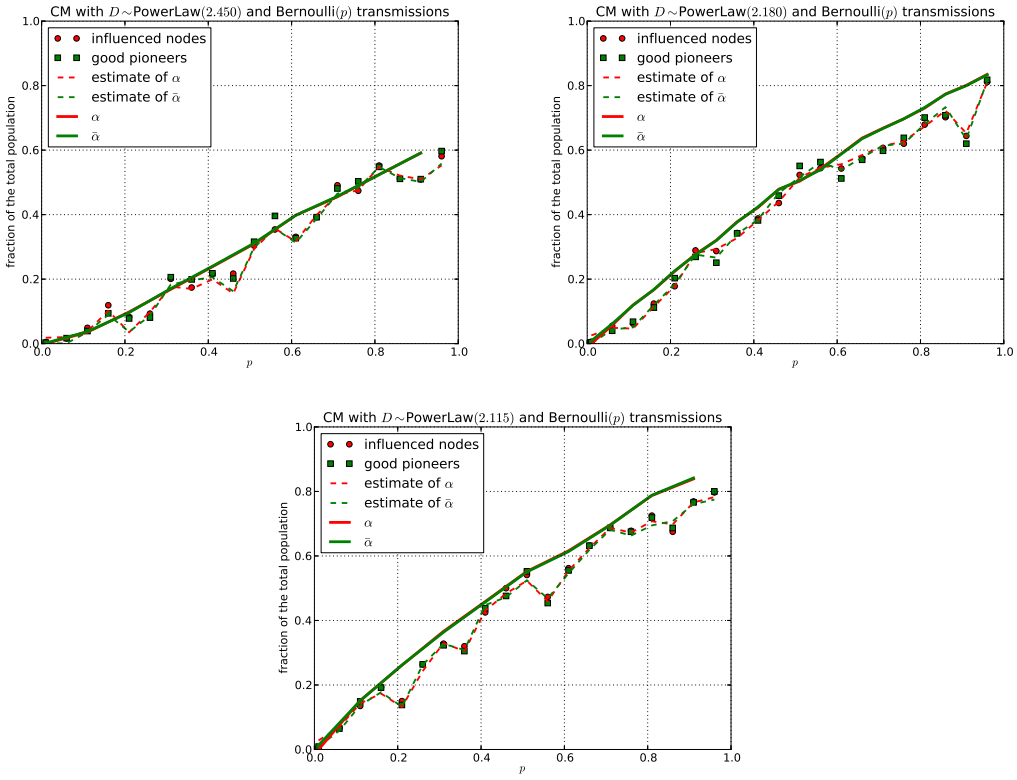

Figure 3.3: CM with Power-Law degree of parameter $\beta = 2.450, 2.180, 2.115$ (corresponding to $\mathbb{E}[D] \approx 2, 4, 6$ and Bernoulli transmissions with probability p. The set of good pioneers and the influenced population are of the same size. Their fraction is strictly positive for all $p > 0$ whenever $\beta \leq 3$.

Figure 3.4 shows again the model with Bernoulli transmissions on CM with Poisson and Power-Law degree distribution, this time however for $\mathbb{E}[D] \approx 1.35$ for which both models exhibit the phase transition in p. A general observation is that the Power-Law degree distribution gives smaller critical values of p for the existence of a positive fraction of influenced population and good pioneers, however for these the size of these sets increase with the transmission probability p more slowly in the Bernoulli model. Obviously the values of $\alpha = \overline{\alpha}$ at $p = 1$ correspond to the size of the biggest connected component of the underlying CM.

Figure 3.5 shows the node percolation (or "apathetic and enthusiastic users) on CM with Poisson and Power-Law degree distribution of mean $\mathbb{E}[D] \approx 2$. Note that the influenced components have the same

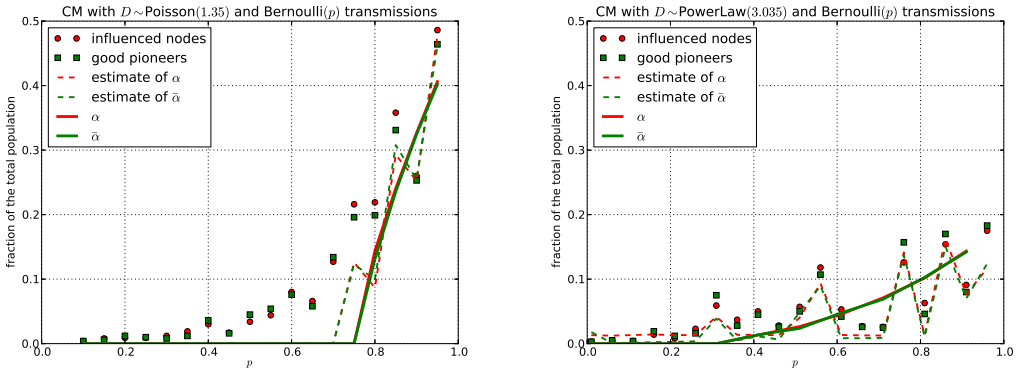

Figure 3.4: CM with Poisson and Power-Law degree of mean $\mathbb{E}[D] \approx 1.35$ ($\lambda = 1.35$ and $\beta = 3.035$) and Bernoulli transmissions. The set of good pioneers and the influenced population are of the same size for each model. One observes the phase transition in both models, at $p = 1/\lambda$ and $p = \zeta(\beta-1)/(\zeta(\beta-2)-\zeta(\beta-1)$, respectively.

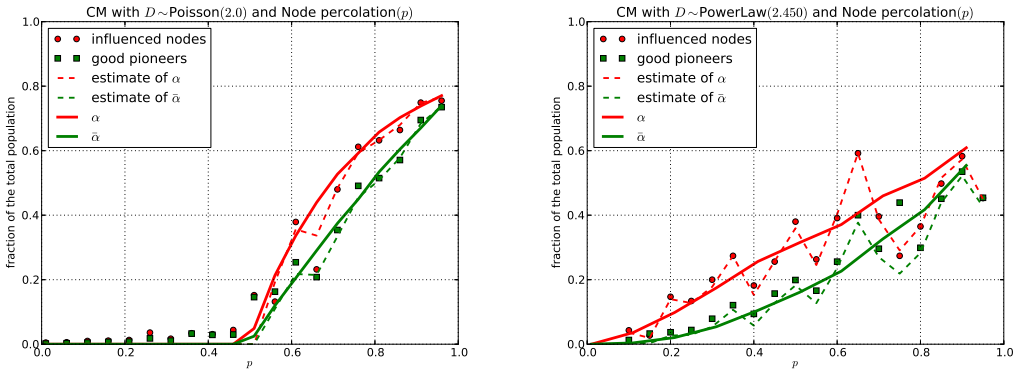

Figure 3.5: Node percolation ("apathetic and enthusiastic users") on CM with Poisson and Power-Law degree of mean $\mathbb{E}[D] \approx 2$ ($\lambda = 2$ and $\beta = 2.45$). The influenced component and the critical values for p are equal to these for the CM with Bernoulli transmissions. The set of good pioneers is smaller than the influenced population. We do not observe the phase transition for the Power-Law model since $\beta < 3$.

size as for Bernoulli transmissions, however good components are smaller. The critical values of p for the phase transition are also the same as for Bernoulli transitions. Note that estimation of the node percolation model is more difficult than the Bernoulli transmissions because of higher variance of the estimators.

Finally, Figure 3.6 shows that the coupon collector dynamics ("absentminded users") on CM produces

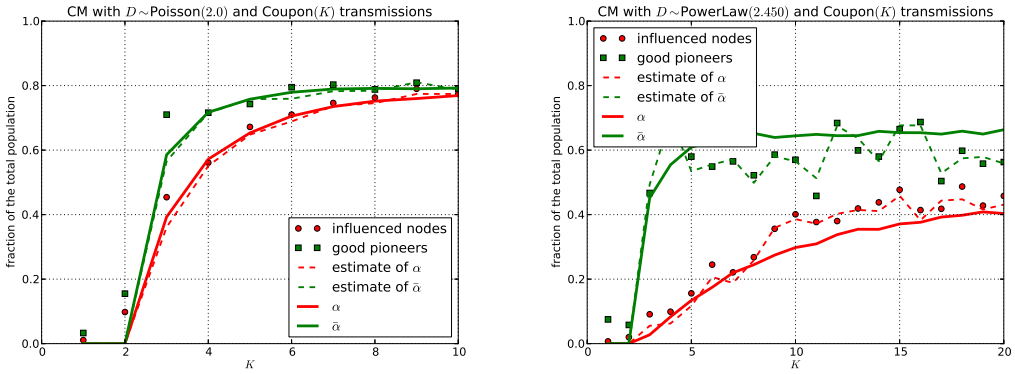

Figure 3.6: Coupon collector dynamics "absentminded users") on CM with Poisson and Power-Law degree of mean $\mathbb{E}[D] \approx 2$ ($\lambda = 2$ and $\beta = 2.45$). The set of good pioneers is bigger than the influenced population.

bigger sets of good pioneers than the influenced population.

3.3 Application to Viral Campaign Evaluation

What does the analysis presented up to now suggest in terms of strategy for a firm which is just about to start a new marketing campaign on an online social network without having any prior information about the network structure?

If the fraction of good pioneers in the network is non-negligible, the firm has a strictly positive probability of picking a good pioneer even when it picks a pioneer uniformly at random from the network. Now when is the fraction of good pioneers non-negligible? Since the firm has no prior information about the network structure and the campaign effectiveness, the best it can do is to collect information from its pioneers regarding the number of friends that they have (total degree) and the number of friends they influence in this campaign (transmitter degree), and then assume that the network is a uniform random network having the sampled total degree and transmitter degree distributions. The collected information, denote it by $(D_i, D_i^{(t)})$, $i = 1, \ldots, N$, allows to estimate various quantities relevant to the potential development of the ongoing campaign, as we did in 3.2.4.

More precisely the results presented in Chapter 2 suggest the following approach.

Network fragmentation The first and foremost question is whether the network is not too fragmented to allow for viral marketing. This is related to the condition (1.18) in Section 1.3.1. In order to answer this question one considers the following estimator of $\mathbb{E}[D^2 - 2D]$

$$\frac{1}{N}\sum_{i=1}^{N}\Big(D_i^2 - 2D_i\Big).$$

If the value of this estimator is not sharply larger than zero then the firm must assume that the network is too fragmented to allow for viral marketing. Natural confidence intervals can be considered in this context too. Evidently, the confidence increase as the firm picks more pioneers and collects more data.

Effectiveness of the campaign If one estimates that the network is not too fragmented, then the firm can evaluate the effectiveness of the ongoing campaign. It is related to condition (2.4). Again one considers the natural estimate of $\mathbb{E}[DD^{(t)} - D^{(t)} - D]$

$$\frac{1}{N}\sum_{i=1}^{N}\Big(D_iD_i^{(t)} - D_i - D_i^{(t)}\Big).$$

If the value of this estimator is sharply larger than zero then the firm can assume that there is a realistic chance of picking a good pioneer via random sampling and make the campaign go viral. Otherwise, the previous phase of the campaign can be considered as non-efficient.

Cost-benefit analysis If the firm deems the campaign to be effective, it can then, exactly as we did in 3.2.4.2, come up with the estimates of the relative fractions of good pioneers and population vulnerable to influnce, and do a cost-benefit analysis.

What we have described is an outline which can be used by the firms to come up with a rational methodology for making decisions in the context of viral marketing.

4

Isolated Vertices and the Longest Edge of the Minimum Spanning Tree of Weighted Configuration Model

In this chapter, we consider weighted random graphs, i.e., the random graphs whose edges are given random i.i.d. weights. This is a way to introduce geometry in otherwise non-geometric graphs. One could ask several questions for such "artificially geometrized" random graphs, for instance, how are they related to the classical geometric graphs, where the geometry is induced naturally by the euclidean distance. The goal can be to approximate the latter by the former, which are more amenable to explicit quantitative analysis, as we have seen in the previous chapters. Indeed, in this chapter, we observe one particular similarity between weighted configuration model and the Gilbert graph on Euclidean space (introduced by Gilbert in [34]), namely, that the longest edge of the minimum spanning tree scales with the graph size exactly as the longest edge of the Nearest Neighbour Graph. This has been observed for the Gilbert graph

by Penrose in [51].

Further motivation for this study comes from the the theory of epidemic spread - how isolated must an individual be in a connected network to survive a contagion spreading in the network?

From the perspective of random graph theory, this study is motivated by the observation that the vertices which are separated from the rest of the graph by a distance exceeding certain threshold play an important role in determining some global properties of the graph like diameter, flooding time etc., in spite of being statistically rare.

4.1 Introduction

Recall that the minimum spanning tree (abbreviated as MST) of a connected graph on n vertices is a connected subgraph in which the sum of all edge-lengths is minimum among all possible connected subgraphs. In this chapter, we are interested in the longest edge of the MST of configuration model, $G(n,(d_i)_1^n)$, which we denote by $\mathbf{M}_n$. Remark that, in contrast to the previous chapters, we work directly with the simple-graph version of the configuration model in this chapter, that is, the uniform random (simple) graph, $G(n,(d_i)_1^n)$ on a given degree sequence $(d_i)_1^n$. For ease of presentation, we will denote it by the abbreviated notation, G_n, in this chapter. However, we will continue to use the results prevously stated for the multi-graph version, $G^*(n,(d_i)_1^n)$, without repeating the arguments of Theorem 1.9.

Every edge (i,j) of G_n is independently given a random edge-length, denoted $Y(i,j)$, distributed according to random variable Y. We thus get a weighted random graph which we denote by $\widetilde{G}_n$. In the context of the epidemic spread in species described above, lower value of $Y(i,j)$ signifies that i and j are closely linked. For $x \geq 0$, let

$$\mathbb{P}(Y \leq x) = F(x)$$

and $\bar{F}(x) = 1 - F(x)$. Since MST is invariant to the precise distribution of edge-weights, we can assume without loss of generality that Y is exponentially distributed with parameter 1, i.e., $F(x) = e^{-x}$. The length of the longest edge of MST for general edge-distribution can be simply obtained from that for the exponential distribution by applying an appropriate function.

To make the notion of isolated (or weakly-linked) vertices precise, we introduce the following definition.

Definition 4.1 *Given $\alpha > 0$, we say that an edge (i, j) is α-long if*

$$Y(i, j) > \alpha, \tag{4.1}$$

and a vertex i in $\widetilde{G}_n$ is said to be α-far if all its connecting edges are α-long. That is, letting

$$M_n(i) := \min_{j:(i,j)\in e(\widetilde{G}_n)} Y(i, j),$$

then i is α-far if

$$M_n(i) > \alpha.$$

We recall that the *nearest neighbour graph* (abbreviated as NNG) of a graph is its subgraph where every vertex is connected by an edge only to its nearest neighbour.

We first study the asymptotic distribution of the number of α_n-far vertices in $\widetilde{G}_n$ when α_n is appropiately scaled as $n \to \infty$. α_n-far vertices are separated from the rest of the graph by the longest distance, so we will informally refer to them as isolated in this chapter, even though, strictly speaking, they may not be isolated, i.e. disconnected, from $\widetilde{G}_n$. Their asymptotic distribution leads to a weak law for the length of the longest edge of NNG of $\widetilde{G}_n$. We finally prove that for $\widetilde{G}_n$, the longest edge of its MST and that of its NNG coincide.

This strategy is similar to the one adopted in [51] to prove a weak law for the longest edge of the Gilbert graph in Euclidean setting. To find the distribution of number of α_n-far vertices, we also use the same analytical tool of *Stein-Chen* method as in the Euclidean setting, which we recall from Section 1.2.2 is also used to prove connectivity in Erdős-Rényi random graph ([44]). However, to prove that the longest edge of MST and NNG coincide, the analysis differs from that in [51] due to the difference in the nature of underlying graphs. We instead prove that once we remove all the α_n-long edges, then the disconnectivity in the subgraph can only be due to the isolated vertices, exactly as one does to prove connectivity in Erdős-Rényi graph. For this, we will use the results on percolation in configuration model proved by Fountoulakis in [29]) and recalled in Section 1.3.2 and the proof itself is based on the proof of connectivity of configuration model (Theorem 1.16) given by van der Hofstad in [54].

4.2 Results

We start by making an additional assumption on d_i, along with Condition 1.8. Let $d_{min} := \min\{d : \mathbb{P}(D = d) > 0\} \geq 3$, and assume that $d_i \geq d_{min}$ for every $i \in [n]$. Then, by Theorem 1.16, G_n is connected whp (again, see [54] for proof). We impose this connectivity constraint because for one thing, MST is not defined for disconnected graphs. However, it is possible to consider the MST of the giant component of G_n (if it exists) instead. The main reason is to keep the analysis simple.

As described earlier, we give i.i.d. weight to every edge of G_n to obtain $\widetilde{G}_n$. Now we state the main theorem proved in Section 4.3.

Theorem 4.2 *Suppose the sequence $(\alpha_n)_{n\in\mathbb{N}}$ is such that when $n \to \infty$,*

$$d_{min}\alpha_n - \log n \to \beta \tag{4.2}$$

for some fixed $\beta \in \mathbb{R}$ and let N_0 be the number of α_n-far vertices in $\widetilde{G}_n$. Then, N_0 converges in the variation distance topology to the Poisson random variable with parameter $p_{d_{min}} e^{-\beta}$ for almost all sequences $\{G_n\}_{n\geq 1}$.

Note that condition (4.2) is equivalent to

$$ne^{-d_{min}\alpha_n} \to e^{-\beta}. \tag{4.3}$$

Therefore, taking $a := 2e^{-\beta}$, there exists k such that for $n > k$,

$$ne^{-d_{min}\alpha_n} < a, \tag{4.4}$$

and

$$e^{-\alpha_n} < \left(\frac{a}{n}\right)^{\frac{1}{d_{min}}}. \tag{4.5}$$

Theorem 4.2 leads to the following result on the longest edge of the NNG of $\widetilde{G}_n$ which we define formally as $\mathbf{M}'_n := \max_{i\in v(\widetilde{G}_n)} M_n(i)$. We have

Corollary 4.3 *For all* $\beta \in \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{P}(d_{min}\mathbf{M}'_n - \log n \le \beta) = e^{-p_{d_{min}} e^{-\beta}}. \tag{4.6}$$

Proof. It follows from the above theorem simply by taking

$$\alpha_n = \frac{1}{d_{min}}[\beta + \log n] \tag{4.7}$$

and remarking that

$$\mathbb{P}(\mathbf{M}'_n \le \alpha_n) = \mathbb{P}(N_0 = 0) \to e^{-p_{d_{min}} e^{-\beta}} \tag{4.8}$$

The above theorem says that the probability that no α_n-long edge exists in NNG tends to $e^{-p_{d_{min}} e^{-\beta}}$.

Recall that NNG $\subset$ MST. The main theorem of the second part gives the comparison between the α_n-long edges of NNG and MST.

Theorem 4.4 *For any* $\beta \in \mathbb{R}$ *and* α_n *satisfying (4.2), every* α_n*-long edge of the MST of* $\widetilde{G}_n$ *is also present in the corresponding NNG whp. Hence, evidently, every such edge has an end at a leaf of the MST, that is, a vertex of degree* 1.

The above theorem leads to the following result on the longest edge of MST, which corresponds exactly to Corollary 4.3 for NNG. Recall that $\mathbf{M}_n$ denotes the longest edge of the MST. Then, we have

Corollary 4.5 *For all* $\beta \in \mathbb{R}$,

$$\mathbb{P}(d_{min}\mathbf{M}_n - \log n \le \beta) \to e^{-p_{d_{min}} e^{-\beta}}. \tag{4.9}$$

Proof. Again, we take

$$\alpha_n = \frac{1}{d_{min}}[\beta + \log n]. \tag{4.10}$$

Theorem 4.4 implies that when $n \to \infty$, the probability that no α_n-long edge exists in MST is greater than or equal to the probability that no α_n-long edge exists in NNG, that is,

$$\liminf_{n\to\infty} \mathbb{P}(d_{min}\mathbf{M}_n - \log n \le \beta) \ge \liminf_{n\to\infty} \mathbb{P}(d_{min}\mathbf{M}'_n - \log n \le \beta). \tag{4.11}$$

But as we remarked earlier, NNG $\subset$ MST, therefore, we have, for all n,

$$\mathbb{P}(d_{min}\mathbf{M}_n - \log n \leq \beta) \leq \mathbb{P}(d_{min}\mathbf{M}'_n - \log n \leq \beta). \tag{4.12}$$

Thus, from (4.6), (4.11) and (4.12), we have Corollary 4.5

Remark 3 In (4.9), if β is replaced by β_n such that $\beta_n n \to \infty$ and $\beta_n / \log n \to 0$ as $n \to \infty$, then we have

$$\frac{\mathbf{M}_n}{\log n} \xrightarrow{\text{p}} \frac{1}{d_{min}} \tag{4.13}$$

We recall from Section 1.3.4 that Theorem 1.17 (proven in [5]) states that, under the conditions on $(d_i)_1^n$ assumed in this chapter, we have

$$\frac{\text{diam}(G_n)}{\log n} \xrightarrow{\text{p}} \frac{1}{\nu - 1} + \frac{2}{d_{min}} \tag{4.14}$$

As we remarked earlier, the first term, $1/(\nu-1)$, is linked to the distance between two uniformly chosen vertices of G_n ([11]), while from (4.13), we can see that the second term, $2/d_{min}$ is closely linked to the length of the longest edge of the MST (or NNG). Indeed, one can conjecture that the end-points of the diameter are formed by the isolated vertices, say, iv_1 and iv_2, both of which are connected by an edge to *typical* vertices of the graph, say v_1 and v_2, respectively. Therefore,

$$\text{diam}(G_n) = \text{dist}(iv_1, v1) + \text{dist}(v_1, v_2) + \text{dist}(v_2, iv_2) \tag{4.15}$$

However, we do not prove this conjecture here.

4.3 Isolated vertices of weighted configuration model

In this section, we fix $\beta \in \mathbb{R}$ and α_n satisfying (4.2).

The intuition behind Theorem 4.2 is as follows. Given a particular realization of G_n, $I_i := \mathbb{1}(\text{vertex } i \text{ is far-out})$ is a Bernoulli random variable with parameter $e^{-d_i\alpha_n}$ and $N_0 = \sum_{i=1}^n I_i$. Taking $\mathbf{p} = [e^{-d_1\alpha_n} e^{-d_2\alpha_n} \dots e^{-d_n\alpha_n}]$ and ignoring the weak dependence among I_i's,

$$N_0 \overset{d}{\approx} Poisson - binomial(n, \mathbf{p}).$$

Also,

$$\mathbb{E}N_0 = \sum_{i=1}^{n} e^{-d_i\alpha_n} \approx n\mathbb{E}[e^{-D\alpha_n}] = ne^{-d_{min}\alpha_n}\mathbb{E}[e^{-\hat{D}\alpha_n}], \tag{4.16}$$

where $\hat{D} = D - d_{min}$. From (4.5), when $n \to \infty$, $e^{-\alpha_n} \to 0$, and therefore,

$$\mathbb{E}[e^{-\hat{D}\alpha_n}] \to \mathbb{P}(\hat{D} = 0) = \mathbb{P}(D = d_{min}) = p_{d_{min}}.$$

This, along with (4.4) and (4.16), implies that for large n,
$\mathbb{E}N_0 \approx p_{d_{min}} e^{-\beta}$ and hence,

$$N_0 \overset{d}{\approx} Poisson(p_{d_{min}} e^{-\beta}).$$

We make the above steps rigorous using the Stein-Chen method in a manner similar to that in [44] for the proof of connectivity in Erdős-Rényi graph, which we briefly described in Section 1.2.2. Recall that P_a denotes Poission r.v. with parameter a, $\mathscr{L}(\boldsymbol{V})$ denotes the law of random vector $\boldsymbol{V}$, and $d_{var}(.,.)$ denotes the *distance in total variation*.

Remark first that Condition 1.8 (ii) implies that there exists a constant A_2 such that for all n,

$$\sum_{i=1}^{n} (d_i)^2 < A_2 n, \tag{4.17}$$

and therefore, letting $d_{max} := \max_i d_i$,

$$d_{max} < \sqrt{A_2}\sqrt{n}. \tag{4.18}$$

Also (1.16) implies that there exists constants A_1 and A_1' such that for all n,

$$A_1' n < \sum_{i=1}^{n} d_i < A_1 n. \tag{4.19}$$

We start by proving the following lemma.

Lemma 4.6 *Given the Condition 1.8 and $(\alpha_n)_{n\in\mathbb{N}}$ satisfying (4.2), we have that when $n \to \infty$,*

$$\sum_{i=1}^{n} e^{-2d_i\alpha_n} \to 0, \tag{4.20}$$

and

$$\sum_{i=1}^{n} d_i e^{-(d_i+1)\alpha_n} \to 0. \tag{4.21}$$

Proof. We have that since $d_i \geq d_{min}$,

$$\sum_{i=1}^{n} e^{-2d_i\alpha_n} \leq ne^{-2d_{min}\alpha_n} = e^{-d_{min}\alpha_n} ne^{-d_{min}\alpha_n} < \frac{a^2}{n},$$

where the last inequality follows from (4.4).

Similarly,

$$\sum_{i=1}^{n} d_i e^{-(d_i+1)\alpha_n} \leq e^{-\alpha_n} e^{-d_{min}\alpha_n} \sum_{i=1}^{n} d_i < A_1 ne^{-d_{min}\alpha_n} e^{-\alpha_n},$$

where the last inequality follows from (4.19). Therefore, by (4.4) and (4.5), we have

$$\sum_{i=1}^{n} d_i e^{-(d_i+1)\alpha_n} < A_1 a \left(\frac{a}{n}\right)^{\frac{1}{d_{min}}}.$$

Proof of Theorem 4.2.

Let $I_i := \mathbb{1}(\text{vertex } i \text{ is } \alpha_n\text{-far in } \widetilde{G}_n)$. Given G_n, I_i is a Bernoulli random variable with parameter $\pi_i := e^{-d_i\alpha_n}$. Therefore, $N_0 = \sum_{i=1}^{n} I_i$ and let $S := \mathbb{E}[N_0] = \sum_{i=1}^{n} e^{-d_i\alpha_n}$.

Further, given G_n, let

$$J_{ij} := \mathbb{1}(\min_{k:k\sim j, k\neq i} Y(k,j) > \alpha_n).$$

From the independence of edge weights, it is clear that (1.8) is satisfied, given G_n. Now,

$$|I_j - J_{ij}| = J_{ij} - I_j = \mathbb{1}(\exists \text{ edge } (i,j) \text{ in } G_n, Y(i,j) \leq \alpha_n \text{ and } \min_{k:k\sim j, k\neq i} Y(k,j) > \alpha_n).$$

Therefore,

$$\mathbb{E}[|I_j - J_{ij}| \big| G_n] = \mathbb{1}(\exists\ (i,j) \text{ in } G_n)(1-e^{-\alpha_n})e^{-(d_j-1)\alpha_n}.$$

We are now in a situation to apply Theorem 1.4. Plugging the above into (1.9), we get

$$\begin{aligned} d_{var}(\mathscr{L}(N_0\big|G_n), P_S) &\le 2\frac{1-e^{-S}}{S}\left[\sum_{i=1}^n \pi_i^2 + \sum_{i=1}^n \pi_i \sum_{j\neq i} \mathbb{1}(\exists\ (i,j))(1-e^{-\alpha_n})e^{-(d_j-1)\alpha_n}\right] \\ &\le 2\min(1,\frac{1}{S})\left[\sum_{i=1}^n e^{-2d_i\alpha_n} + \sum_{i=1}^n e^{-d_i\alpha_n}\sum_{j\sim i}(1-e^{-\alpha_n})e^{-(d_j-1)\alpha_n}\right] \\ &\le 2\min(1,\frac{1}{S})\left[\sum_{i=1}^n e^{-2d_i\alpha_n} + \sum_{i=1}^n e^{-(d_i+1)\alpha_n} d_i\right]. \end{aligned}$$

Thus, by Lemma 4.6, we have when $n \to \infty$,

$$d_{var}(\mathscr{L}(N_0\big|G_n), P_S) \to 0. \tag{4.22}$$

Now, let u_k denote the number of vertices with degree k, that is, $u_k = |\{i : d_i = k\}|$. Then, we have

$$\begin{aligned} S = \sum_{i=1}^n e^{-d_i\alpha_n} &= \sum_{k=d_{min}}^{\infty} e^{-k\alpha_n} u_k \\ &\le n e^{-d_{min}\alpha_n}\frac{u_{d_{min}}}{n} + e^{-\alpha_n} n e^{-d_{min}\alpha_n}\frac{\sum_{k\ge 0} u_k}{n}. \end{aligned}$$

By Condition 1.8 (i), we have that

$$\frac{u_{d_{min}}}{n} \to p_{d_{min}},$$

and by (1.16), we have that

$$\frac{\sum_{k\ge 0} u_k}{n} = \frac{\sum_{i=1}^n d_i}{n} \to \mathbb{E}[D],$$

which, along with (4.3) and (4.5), show that $S \to p_{d_{min}} e^{-\beta}$. Thus, by Lemma 1.5,

$$d_{var}(P_{p_{d_{min}}e^{-\beta}}, P_S) \to 0. \tag{4.23}$$

Moreover by triangle inquality,

$$d_{var}(\mathscr{L}(N_0\big|G_n), P_{p_{d_{min}}e^{-\beta}}) \le d_{var}(\mathscr{L}(N_0\big|G_n), P_S) + d_{var}(P_{p_{d_{min}}e^{-\beta}}, P_S).$$

Therefore, from (4.22) and (4.23), we finally have that,

$$d_{var}(\mathscr{L}(N_0\big|G_n), P_{p_{d_{min}}e^{-\beta}}) \to 0.$$

This completes the proof.

4.4 Longest edge of MST

Again, we fix $\beta \in \mathbb{R}$ and α_n satisfying (4.2).

In this section, we turn our attention to the MST of $\widetilde{G}_n$. Theorem 4.4 is equivalent to saying that every α_n-long edge of the MST has an α_n-far vertex as one end and no other edge of $\widetilde{G}_n$ incident to this vertex lies in the MST. To prove this, let $\widetilde{G}_n^\alpha$ denote the graph obtained by keeping only those edges of $\widetilde{G}_n$ whose length is less than or equal to α_n. In other words, $\widetilde{G}_n^\alpha$ is obtained from $\widetilde{G}_n$ after bond percolation with probability $\pi_n^\alpha = 1 - e^{-\alpha_n}$. It is clear that α_n-far vertices are disconnected from $\widetilde{G}_n^\alpha$. We would like to prove that the rest of the vertices form one giant component of $\widetilde{G}_n^\alpha$, or equivalently, there does not exist in $\widetilde{G}_n^\alpha$, a connected component of size strictly greater than 1, other than its unique giant component. Clearly, by the preceding discussion, this proves Theorem 4.4.

Note that, in what follows, $\mathbf{d} = (d_i)_1^n$ will continue to denote the degree sequence of $\widetilde{G}_n$, where d_i is the degree of vertex i, while the (random) degree sequence obtained after percolation in $\widetilde{G}_n^\alpha$ will be denoted by $\mathbf{D}^\alpha = (D_i^\alpha)_i^n$, where D_i^α is the degree of vertex i.

Now, by (4.5), $\pi_n^\alpha \to 1$ when $n \to \infty$. Moreover, since $d_{min} \geq 3$, we have that

$$\frac{\mathbb{E}D}{\mathbb{E}D(D-1)} \leq \frac{1}{2} \tag{4.24}$$

Therefore, trivially, we have that

$$\liminf_{n\to\infty} \pi_n^\alpha > \frac{\mathbb{E}D}{\mathbb{E}D(D-1)}. \tag{4.25}$$

Thus, denoting the second largest component of $\widetilde{G}_n^\alpha$ by C_2^α, we have by Theorem 1.13 that

$$|C_2^\alpha| < \nu_2 \log n, \tag{4.26}$$

for some constant ν_2.

By Theorem 4.2, we already have an estimate of isolated (degree 0) vertices in $\widetilde{G}_n^{\alpha}$. Next, we will estimate the number of vertices with degree 1 and degree 2. Let N_1 and N_2 be the number of vertices with degree 1 and 2, respectively, and N_3 be the number of vertices with degree strictly greater than 2 in $\widetilde{G}_n^{\alpha}$. Recall that the result would follow directly from Theorem 1.16, if N_1 and N_2 were equal to 0. We prove, however, that this is not the case.

Lemma 4.7 *Then, we have that whp, given G_n,*

(i) $N_1 < \rho_1 n^{1/3}$,

(ii) $N_2 < \rho_2 n^{2/3}$,

for some constants ρ_1 and ρ_2.

Proof. (i) Let

$$I_i' := \mathbb{1}(\text{vertex } i \text{ has degree 1 in } \widetilde{G}_n^{\alpha}). \tag{4.27}$$

Note that given G_n, I_i' is a Bernoulli random variable with parameter π_i' with

$$\pi_i' \leq d_i e^{-(d_i-1)\alpha_n}. \tag{4.28}$$

Now, since $N_1 = \sum_{i=1}^n I_i'$, we have that given G_n,

$$\begin{aligned}\mathbb{E}[N_1] = \sum_{i=1}^n \pi_i' &\leq \sum_{i\in[n]:d_i>d_{min}+2} d_i e^{-(d_{min}+1)\alpha_n} + \sum_{i\in[n]:d_i\leq d_{min}+2} (d_{min}+2)e^{-(d_i-1)\alpha_n} \\ &\leq \sum_{i=1}^n d_i e^{-(d_{min}+1)\alpha_n} + \sum_{i=1}^n (d_{min}+2)e^{-(d_i-1)\alpha_n}.\end{aligned}$$

Now, by (4.17), we have

$$\sum_{i=1}^n d_i e^{-(d_{min}+1)\alpha_n} < A_1 n e^{-d_{min}\alpha_n} e^{-\alpha_n},$$

and since $d_i \geq d_{min}$,

$$\sum_{i=1}^{n}(d_{min}+2)e^{-(d_i-1)\alpha_n} \leq n(d_{min}+2)e^{-(d_{min}-1)\alpha_n}.$$

Therefore, we have that given G_n,

$$\mathbb{E}[N_1] < A_1 n e^{-d_{min}\alpha_n}e^{-\alpha_n} + n(d_{min}+2)e^{-(d_{min}-1)\alpha_n},$$

which, by (4.4) and (4.5), implies that for $n > k$, given G_n,

$$\mathbb{E}[N_1] < A_1 a\left(\frac{a}{n}\right)^{\frac{1}{d_{min}}} + n(d_{min}+2)\left(\frac{a}{n}\right)^{\frac{d_{min}-1}{d_{min}}}.$$

Now, since $d_{min} \geq 3$, taking $c_1 := \left(A_1 a^{\frac{1}{d_{min}}} + (d_{min}+2)a^{-\frac{1}{d_{min}}}\right)a$, we have that for $n > k$, given G_n,

$$\mathbb{E}[N_1] \leq c_1 n^{1/3}. \tag{4.29}$$

Now, for $i \neq j$, given G_n,

$$\begin{aligned}\operatorname{cov}(I_i', I_j') &= \mathbb{E}[I_i' I_j'] - \mathbb{E}[I_i']\mathbb{E}[I_j'] \\ &= \left(\mathbb{E}[I_j' | I_i' = 1] - \mathbb{E}[I_j']\right)\mathbb{E}[I_i'].\end{aligned}$$

That is, given G_n,

$$\begin{aligned}\operatorname{cov}(I_i', I_j') = &\big(\mathbb{P}(j \text{ has degree 1 in } \widetilde{G}_n^{\alpha} \,\big|\, i \text{ has degree 1 in } \widetilde{G}_n^{\alpha}) \\ &- \mathbb{P}(j \text{ has degree 1 in } \widetilde{G}_n^{\alpha})\big)\mathbb{P}(i \text{ has degree 1 in } \widetilde{G}_n^{\alpha}).\end{aligned}$$

Since $\mathbb{P}(i \text{ has degree 1 in } \widetilde{G}_n^{\alpha}) = \pi_i'$, we have that given G_n,

$$\operatorname{cov}(I_i', I_j') \leq \mathbb{1}(\exists\ (i,j) \text{ in } G_n)\pi_i'. \tag{4.30}$$

Therefore, given G_n,

$$\begin{aligned}
\operatorname{var}(N_1) &= \sum_{i=1}^{n} \pi_i'(1-\pi_i') + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \operatorname{cov}(I_i', I_j') \\
&\leq \sum_{i=1}^{n} \left(1 - \pi_i' + \sum_{\substack{j=1 \\ j\neq i}}^{n} \mathbb{1}(\exists\ (i,j) \text{ in } G_n) \right) \pi_i' \\
&\leq \sum_{i=1}^{n} (1+d_i)\, \pi_i' \\
&\leq \sum_{i=1}^{n} 2(d_i)^2 e^{-(d_i-1)\alpha_n} \\
&\leq \sum_{i\in[n]:d_i>d_{min}+2} 2(d_i)^2 e^{-(d_{min}+1)\alpha_n} + \sum_{i\in[n]:d_i\leq d_{min}+2} 2(d_{min}+2)^2 e^{-(d_i-1)\alpha_n} \\
&\leq \sum_{i=1}^{n} 2(d_i)^2 e^{-(d_{min}+1)\alpha_n} + \sum_{i=1}^{n} 2(d_{min}+2)^2 e^{-(d_i-1)\alpha_n}.
\end{aligned}$$

Now, as before, by (4.17), we have

$$\sum_{i=1}^{n} 2(d_i)^2 e^{-(d_{min}+1)\alpha_n} < 2A_2 n e^{-d_{min}\alpha_n} e^{-\alpha_n}$$

and since $d_i \geq d_{min}$,

$$\sum_{i=1}^{n} 2(d_{min}+2)^2 e^{-(d_i-1)\alpha_n} \leq 2n(d_{min}+2)^2 e^{-(d_{min}-1)\alpha_n}.$$

Therefore, as before, by (4.4) and (4.5), we have that given G_n,

$$\operatorname{var}[N_1] < 2A_2 a \left(\frac{a}{n}\right)^{\frac{1}{d_{min}}} + 2(d_{min}+2)^2 n \left(\frac{a}{n}\right)^{\frac{d_{min}-1}{d_{min}}}.$$

Taking $c_2 := 2a\left(A_2 a^{\frac{1}{d_{min}}} + (d_{min}+2)^2 a^{-\frac{1}{d_{min}}}\right)$, we have that for $n > k$, given G_n,

$$\text{var}[N_1] < c_2 n^{1/3}. \tag{4.31}$$

Finally, by Chebyshev's inequality, we have that for $n > k$, given G_n,

$$\mathbb{P}\left\{|N_1 - \mathbb{E}N_1| > \sqrt{c_2} n^{1/6} \log n\right\} \leq \left(\frac{1}{\log n}\right)^2. \tag{4.32}$$

This implies that whp, given G_n,

$$N_1 < \mathbb{E}N_1 + \sqrt{c_2} n^{1/6} \log n.$$

Thus, by (4.29), we have that whp, given G_n,

$$\begin{aligned} N_1 &< c_1 n^{1/3} + \sqrt{c_2} n^{1/6} \log n \\ &< \rho_1 n^{1/3}, \end{aligned}$$

where $\rho_1 := c_1 + \sqrt{c_2}$.

(ii) Let

$$I_i'' := \mathbb{1}(\text{vertex } i \text{ has degree 2 in } \widetilde{G}_n^{\alpha}). \tag{4.33}$$

Note that given G_n, I_i'' is a Bernoulli random variable with parameter π_i'' with

$$\pi_i'' \leq \binom{d_i}{2} e^{-(d_i-2)\alpha_n}. \tag{4.34}$$

As in the proof of (i), since $N_2 = \sum_{i=1}^n I_i''$, we have

$$\mathbb{E}[N_2] = \sum_{i=1}^n \pi_i'' \leq \sum_{i=1}^n (d_i)^2 e^{-(d_{min}+1)\alpha_n} + \sum_{i=1}^n (d_{min}+3)^2 e^{-(d_i-2)\alpha_n}.$$

Now, by (4.17), we have

$$\sum_{i=1}^{n}(d_i)^2 e^{-(d_{min}+1)\alpha_n} < A_2 n e^{-d_{min}\alpha_n} e^{-\alpha_n},$$

and since $d_i \geq d_{min}$,

$$\sum_{i=1}^{n}(d_{min}+3)^2 e^{-(d_i-2)\alpha_n} \leq n(d_{min}+3)^2 e^{-(d_{min}-2)\alpha_n}.$$

Therefore, we have that given G_n,

$$\mathbb{E}[N_2] < A_2 n e^{-d_{min}\alpha_n} e^{-\alpha_n} + n(d_{min}+3)^2 e^{-(d_{min}-2)\alpha_n},$$

which, by (4.4) and (4.5), implies that for $n > k$, given G_n,

$$\mathbb{E}[N_2] < A_2 a \left(\frac{a}{n}\right)^{\frac{1}{d_{min}}} + n(d_{min}+3)^2 \left(\frac{a}{n}\right)^{\frac{d_{min}-2}{d_{min}}}.$$

Now, since $d_{min} \geq 3$, taking $b_1 := a\left(A_2 a^{\frac{1}{d_{min}}} + (d_{min}+3)^2 a^{-\frac{2}{d_{min}}}\right)$, we have that for $n > k$, given G_n,

$$\mathbb{E}[N_2] \leq b_1 n^{2/3}. \tag{4.35}$$

Again, as in the proof of (i), we have that given G_n,

$$\mathrm{cov}(I_i'', I_j'') \leq \mathbb{1}(\exists\ (i,j) \text{ in } G_n)\pi_i''. \tag{4.36}$$

Therefore, given G_n,

$$\begin{aligned}
\operatorname{var}(N_2) &= \sum_{i=1}^{n} \pi_i''(1-\pi_i'') + \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} \operatorname{cov}(I_i'', I_j'') \\
&\leq \sum_{i=1}^{n} \left(1 - \pi_i'' + \sum_{\substack{j=1\\ j\neq i}}^{n} \mathbb{1}(\exists\ (i,j) \text{ in } G_n) \right) \pi_i'' \\
&\leq \sum_{i=1}^{n} (1+d_i)\, \pi_i'' \\
&\leq \sum_{i=1}^{n} (d_i)^3 e^{-(d_i-2)\alpha_n} \\
&\leq \sum_{i=1}^{n} (d_i)^3 e^{-(d_{min}+1)\alpha_n} + \sum_{i=1}^{n} (d_{min}+3)^3 e^{-(d_i-2)\alpha_n}.
\end{aligned}$$

Now, as before, by (4.17) and bounding one of d_i by d_{max} and then using (4.18), we have

$$\begin{aligned}
\sum_{i=1}^{n} (d_i)^3 e^{-(d_{min}+1)\alpha_n} &< d_{max} A_2 n e^{-d_{min}\alpha_n} e^{-\alpha_n} \\
&< A_2 \sqrt{A_2}\, n^{\frac{1}{2}} n e^{-d_{min}\alpha_n} e^{-\alpha_n},
\end{aligned}$$

and since $d_i \geq d_{min}$,

$$\sum_{i=1}^{n} (d_{min}+3)^3 e^{-(d_i-2)\alpha_n} \leq n (d_{min}+3)^3 e^{-(d_{min}-2)\alpha_n}.$$

Therefore, as before, by (4.4) and (4.5), we have that given G_n,

$$\operatorname{var}[N_2] < A_2 \sqrt{A_2}\, a \left(\frac{a}{n}\right)^{\frac{1}{d_{min}}} n^{\frac{1}{2}} + n (d_{min}+3)^3 \left(\frac{a}{n}\right)^{\frac{d_{min}-2}{d_{min}}}.$$

Taking $b_2 := a\left(A_2\sqrt{A_2}a^{\frac{1}{d_{min}}} + (d_{min}+3)^3 a^{-\frac{2}{d_{min}}}\right)$, we have that for $n > k$, given G_n,

$$\mathrm{var}[N_2] < b_2 n^{2/3}. \tag{4.37}$$

Finally, by Chebyshev's inequality, we have that for $n > k$, given G_n,

$$\mathbb{P}\left\{|N_2 - \mathbb{E}N_2| > \sqrt{b_2}n^{1/3}\log n\right\} \leq \left(\frac{1}{\log n}\right)^2. \tag{4.38}$$

This implies that whp, given G_n,

$$N_2 < \mathbb{E}N_2 + \sqrt{b_2}n^{1/3}\log n.$$

Thus, by (4.35), we have that whp, given G_n,

$$\begin{aligned} N_2 &< b_1 n^{2/3} + \sqrt{b_2}n^{1/3}\log n \\ &< \rho_2 n^{2/3}, \end{aligned}$$

where $\rho_2 := b_1 + \sqrt{b_2}$.

Now, we give the following key result of this section, whose proof is inspired by the proof of Theorem 1.16 in [54].

Proposition 4.8 *For any $\beta \in \mathbb{R}$ and α_n satisfying (4.2), whp, given G_n, there are no connected components of size strictly greater than 1 in $\widetilde{G}_n^{\alpha}$, except the biggest component.*

Proof. We recall from Section 1.3 that a *configuration* denotes a pairing of all the half-edges. We also recall that the total number of configurations for a degree sequence, $(d_i)_1^n$, is given by $(l_n - 1)!!$, where $l_n = \sum_{i=1}^n d_i$ and !! is a symbol for double factorial. Therefore, the probability of any one configuration equals $1/(l_n - 1)!!$.

By (4.26), we need only prove that whp, given G_n, there are no connected components of size strictly greater than 1 and less than $\nu_2 \log n$.

Now, on the event that $\widetilde{G}_n^{\alpha}$ has a component of size between 2 and $\nu_2 \log n$, there exists a set of indices $\mathscr{I} \subset [\![n]\!]$ with $|\mathscr{I}| \leq \lfloor \nu_2 \log n \rfloor$ such that all half edges incident to vertices in $\mathscr{I}$ are only paired to other

half-edges incident to other vertices in $\mathscr{I}$. For $\mathscr{I} \subset [\![n]\!]$, let

$$l_n(\mathscr{I}) = \sum_{i \in \mathscr{I}} d_i. \tag{4.39}$$

Clearly, in order for the half-edges incident to vertices in $\mathscr{I}$ to be paired only to other half-edges incident to vertices in $\mathscr{I}$, $l_n(\mathscr{I})$ needs to be even, and the number of configurations for which this happens is bounded above by $(l_n(\mathscr{I})-1)!!((l_n(\mathscr{I}))^c-1)!!$.

Therefore, letting

$$p_n^{dc} := \mathbb{P}\big(\widetilde{G}_n^{\alpha} \text{ has a component of size between 2 and } \nu_2 \log n\big), \tag{4.40}$$

we have that

$$\begin{aligned} p_n^{dc} &\leq \sum_{\substack{\mathscr{I} \subset [\![n]\!] \\ |\mathscr{I}| \leq \nu_2 \log n}} \frac{(l_n(\mathscr{I})-1)!!((l_n(\mathscr{I}))^c-1)!!}{(l_n-1)!!} \\ &= \sum_{\substack{\mathscr{I} \subset [\![n]\!] \\ |\mathscr{I}| \leq \nu_2 \log n}} \prod_{j=1}^{l_n(\mathscr{I})/2} \frac{l_n(\mathscr{I})-2j+1}{l_n-2j+1}. \end{aligned}$$

Evidently, a component of size between 2 and $\nu_2 \log n$ does not have degree 0 vertices. Let $\mathscr{I}$ be a component with μ_1 degree 1 vertices, μ_2 degree 2 vertices and μ_3 vertices of degree strictly greater than 2. Therefore, we have

$$|\mathscr{I}| = \mu_1 + \mu_2 + \mu_3 \leq \nu_2 \log n, \tag{4.41}$$

and,

$$l_n(\mathscr{I}) \geq \mu_1 + 2\mu_2 + 3\mu_3. \tag{4.42}$$

Moreover, by (4.18), we have that

$$l_n(\mathscr{I}) \leq d_{max}\, \nu_2 \log n < \sqrt{A_2}\, \nu_2 \sqrt{n} \log n. \tag{4.43}$$

In particular, since $l_n > A_1' n$ by (4.19), there exists k_1, such that for $n > k_1$, we have that

$$l_n(\mathscr{I}) < \frac{l_n}{2} \tag{4.44}$$

Now, let

$$f(x) := \prod_{j=1}^{x} \frac{2x-2j+1}{l_n-2j+1}. \tag{4.45}$$

This can be rewritten as

$$f(x) = \frac{\prod_{j=1}^{x}(2x-2j+1)}{\prod_{j=1}^{x}(l_n-2j+1)} = \frac{\prod_{i=0}^{x-1}(2i+1)}{\prod_{k=0}^{x-1}(l_n-2k-1)} = \prod_{i=0}^{x-1} \frac{2i+1}{l_n-2i-1} \tag{4.46}$$

where $i = x-j$ and $k = j-1$ in the second equality. Since, for $x \le \frac{l_n}{4}$,

$$\frac{f(x+1)}{f(x)} = \frac{2x+1}{l_n-2x-1} \le 1, \tag{4.47}$$

we have that for $n > k_1$ and $x \le l_n(\mathscr{I})/2 \le l_n/4$, $x \mapsto f(x)$ is decreasing. Moreover, for $x \le (\mu_1+2\mu_2+3\mu_3)/2 \le \frac{3}{2}\nu_2 \log n$, there exists k_2 such that for $n > k_2$, we have that

$$\frac{2x+1}{l_n-2x+1} < \frac{4\nu_2 \log n}{A_1' n} < n^{-4/5}, \tag{4.48}$$

using (4.19).

Therefore, for $x \le \frac{3}{2}\nu_2 \log n$ and $n > k_2$,

$$f(x) \le n^{-4x/5}. \tag{4.49}$$

Therefore, for $n > k_3 := \max\{k_1, k_2\}$, we have that

$$f\left(\frac{l_n(\mathscr{I})}{2}\right) \le f\left(\frac{\mu_1+2\mu_2+3\mu_3}{2}\right) \le n^{-\frac{4}{10}(\mu_1+2\mu_2+3\mu_3)} = n^{-\frac{2(\mu_1+2\mu_2+3\mu_3)}{5}}. \tag{4.50}$$

Now, taking $\rho = \max(\rho_1, \rho_2)$, we have from Lemma 4.7 that given G_n, whp,

$$\binom{N_1}{\mu_1}\binom{N_2}{\mu_2}\binom{N_3}{\mu_3} \leq \rho^2 n^{\frac{\mu_1+2\mu_2+3\mu_3}{3}}. \tag{4.51}$$

Therefore, we have that for $n > k_3$, given G_n,

$$\begin{aligned}
p_n^{dc} &\leq \sum_{m=2}^{\nu_2 \log n} \sum_{\substack{\mu_1,\mu_2,\mu_3 \\ \mu_1+\mu_2+\mu_3=m}} \binom{N_1}{\mu_1}\binom{N_2}{\mu_2}\binom{N_3}{\mu_3} f\left(\frac{l_n(\mathscr{F})}{2}\right) \\
&\leq \sum_{m=2}^{\nu_2 \log n} \sum_{\substack{\mu_1,\mu_2,\mu_3 \\ \mu_1+\mu_2+\mu_3=m}} \rho^2 n^{\frac{\mu_1+2\mu_2+3\mu_3}{3}} n^{-\frac{2(\mu_1+2\mu_2+3\mu_3)}{5}} + \mathbb{P}\left(N_1 \geq \rho_1 n^{1/3} \text{ or } N_2 \geq \rho_2 n^{2/3}\right) \\
&\leq \sum_{m=2}^{\nu_2 \log n} \sum_{\substack{\mu_1,\mu_2,\mu_3 \\ \mu_1+\mu_2+\mu_3=m}} \rho^2 n^{-\frac{\mu_1+2\mu_2+3\mu_3}{15}} + o(1) \\
&\leq \sum_{m=2}^{\nu_2 \log n} \sum_{\substack{\mu_1,\mu_2,\mu_3 \\ \mu_1+\mu_2+\mu_3=m}} \rho^2 n^{-\frac{m}{15}} + o(1) \\
&\leq \sum_{m=2}^{\nu_2 \log n} \rho^2 (\nu_2 \log n)^3 n^{-\frac{m}{15}} + o(1).
\end{aligned}$$

Now, there exists k_4 such that for $n > k_4$, $\rho^2(\nu_2 \log n)^3 n^{-\frac{m}{15}} \leq n^{-\frac{m}{16}}$. Thus, for $n > \max(k_4, k_3)$, we finally have that given G_n,

$$\begin{aligned}
p_n^{dc} &\leq \sum_{m=2}^{\nu_2 \log n} n^{-\frac{m}{16}} + o(1) \\
&\leq \frac{n^{-\frac{1}{8}}}{1 - n^{-\frac{1}{16}}} + o(1).
\end{aligned}$$

This proves Proposition 4.8

Proof of Theorem 4.4. As discussed earlier, Theorem 4.4 follows from Proposition 4.8.

5

Future Work: Convex comparison of Random Graphs

The random structures that are amenable to theoretical analysis (including those described in this thesis) tend to be relatively simplistic, and don't always capture the essence of the real world phenomenon. Therefore, it is natural to define some kind of stochastic ordering, that would allow us to compare the *essential* properties of the realistic models to those of their simpler, fully-understood counterparts. Many random models are parametrized by the *size* of the model, and the essential properties of the model are the asymptotic ones as the *size* of the graph tends to infinity. So the theory of *local weak convergence* introduced in Section 1.4 provides a natural setting to investigate any stochastic order on random graphs.

The orders that we will focus on will be *convex*-like orders which compare two random variables with the same mean according to how *spread-out* their distributions are. This is motivated by the successful application of *directionally-convex* ordering in the context of point processes by Błaszczyszyn and Yogeshwaran in [14], [12]. We investigate these ideas in the context of branching process and configuration

model introduced in Section 1.1 and Section 1.3, respectively.

5.1 Convex Comparison of Random Graphs

5.1.1 Convex Order on Galton-Watson Tree and Implications

Since many important random graph models approach a random tree in the sense of local weak convergence, we start with an interesting observation on the Galton Watson branching process introduced in Section 1.1.

The Theorem 1.1 relates the p_{ext} to the mean of the offspring distribution, but it is natural to ask that in the supercritical regime, what will be the impact of *spreading-out* the offspring distribution on p_{ext}, while keeping the mean constant. Informally, we expect that a species whose reproduction distribution fluctuates more than that of another species is also more likely to die out, even if on average an individual in either species produces the same number of offspring. To confirm this suspicion, we examine the extinction probabilities of two Galton-Watson processes whose offspring distributions are convexly ordered. Convex order captures the idea of *spreading-out* of distributions (more completely than the comparison of the variances alone, for example). We define it as follows.

Definition 5.1 *Given two random variables, X and Y, X is said to be convexly smaller than Y, and we write $X \leq_{cx} Y$, if for every convex function f such that $\mathbb{E}[f(X)]$ and $\mathbb{E}[f(Y)]$ are finite, we have*

$$\mathbb{E}[f(X)] \leq \mathbb{E}[f(Y)].$$

If the above inequality holds only for every increasing and convex function, then the corresponding order is called increasing convex order, or icx order and we write $X \leq_{icx} Y$. Analogously we define decreasing convex order(dcx).

Remark 4 $X \leq_{cx} Y \Rightarrow \mathbb{E}[X] = \mathbb{E}[Y]$ (Taking f, by turns, to be the identity function and its negative).

We can now answer our question about the impact of spreading-out the offspring distribution on the extinction probability.

Theorem 5.2 *Consider two supercritical Galton-Watson processes with offspring distributions, Z_1 and Z_2, that are convexly ordered:*

$$Z_1 \leq_{cx} Z_2.$$

Then their extinction probabilities, p^1_{ext} and p^2_{ext} resp., are ordered too:

$$p^1_{ext} \leq p^2_{ext}.$$

Proof. From the characterization of p_{ext} by the smallest solution of equation(1.4), all we need to show is that for all $s < 1$,

$$\mathbb{E}(s^{Z_1}) \leq \mathbb{E}(s^{Z_2}). \tag{5.1}$$

But for all $s > 0$, $\phi_s : x \to s^x$ is a convex function. Therefore, by the definition of convex order, we have (1.4).

Remark 5 If we progressively make the offspring distribution of a supercritical Galton-Watson process to be convexly smaller, it will approach the extreme case when the offspring distribution is deterministic. In that case, p_{ext} is evidently 0, which is consistent with our result.

5.1.2 Convex Order on Sequences of Finite Random Graphs and Implications in Configuration Model

Having studied the impact of convex ordering in the simplest case of Galton-Watson branching processes, we would like to study its impact in more general random graph models. In particular, the sequences of finite uniformly rooted random graphs $[G(n) : n \in \mathbb{N}]$ which converge in the local weak sense present an interesting context where convex ordering might prove useful in comparing the asymptotic properties.

Definition 5.3 *We call two uniformly rooted random graphs on n nodes, $G_1(n)$ and $G_2(n)$, to be convexly ordered if the distributions of the corresponding root degrees, $D^r_1(n)$ and $D^r_2(n)$ are convexly ordered,*

$$G_1(n) \leq_{cx} G_2(n) \qquad \text{if} \qquad D^r_1(n) \leq_{cx} D^r_2(n).$$

Extending the definition to two sequences of uniformly rooted random graphs,

$$[G_1(n) : n \in \mathbb{N}] \leq_{cx} [G_2(n) : n \in \mathbb{N}] \qquad \textit{if} \qquad G_1(n) \leq_{cx} G_2(n) \textit{ for all } n \in \mathbb{N}.$$

The asymptotic property that we will study on $[G(n) : n \in \mathbb{N}]$ is the percolation probability, which we define as follows,

Definition 5.4 *Let $G_c(n)$ denote the connected component containing the uniformly chosen root in $G(n)$, and $|G_c(n)|$ its size. We define the percolation probability in $[G(n) : n \in \mathbb{N}]$ by*

$$\lim_{k\to\infty} \liminf_{n\to\infty} \mathbb{P}\{|G_c(n)| > k\} =: \theta(G). \tag{5.2}$$

In a number of random graph models, $G_c(n)$ converges in a local weak sense to a branching process. In such cases, the percolation probability of the model relates in a very simple way to the extinction probability of the limiting branching process.

Theorem 5.5 *Suppose $G_c(n)$ converges in a local weak sense to a branching process, LBP. Let $p_{ext}(LBP)$ denote the extinction probability of LBP and as before, $\theta(G)$ be the percolation probability of $[G(n) : n \in \mathbb{N}]$. Then,*

$$\theta(G) = 1 - p_{ext}(LBP). \tag{5.3}$$

Proof. By the definition of local weak convergence, for fixed k,

$$\begin{aligned}\liminf_{n\to\infty} \mathbb{P}\{|G_c(n)| > k\} &= \lim_{n\to\infty} \mathbb{P}\{|G_c(n)| > k\} \\ &= \mathbb{P}\{|LBP| > k\}.\end{aligned}$$

Therefore, (5.2) simplifies to

$$\begin{aligned}\theta(G) &= \lim_{k\to\infty} \mathbb{P}\{|LBP| > k\} \\ &= \mathbb{P}\{|LBP| = \infty\} \\ &= 1 - p_{ext}(LBP).\end{aligned}$$

where the second step is due to the monotone convergence theorem.

Since we've already seen some consequences of convex ordering on simple branching processes, we focus our attention on those random graph models where we have some kind of convergence to a branching process. Recall from Section 1.4 that arguably the simplest of such models is the sparse Erdős-Rényi model, $[ER(n,p); np = \lambda : n \in \mathbb{N}]$ which converges in the local weak sense to the λ-Poisson Galton-Watson process (i.e., branching process whose offspring distribution is Poisson with parameter λ). Recall that $ER(n,p)$ is a random graph on n vertices, in which a pair of vertices is connected by an edge with probability p, independently of all other pairs. It is evident that we cannot put two Erdős-Rényi models in convex order, because λ determines the distribution of the random root degree. But remark that we cannot put the offspring distribution of the limiting Poisson Galton-Watson process in convex order either for the same reason.

This leads us to study the impact of convex ordering on the generalization of Erdős-Rényi model introduced in Section 1.3: configuration model , which we will denote in this chapter by $[G^*(n) : n \in \mathbb{N}]$.

Recall from Section 1.4 that if we pick a root in this graph uniformly, the connected component containing the random root, $G_c^*(n)$, converges in local weak sense to a branching process. This limiting branching process is the same as the approximating branching process described in Section 1.3.1. We recall its construction: The founding father has a random number of children, according to the same distribution as that of the asymptotic degree distribution, D. But from the first generation onwards, each individual, independently of its siblings, produces offspring according to the distribution of V which is given by

$$\mathbb{E}[f(V)] = \mathbb{E}[Df(D-1)]/\mathbb{E}(D) \tag{5.4}$$

for all bounded,continuous functions, f.

We refer to this limiting two-stage branching process as *size-biased* branching process, and denote it by SB.

Remark 6 Because of the similarity of the limiting branching process of the configuration model to the Galton-Watson process, we might naively expect that the convex ordering of configuration models would lead to an ordering of their percolation probabilities. Unfortunately, this seems not to be true in general. To see why, consider two configuration models, $[G_1^*(n) : n \in \mathbb{N}] \leq_{cx} [G_2^*(n) : n \in \mathbb{N}]$, i.e., $D_1 \leq_{cx} D_2$. We work with the extinction probability of the limiting branching process since it is related to the percolation probability of the configuration model by equation(5.3). Denote this extiction probability for the first model by $\tilde{\eta}_{D_1}$, and the extiction probability for the Galton-Watson process with size-biased distribution V_1

by η_{V_1}. The notation for the second model is analogous. Then,

$$\tilde{\eta}_{D_1} = \mathbb{P}(D_1 = 0) + \sum_{k=1}^{\infty} \mathbb{P}(D_1 = k)(\eta_{V_1})^k = \mathbb{E}[(\eta_{V_1})^{D_1}]. \tag{5.5}$$

Now by the convex ordering of D_1 and D_2, we could have concluded that $\tilde{\eta}_{D_1} \leq \tilde{\eta}_{D_2}$ if we had $\eta_{V_1} \leq \eta_{V_2}$. To prove this, we try, as in theorem 5.2, to prove that $\mathbb{E}(s^{V_1}) \leq \mathbb{E}(s^{V_2})$ for all $s > 0$. For this, we need

$$\frac{\mathbb{E}[D_1 s^{D_1 - 1}]}{\mathbb{E}[D_1]} \leq \frac{\mathbb{E}[D_2 s^{D_2 - 1}]}{\mathbb{E}[D_2]}$$

or,

$$\mathbb{E}[D_1 s^{D_1 - 1}] \leq \mathbb{E}[D_2 s^{D_2 - 1}].$$

But the function $\xi_s : x \to x s^{x-1}$ is not a convex function, so we cannot obtain the above inequality by just the convex ordering of D_1 and D_2.

To summarise, convex ordering of D_1 and D_2 implies the icx ordering of V_1 and V_2, but the function $\xi_s : x \to s^x$ that we need for the comparison of extinction probabilities is decreasing and convex for $s \in (0, 1)$.

Even if we are not able to conclude anything about the ordering of percolation probabilities from the convex ordering of configuration models, we can say something about the *percolation threshold*. We do not explicitly define percolation threshold here, but the idea should be clear from the following theorem.

Theorem 5.6 *Consider two convexly ordered configuration models,*
$[G_1^*(n) : n \in \mathbb{N}] \leq_{cx} [G_2^*(n) : n \in \mathbb{N}]$*, i.e.,* $D_1 \leq_{cx} D_2$
then,

$$\theta(G_2^*) = 0 \;\Rightarrow\; \theta(G_1^*) = 0. \tag{5.6}$$

Proof. We have

$$\begin{aligned} D_1 \leq_{cx} D_2 &\Rightarrow \mathbb{E}[{D_1}^2] \leq \mathbb{E}[{D_2}^2] \text{ and } \mathbb{E}[D_1] = \mathbb{E}[D_2] \\ &\Rightarrow \frac{\mathbb{E}[{D_1}^2] - \mathbb{E}[D_1]}{\mathbb{E}[D_1]} \leq \frac{\mathbb{E}[{D_2}^2] - \mathbb{E}[D_2]}{\mathbb{E}[D_2]} \\ &\Rightarrow \mathbb{E}[V_1] \leq \mathbb{E}[V_2]. \end{aligned}$$

From the above inequality, equation (5.3) and theorem 1.1,

$$\begin{aligned}\theta(G_2^*) = 0 &\Rightarrow \eta_{V_2} = 1 \\ &\Rightarrow \mathbb{E}[V_2] \leq 1 \\ &\Rightarrow \mathbb{E}[V_1] \leq 1 \\ &\Rightarrow \eta_{V_1} = 1 \\ &\Rightarrow \theta(G_1^*) = 0.\end{aligned}$$

This result gives an insight into the emperical result of Newman (2002) [47], where a corelation was introduced between the degree distributions of different vertices. Their main result could be framed in terms of the *supermodular order*, another *convex-like order*, on two random graph models. Then, the result would imply that a random graph model lower in supermodular order has a higher percolation threshold. This suggests that a *directionally convex ordering* on random graph models can be defined which would order the percolation thresholds in the same way as theorem 5.6.

We leave the further analysis for future work.

References

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

[2] Aldous, D. and Steele, M. (2004). The objective method: probabilistic com- binatorial optimization and local weak convergence. In *Probability on discrete structures*. vol. 110 of *Encyclopaedia Math. Sci.* Springer, Berlin pp. 1–72. 15

[3] Alon, N. and Spencer, J. H. (2000). *The Probabilistic Method*. Wiley, New York. 5

[4] Amini, H., Draief, M. and Lelarge, M. (2009). Marketing in a random network. *Proc. Network Control & Optimization* **LNCS 5425,** 17–25. 20

[5] Amini, H. and Lelarge, M. (2015). The diameter of weighted random graphs. *Ann. Appl. Probab.* **25,** 1686–1727. 14, 70

[6] Bailey, N. (1975). *The Mathematical Theory of Infectious Diseases*. Books on cognate subjects. Griffin. 53

[7] Banerjee, A., Chandrasekhar, A. G., Duflo, E. and Jackson, M. O. (2013). The diffusion of microfinance. *Science* **341,**. 53

[8] Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics* **107,** 797–817. 53

[9] Barbour, A. D. and Reinert, G. (2013). Approximating the epidemic curve. *Electron. J. Probab* **18,** 1–30. 21

[10] Bender, E. A. and Canfield, E. R. (1978). The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory* 24:296–307. 7

[11] Bhamidi, S., van der Hofstad, R. and Hooghiemstra, G. (2010). First passage percolation on random graphs with finite mean degrees. *Ann. Appl. Probab.* **20,** 1907–1965. 15, 70

[12] Błaszczyszyn, B. and Yogeshwaran, D. (2009). Directionally convex ordering of random measures, shot-noise fields and some applications to wireless networks. *Adv. Appl. Probab.* **41,** 623–646. xi, 85

[13] Błaszczyszyn, B. and Gaurav, K. Viral marketing on configuration model. arxiv 1309.5779 2013. submitted. ix

[14] Błaszczyszyn, B. and Yogeshwaran, D. (2013). Clustering and percolation of point processes. *Electron. J. Probab.* **18,** 1–20. xi, 85

[15] Bollobás, B. (2001). *Random graphs* vol. 73. Cambridge university press. 7

[16] Bollobás, B. (1981). Degree sequences of random graphs. *Discrete Mathematics* **33,** 1 – 19. 4

[17] Bollobás, B. (1984). The evolution of random graphs. *Transactions of the American Mathematical Society* 286(1):257–274. 5

[18] Britton, T., Janson, S., Martin-Löf, A. et al. (2007). Graphs with specified degree distributions, simple epidemics, and local vaccination strategies. *Advances in Applied Probability* **39,** 922–948. 12, 13, 28

[19] Cheng, J., Adamic, L., Dow, P. A., Kleinberg, J. M. and Leskovec, J. (2014). Can cascades be predicted? In *Proceedings of the 23rd international conference on World wide web*. International World Wide Web Conferences Steering Committee. pp. 925–936. 21

[20] Comets, F., Delarue, F. and Schott, R. (2014). Information Transmission under Random Emission Constraints. *Combinatorics, Probability and Computing* **23,** 973–1009. ix, 20

[21] Cooper, C. and Frieze, A. (2004). The Size of the Largest Strongly Connected Component of a Random Digraph with a Given Degree Sequence. *Combinatorics, Probability and Computing* **13,** 319–337. 20, 27

[22] Coupechoux, E. and Lelarge, M. (2014). How clustering affects epidemics in random networks. *Adv. in Appl. Probab.* **46,** 985–1008. 11, 20

[23] Coupechoux, E. and Lelarge, M. (2015). Contagions in random networks with overlapping communities. *Adv. in Appl. Probab.* **47,** 973–988. 11, 20

[24] Dembo, A., Montanari, A. et al. (2010). Gibbs measures and phase transitions on sparse random graphs. *Brazilian Journal of Probability and Statistics* **24,** 137–211. 15

[25] Domingos, P. and Richardson, M. (2002). Mining the network value of customers. In *In Proceedings of the Seventh International Conference on Knowledge Discovery and Data Mining*. ACM Press. pp. 57–66. 53

[26] Draief, M. and Massoulié, L. (2010). *Epidemics and Rumors in Complex Networks* vol. 369 of *London Mathematical Society Lecture Notes*. Cambridge University Press. 5

[27] Erdös, P. and Rényi, A. (1959). On random graphs, I. *Publicationes Mathematicae (Debrecen)* **6,** 290–297. 3

[28] Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. In *PUBLICATION OF THE MATHEMATICAL INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES*. pp. 17–61. 3, 4

[29] Fountoulakis, N. (2007). Percolation on sparse random graphs with given degree sequence. *Internet Mathematics* **4,** 329–356. x, 12, 18, 20, 67

[30] Fountoulakis, N. and Panagiotou, K. (2010). Rumor spreading on random regular graphs and expanders. In *APPROX-RANDOM*. ed. M. J. Serna, R. Shaltiel, K. Jansen, and J. D. P. Rolim. vol. 6302 of *Lecture Notes in Computer Science*. Springer. pp. 560–573. 12, 13

[31] Fountoulakis, N. and Panagiotou, K. (2013). Rumor spreading on random regular graphs and expanders. *Random Structures & Algorithms* **43,** 201–220. 20

[32] Fountoulakis, N. and Reed, B. A. (2009). A general critical condition for the emergence of a giant component in random graphs with given degrees. *Electronic Notes in Discrete Mathematics* 34:639–645. 11

[33] Gaurav, K., Błaszczyszyn, B. and Keeler, P. (2014). Pioneers of influence propagation in social networks. In *Proc. of CSoNet, collocated with COCOON*. see also http://arxiv.org/abs/1310.2441. x

[34] Gilbert, E. N. (1959). Random graphs. *Ann. Math. Stat* **30,** 1141–1144. 65

[35] Janson, S. (2009). On percolation in random graphs with given vertex degrees. *Electron. J. Probab.* **14,** no. 5, 86–118. 12, 13, 18, 20

[36] Janson, S. (2014). The probability that a random multigraph is simple. ii. *J. Appl. Probab.* **51A,** 123–137. 9

[37] Janson, S., Knuth, D. E., Pittel, B. and et al. The birth of the giant component 1993. 4

[38] Janson, S. and Luczak, M. J. (2008). A new approach to the giant component problem. *Random Structures and Algorithms* **34,** 197–216. ix, 8, 10, 11, 17, 19, 20, 27, 28, 29, 38

[39] Janson, S. and Spencer, J. (2007). A point process describing the component sizes in the critical window of the random graph evolution. *Combinatorics, Probability and Computing* 16(4):631–658. 5

[40] Kang, M. and g. Seierstad, T. (2008). The critical phase for random graphs with a given degree sequence. *Combinatorics, Probability and Computing* 17:67–86. 11

[41] Karp, R. M. (1990). The transitive closure of a random digraph. *Random Structures and Algorithms* **1,** 73–93. 20, 27

[42] Kempe, D., Kleinberg, J. and Tardos, E. (2003). Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. KDD '03. ACM, New York, NY, USA. pp. 137–146. 53

[43] Lelarge, M. (2012). Diffusion and cascading behavior in random networks. *Games and Economic Behavior* **75,** 752 – 775. 20

[44] Lindvall, T. (1992). *Lectures on the Coupling Method*. Wiley, New York. x, 5, 67, 71

[45] Łuczak, T. (1990). Component behavior near the critical point of the random graph process. *Random Structures & Algorithms* 1:287–310. 5

[46] Łuczak, T., Pittel, B. and Wierman, J. C. (1994). The structure of a random graph at the point of the phase transition. *Transactions of the American Mathematical Society* **341,** 721–748. 5

[47] M. E. J. Newman (2002). Assortative mixing in networks. *Phys. Rev. Lett.* **89,**. 91

[48] Molloy, M. and Reed, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms* **6,** 161–180. 8, 9, 17, 19, 20

[49] Molloy, M. and Reed, B. (1998). The size of the giant component of a random graph with a given degree sequence. *Comb. Probab. Comput.* **7,** 295–305. 9, 10, 11, 17, 19, 20

[50] Newman, M. E. J. (2002). Spread of epidemic disease on networks. *Phys. Rev. E* **66,** 016128. 53

[51] Penrose, M. D. (1997). The longest edge of the random minimal spanning tree. *The Annals of Applied Probability* **7,** 340–361. x, 66, 67

[52] Sushil Bikhchandani, D. H. and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy* **100,** 992–1026. 53

[53] T.E.Harris (1963). *The theory of branching processes*. Springer, Berlin. 3

[54] Van Der Hofstad, R. (2014). *Random graphs and complex networks*. Available on `http://www.win.tue.nl/rhofstad/NotesRGCN.pdf`. x, 13, 20, 67, 68, 81